Economics of Scaling Agricultural Research Recommendations to Up-Scale Adoption and Impact

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Maxwell Mkondiwa

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Terrance Hurley, Co-Advisor: Philip G. Pardey

June 2019

Acknowledgements

This work and my life as a PhD student wouldn't be possible without the generous support of time and resources from many people. If I were to write all about them, they would fill out the entire pages of the dissertation. A book, a biography perhaps, will be written about that experience. Let me start by thanking my co-advisors: Phil Pardey and Terry Hurley. Phil basically believed in me without ever meeting me, he trusted I can do research when he was not obligated to. That level of generosity guided me throughout my studies. In times when I fell short on my studies, he still trusted me and offered his encouragement. This was surely a student-advisor matching made in heaven! Terry offered many ideas starting with my early MS work and building on to this work. His open-door policy and gracious help with whatever ideas I had along the way gave me the confidence to pursue the ideas further. I acknowledge the wonderful support of the other committee members, Jeff Apland and Jeff Coulter. I thank Jeff Apland for his excellent help with GAMS programming for chapter 4. The paper started as a small project in his well-organized mathematical optimization class and through his continued advice will likely be a long-term policy model for Malawi. When I was hiding because I couldn't get things done on the model as expected, he found a way of getting me back up. Jeff Coulter read early drafts of the papers and gave me advice that improved the papers. I admire his responsiveness and thoughtfulness to my questions.

I wish to thank fellow graduate students, faculty and staff in the department of applied economics. I feel particularly blessed to have met amazing friends, too many to list here (my attempt at a list is biased towards those currently in the program). Thanks to Ibrahim Keita (past, present and future president of Mali; when I was stuck cashless at a hotel in D.C., he literary saved me from homelessness!), Sinafikeh Germessa (fellow Malawian economist by association), Silver Namunene (Achimwene), Divine Ndifongwa (passed micro prelims because of you), Khadidja (best friend forever, remember that), Julieth Santamaria (I knew you before you were/will be famous!), Celestine, Mario Renato, Ashwini Shankar, Tiesta Thakur, Haseeb Ali, Aditi Kadam, Harshada Karnik and the list goes on and on. Some faculty (beyond teaching) particularly helped me on research. I thank Jay Coggins, amazing and dedicated teacher for teaching me the most classes (4!), and three years of working with a group of us (Ibrahim, Ian and Sina) on a geoengineering paper (we need to get the paper out!). Theorem 1 in chapter 3 derives from those conversations. Thanks to the following faculty: Jason Kerwin (fellow Malawian by association), Terry Roe (amazing stories and chats about Africa's growth), Rodney Smith (for reminding me to think about general equilibrium implications all the time), Dick Levins (my soccer agent), Metin Cakir (conversations on chapter 3 and for organizing food group meetings), Hikaru Peterson (for organizing food group meetings where I presented all the essays), Marc Bellemare (for comments on the early versions of ideas in chapter 2), Ford Runge (introducing me to a family where I had my thanksgiving dinners) and anonymous second year paper faculty reviewer. I thank Dipankar Bandyopadhyay and Lin Zhang, biostatistics professors who taught me Bayesian statistics that I have used in chapter 2 and 3. My gratitude to the Heydingers (Beverly, Rick, John and Erin) and their relations for the thanksgiving dinners.

Special thanks to colleagues in InSTePP-my immediate family while in Minnesota. These include Michelle Hallaway (basically baby-sitted me, couldn't have survived entire PhD life without her administrative support), Ian Luby (a great 4 years office mate! Watched and chatted champions league more than working), Yuan Chai (many hotpots and read chapter 2), Robert Andrade (more soccer after this), Ali Joglekar (my mentor), Xudong Rao (we need to write that paper), Huichun Sun, Kyuseon (Kristy) Lee ,Devon Phillips, Senait Senay, Jan Greyling and Mason Hurley.

Soccer has been part of this dissertation. I thank fellow soccer players in Unreal Madrid (all time champs, oldest players! Great core group of Rodrigo, Giovanni and Jose), Oaxaca (three years of 1 hour drive every Sunday to the field with a coach who spoke Spanish only but we chatted throughout!), Northrop (I lost 4 teeth in the end but was worth the great games and friendships), and Bunda Alumni F.C (always great group of guys). I also thank many colleagues at the pickup games.

I sincerely thank the McKnight Foundation for fully funding both my MS and PhD studies from 2013 to 2019.Whenever I saw any homeless American, I was humbled by the great privilege of being generously fed, housed, and educated through the funding. Hopefully, this will pay off beyond my inner circle. I specifically wish to thank several people within the ranks of McKnight Foundation who personally helped me. Ric Coe, Rebecca Nelson, Charlie Riches and Jane Maland Cady offered much of their time and dozens of emails to ensure that I get funding and a better university for my graduate studies. Ric Coe basically chose me as his protégé in research methodology and brought the idea for the potential for funding for PhD studies even without me requesting it. Rebecca Nelson suggested UMN (a decision I will forever cherish) and encouraged me to be more assertive with my research ideas.

I thank my family for the patience with my “loud” silence over the years. I thank my family members in Canada for epic holidays. Thanks to Aunt Cecilia for the care and for coming for my commencement. I thank my dad (beyond being a father he is my academic hero for making me skip two secondary school grades), maternal aunts (Mrs. F. Masimbe, another academic hero and mum for many chats on the ideas in this dissertation and Mrs. Chitsundi, mum of me and young brother for teaching me important virtues of life). Thanks to my sister Msekiwa Matsimbe for paying graduate school application fees and all the care whenever I get back to Malawi. Thanks to all my brothers, sisters, and cousins for being there always. To my grandmother, I had to finish this on time because you kept track of the so many years of schooling. Being the first PhD in family wasn’t easy but hopefully many more will find it easy in the future.

Dedication

To those who contributed immensely to my life but are not here to experience my progress: mum (Jane Sagawa), granddad (Clyton Mkondiwa), brother (Clyton Mkondiwa) and mentors (Dr. Jason Beddow and Dr. George Mataya).

Abstract

A fundamental challenge of agricultural development in sub-Saharan Africa (SSA) is that technologies which prove successful at a small scale, in limited locations, and with few farmers, often fail to scale to encompass the preponderance of poor farmers. This study focuses on the economics of deploying technologies and recommendations that are then scaled beyond their initial targeted groups. The dissertation is composed of three essays. In the first essay, we address the stylized fact that experimental crop responses are typically higher than observational crop responses obtained in farmers' fields. This is arguably a canonical example of a failure to scale from experimental plots. To close these crop response gaps—necessary goal assuming general constant long-term trends in maize output/fertilizer price ratios—, we propose that fertilizer recommendations be based on a Bayesian combination of experimental and observational crop response estimates. We use Bayesian econometric methods to combine estimates from experimental and observational evidence. In the second essay, we build on the first to determine the likelihood that farmers will adopt new varietal technologies. We modify the differentiated product demand models used in the industrial organization literature to the economics of hybrid maize varietal adoption in Malawi. By focusing on the characteristics space of maize varieties, our approach can help in ex-ante evaluation of the scaling-up potential of new crop varieties. The final essay calibrates inter-district food flows in Malawi thereby providing statistics for improving the targeting of national and regional food policies and technology commercialization strategies. We develop a food sector model for Malawi and use it to analyze the impacts of varying transport costs on food traded among districts within the country.

Table of contents

List of Tables

List of Figures

1. Introduction

This dissertation focuses on econometric and mathematical models for scaling up technologies or recommendations to improve their widespread adoption and development impacts. In the first essay, we propose that recommendations on fertilizer application rates and evaluations be based on a Bayesian combination of experimental and observational evidence to ensure that the recommendations can work at scale. We illustrate the approach with an application to spatially variable maize yield responses throughout Malawi. In the second essay, we propose a characteristics space analysis of quality-differentiated agricultural inputs—in this case alternative maize seed varieties—that allows analysis of the willingness to pay for various characteristics, thus revealing the demand for particular new crop varieties. In the final essay, we develop a food sector model that calibrates food flows across districts within Malawi thereby providing important information for making targeted food policy and technology commercialization decisions.

Essay 1: A stylized fact of African agriculture is that crop responses to inorganic fertilizer application derived from experimental studies are often substantially greater than those from observational studies (e.g., farm surveys and administrative data). The divergence between experimental and observational crop responses reported in the literature, coupled with the recent debates on the relative costs and benefits of the expensive farm input subsidy programs in Africa, have raised the importance of reconciling these estimates. This essay argues that progress on closing the gaps has been impeded by focusing only on mean crop response differences while ignoring the enormous uncertainty and heterogeneity arising from variations in soil types, environmental conditions, and management practices. We show in this essay that a novel way of dealing with different crop response function estimates is to use a Bayesian approach that combines information from both experimental and observational data to model uncertainty and heterogeneity in crop yield responses. Our Bayesian approach has the advantage over conventional approaches to assessing (spatially variable) yield responses in that it addresses the potential concern that the experimental process tends to overstate observational crop response while observational methods tend to be muddled with behavioral (or crop management) factors. Using nationally representative experimental, survey, and administrative datasets from Malawi, we find that: (1) crop responses to inorganic fertilizer use are uniformly low in observational data, (2) there are large heterogeneities in yield responses across space, and (3) ignoring parameter uncertainty and spatial heterogeneity in crop responses can lead to questionable policy prescriptions.

Essay 2: There are always at least two sides to every story and an economic story of the adoption of agricultural technologies throughout sub-Saharan Africa (SSA) is no exception. One familiar story highlights the "acceptance problem," namely that fundamental constraints in remoteness, weak markets, inappropriate policies, low education, cultural factors and many other related constraints are key drivers to the low rates of technology adoption that prevail in many parts of SSA. The other story focuses on the "availability problem," which maintains that many newly available technologies fail to provide any relative advantage in terms of their performance related attributes compared with the other (often status quo) alternatives. This essay proposes a pure characteristics space analysis for both the acceptance and availability problems. We illustrate this model using an application of the adoption of maize varieties in Malawi, a rapidly changing differentiated input market. We find that farmers are willing to pay more for complex traits like drought tolerance and flint texture than yield differentials per se. Our results (and the analytical approach we develop) have direct implications for maize breeding programs in Malawi.

Essay 3: This essay develops a spatially-explicit, mathematical-programming model for the Malawian food sector to calibrate inter-district food flows and to assess how transport cost variations affect these flows. Data on inter-district commodity trade flows are typically not collected and are thus unavailable for most sub-Saharan African (SSA) countries and for many parts of world. However, access to such data would present opportunities for smarter and better targeted development policies that allow for the spatial spillover of interventions targeted to a specific locale. The food sector modeling approach we develop and implement allows for a natural estimation of inter-district trade flows in data sparse environments where the lack of such data preclude estimation of intra-national gravity trade models. Our modeling method is consistent with a modified von Thünen "arrows" approach in which transport costs determine the quantities and types of inputs or outputs that flow across (spatially) “separated” but not “isolated” districts. The calibration results for our baseline model indicate that about 7% of Malawian maize production flows among districts as compared with more than 40% for rice, beans and groundnuts, and 0% for cassava and potatoes. A simulation experiment of varying unit transport costs shows that reductions in per unit transport costs nonlinearly increase the share of production that is traded inter-regionally, although the traded shares vary among the crops included in our model.

2. Closing the Gaps in Experimental and Observational Crop Response Estimates: A Bayesian Approach

The best fertilizer on any farm is the footsteps of the owner.

(Taken from Scott 1998, p. 284, attributed to Confucius)

2.1. Introduction

A stylized fact of African agriculture is that experimentally derived crop responses to inorganic fertilizer application are often substantially greater than those obtained from observational studies (e.g., using farm survey or official administrative data). There is also a long history of description of this yield gap, which can be reduced to the presence or absence of positive or negative confounding factors such as biologically optimal crop management by researchers versus biologically sub-optimal (albeit possibly optimal bio-economic) management by farmers; smaller, more uniform, plot sizes used in experiments versus larger and heterogeneous plot sizes used by farmers; biases or spatial inconsistencies in site or sample selection of scientific versus farmer plots; and observer bias (see, for example, Coe et al. 2016; Snapp et al. 2014). In addition to these factors, sample sizes (or the number of replicates) also differ, with farm surveys that can span thousands of households versus experiments that often include a few hundred sites at most. According to Bullock and Bullock (2000, p.97), "…the simple fact is that most agronomic experiments are not run for enough years and enough locations to obtain many different observations of weather and possible field characteristics". Because a field experiment at a few locations cannot capture all this nuance in variation, the representativeness of agronomic experiments is often questioned because crop response estimates and recommendations derived from them are different from what is experienced under farmers' conditions.

Economists have had long-standing debates on both the causes and solutions related to yield gaps.[1] Some of the early work on this topic includes Davidson and Martin (1965 and 1967), Dillon and Anderson (1995), and Anderson (1992). This study contributes to this

[1] See Beddow et al. (2014) for a detailed bio-economic review and evaluation of the yield gap literature.

prior literature by using experimental and observational evidence in Malawi to characterize the crop response gaps and proposes Bayesian linear and hierarchical models to combine the estimates from observational and experimental studies.

The discrepancy in experimental versus commercial yield response can have profound policy implications. For example, a recent study by Jayne et al. (2015) using observational crop responses of the social benefits versus costs of the Malawian farm input subsidy program—one of the largest targeted national farm input subsidy programs in Africa—found it to be unduly costly.[2] In direct contrast, using experimental crop response data, Chirwa and Dorward (2013) found the program to be economically beneficial relative to its costs. According to Jayne et al. (2015), the use efficiency of the nitrogen applied to maize is perhaps the most important factor determining the benefits of the Malawi farm input subsidy program. The crux of their case largely hinges on the following yield response relativities:

> "These (crop response) estimates (3.4 – 9.9kg of maize output per unit of fertilizer applied per ha) are based on farm survey data and not researcher-influenced plots, and they reflect the range of management practices and production constraints found within Malawi's smallholder farm sector… Unfortunately, Dorward and Chirwa (2015) maize response estimates of 16 – 18 kg are derived from researcher-influenced farm trials undertaken in the late 1990s with participants who were largely master farmers" (Jayne et al., 2015, p. 746)

While noting the challenges of reconciling these crop responses, Arndt et al. (2016) evaluated the subsidy program using crop responses ranging from 11.8 to 18.5 kg of maize per ha per kg of nitrogen fertilizer. They settled on this range of responses, more or less arbitrarily, to cover reported rates from the observational and experimental evidence they

[2] Jayne et al. (2018) reviews subsidy programs for 10 African countries and reports that between 2011 and 2014 the farm input subsidy program in Malawi accounted for 21 to 44 percent of the country's total spending on agriculture.

reviewed[3]. Arndt and co-authors further comment that reconciling the experimental and observational crop responses remains an important and unresolved problem. The difficulty of fully reconciling the estimates from these multiple sources of data, which were collected at different time periods in different locations, with different varieties and using different research methods, is exacerbated by the reliance on mean response comparisons that completely ignore the substantive spatial and temporal heterogeneity in these responses.

Given these challenges, policy making is usually left to guesswork regarding the true crop responses and a reliance on arbitrary approaches to re-adjusting experimental yield responses to better reflect farmer conditions. In this chapter, we propose a simple and replicable method of bringing all these subjective judgements into a formal estimation framework. The approach is based on the Bayesian paradigm of combining prior information and observational data. We use a Bayesian hierarchical model to incorporate both parameter uncertainty and heterogeneity in crop response functions and fertilizer recommendations. The Bayesian approach of combining different evidence on the same phenomenon has recently been used by Fessler and Kasy (2018) to combine predictions of labor demand and wage inequality derived from economic theory and empirically derived estimates, by Meager (2019) to combine results from various randomized control trials of micro-credit interventions across countries, and by Rosas et al. (2018) to impose duality theory restrictions based on experimental trial data to assess market level crop yield responses to prices in the U.S. In macroeconomic forecasting, the idea of combining evidence using the Bayesian paradigm has been implemented in Dynamic Stochastic General Equilibrium Modeling (DSGE) as well as New Keynesian Macroeconomics. It has also been used in the economics of education literature to combine teacher value added measures that are precise but biased with alternative measures based on admission lotteries for students that are unbiased but imprecise (Angrist et al. 2017).

[3] Notice here that some of the crop response gaps are due to semantics. While much of what Jayne et al. (2015) refer to as crop responses are based on total fertilizer applied, Dorward and Chirwa (2015) specifically refer to nitrogen applied. This is a common cause of ostensible differences in yield responses given that agronomic experiments typically report nutrient specific responses while observational evidence often report crop output responses to the total fertilizer applied.

This chapter contributes to the crop responses literature by exploring the possibility of improving soil fertility recommendations through the careful combination of experimental and observational crop response evidence, while also taking into account parameter uncertainty and heterogeneity. Specifically, the study analyzes the effect of observational crop responses when conditioned on prior (experimental) crop responses. The applications from this modeling approach are many, especially given a lack of directly comparable experimental data over time due to changes in experimental designs. Using our proposed approach, researchers can simply use previous estimates as priors in a new analysis. Similarly, in many agronomic research projects (e.g., the Drought Tolerant Maize for Africa project by CIMMYT), scientists are asked to conduct household surveys prior to or while conducting experiments as part of learning the environment. With this approach they can formally use the household survey estimates as priors in their experimental analysis. The key rationale is that Bayesian estimates weight the estimate from the present and prior data using an inverse of the variance parameters so that the uncertainty of the parameters determine whether the prior or present data dominate.

This chapter makes three main contributions. The first is that we incorporate parameter uncertainty in single output and multi-output crop response function estimation, which provides a more complete description of the crop response parameters. Specifically, we contribute to the on-going debates on the use of experimental versus observational mean crop responses by showing that using the mean response function in combination with arbitrary adjustments may result in suboptimal policy prescriptions in most cases because the inherent unobserved heterogeneity within and across farms requires site-specific optimization. Instead, researchers are likely better off using the entire distribution of the parameters (as this distribution contains more information than the mean), which entails comparing the distributions of benefit-cost ratios and profits obtained from the different alternatives being studied. Second, we illustrated the empirical implications of considering the spatial heterogeneity of crop responses among districts within Malawi. Third, we propose and validate the use of Bayesian recommendations that take into account parameter uncertainty and heterogeneity.

Although endogeneity issues from measurement error, simultaneity, and omitted variables require close consideration when estimating crop response functions (and such concerns bedevil all prior crop response assessments that use observational data), the use of district-level fixed effects allows comparisons of within district differences of the sources of evidence while accounting for uncertainty of parameter estimates, and heterogeneity of crop response estimates. Furthermore, this chapter follows a partial identification strategy to test if the crop response parameter is observationally equivalent under various prior specifications.

Debates on whether crop response estimates are low or high in Malawi and other African countries are difficult if not impossible to resolve when uncertainty and heterogeneity of the estimates across time and space is ignored. Most importantly, the results we obtain below show that even with an extremely high prior mean yield response (e.g., 30kg/ha of maize output per kg of fertilizer applied) and level of precision (e.g., a value of 10, which is equivalence to a variance of 0.1), the posterior crop response estimates can only go as high as 20kg of maize output per ha for a unit of nitrogen (N) fertilizer applied. In addition, the lowest is around 2kg of maize output per kg of fertilizer. This implies that there is a 95% probability that the mean crop responses are between 2-20kg of maize output per kilogram of N fertilizer applied[4]. Further analysis in the chapter shows that there is huge spatial heterogeneity in the crop responses, which should be of importance in policy design because some of the districts are non-responsive to fertilizer application.

This evidence therefore suggests that resolving policy debates that depend on crop responses should consider variances and heterogeneity in these responses. In summary, the results presented below illustrate that Malawian maize yield responses are generally low and highly variable (over time and space). This underscores the need for evidence-based targeting of locations and beneficiaries if farm input subsidy programs such as that presently operating in Malawi are to constitute a cost-effective public policy and be profitable for smallholder farmers.

[4] The interpretation holds because we are using a Bayesian credible interval not a frequentist confidence interval.

2.2. Model

2.2.1 Theoretical model

The standard neoclassical approach to production economics on crop response to inputs like fertilizer is a primal approach based on deterministic profit maximization. Following Hartley (1983), the deterministic conditional neoclassical model of fertilizer usage and output response, given that a positive area of land has been allocated to crop j, assumes that farmers maximize profits with respect to all variable input levels associated with the area of land a_{ij} which is usually normalized to a unit hectare. The profit associated with crop j in each plot i is defined as

$$\pi_{ij} = p_{ij}y_{ij} - w_{ij}x_{ij} - FC_{ij} \text{ subject to: } y_{ij} = f_{ij}\left(x_{ij}, a_{ij}, z_{ij};\ \theta_{ij}\right) \qquad \textbf{(1)}$$

Where π_{ij} is the profit per unit (hectare) for each plot i and crop j. p_{ij} and w_{ij} are prices of crop outputs and fertilizer respectively. y_{ij} is the crop specific yield (kg/ha) and $y_{ij} = f_{ij}(x_{ij}, a_{ij}, z_{ij};\ \theta_{ij})$ describes the production technology where x_{ij} is the quantity of fertilizer applied (kg/ha), a_{ij} represents area under crop j in plot i, z_{ij} represents the quantity of other inputs like labor, and θ_{ij} represents the set of relevant response parameters that are usually estimated from the data. FC_{ij} represents fixed costs.

Under assumptions of twice continuously differentiability, convexity of the production possibilities set, strict concavity of the objective function, the economic condition for optimality is

$$\begin{gathered} \frac{\boldsymbol{\partial}\pi_{ij}}{\boldsymbol{\partial}x_{ij}} = p_{ij}\frac{\partial f_{ij}\left(x_{ij}{}^{*}, a_{ij}, z_{ij};\ \theta_{ij}\right)}{\partial x_{ij}} - w_{ij} = 0 \\ \frac{\partial f_{ij}\left(x_{ij}{}^{*}, a_{ij}, z_{ij};\ \theta_{ij}\right)}{\partial x_{ij}} = \frac{w_{ij}}{p_{ij}} \end{gathered} \qquad \textbf{(2)}$$

Using the implicit function theorem or assuming conventional functional forms for $f_{ij}\left(x_{ij}{}^{*}, a_{ij}, z_{ij};\ \theta_{ij}\right)$, it is easy to find the optimal $x_{ij}{}^{*}$ and this approach has been used extensively in practice to make fertilizer use recommendations.

There are fundamental flaws using the neoclassical production model. First, the parameter θ_{ij}, which essentially drives the optimality as well as heterogeneity across farms, is assumed to be known and certain such that it is usually not included in the optimization. But these parameters are rarely if ever known (either in an agronomical or statistical sense), which implies that economic decisions made on the basis of this assumption are suspect. The fundamental problem in agricultural settings is that the crop is typically grown on soils with an "inherent soil fertility gradient," which implies that yields even under heavily controlled environments will be uncertain because of the unobserved heterogeneity in the soil even a few centimeters apart (Zingore et al. 2003).[5]

In a statistical sense, θ_{ij} is usually a set of unknown parameters, about which farmers may have some prior information based on the performance of the same or different crops under similar or different input regimes. In conventional theory, there is no provision to incorporate this prior. Second, this conventional approach does not provide any direction as to what type of data would be required to estimate the production relation $f_{ij}\left(x_{ij}{}^{*}, a_{ij}, z_{ij};\ \theta_{ij}\right)$. A researcher can conduct experiments to decipher some $x_{ij}{}^{*}$, but no known experimental design can comprehensively investigate the effect of each of the x_{ij} on yield while also controlling for all other effects in x_{ij} and z_{ij}. Occasionally, multi-factorial experiments are conducted to (partially) address this challenge. Another line of research uses farm surveys to analyze the observable determinants of yield. Under this approach, the farmer has already made a set of input and crop management choices depending on their observed and unobserved circumstances. Using these different approaches result in different estimates of θ_{ij}.

This study proposes an extensive prior robust Bayesian analysis to investigate combined estimates of θ_{ij} that are consistent with theory and the practical challenges (and the relative prices) faced by farmers.

[5] In reality it is not just soil quality that varies with location. Many other climate, terrain, and physical aspect factors vary by location, often in ways that are imprecisely if at all measured, and variations in these factors also affect crop yields.

2.2.1.1 Are Bayesian recommendations the most profitable?

In this section, we show analytically and numerically whether Bayesian recommendations are better than relying solely on experimental or observational based recommendations. Consider, two profit functions, one derived using experimental evidence while the other using observational evidence,

$$\pi^e = \max \int py(\beta_{experiment}, x) - wx$$

$$\pi^s = \max \int py(\beta_{survey}, x) - wx$$

Now consider another possibility in which the results are based on a Bayesian combination of the evidence. Following Carlin and Louis (2009), we can think of a simple case where μ is the prior mean of crop responses, y is the likelihood mean of crop responses, τ^2 is the variance of the prior and σ^2 be the variance of the likelihood, and $\sigma^{2^*} = \sigma^2/n$. The normal-normal conjugate posterior distribution is given by

$$p(\theta|y) = N(\theta| \frac{\sigma^{2^*}\mu + \tau^2 y}{\sigma^{2^*} + \tau^2}, \frac{\sigma^{2^*}\tau^2}{\sigma^{2^*} + \tau^2})$$

Letting $\omega = \frac{\sigma^{2^*}}{\sigma^{2^*} + \tau^2} \in [0,1]$, then the posterior mean is a weighted average, $\omega\mu + (1-\omega)y$ and posterior variance is given by $var(\theta|y) = B\tau^2 \equiv (1-B)\sigma^{2^*}$ and is smaller than $\tau^2$and σ^2. The advantage of the posterior is that precision just as information is additive, $var(\theta|y)^{-1} = var(y|\theta) + var\ (\theta)$.

Without loss of generality, we can write the Bayes posterior mean of crop responses using the notation from above in a similar way,

$$\beta_{\ bayes} = \omega\beta_{experiment} + (1-\omega)\beta_{survey}$$

$$\pi^B(p, w, \beta) = \max \int py(\beta_{\ bayes}, x) - wx$$

A profit function is convex in p, w and therefore $\pi^B \geq \pi^s$ and $\pi^B \geq \pi^e$. We can also represent this relationship using a value-to-cost ratio (VCR) as $\frac{py'(\beta_{\ bayes}, x)}{w} \geq$

$\frac{py'(\beta\ survey, x)}{w}$. This implies that we can simply compare the distribution of the VCR at the economically optimal Bayesian recommendations and the distribution of VCR at the economically optimal survey/experimental recommendations. For a quadratic crop response function, this is calculated as $VCR - EOBR = \frac{p\left(\beta_1^{Bayes} + \beta_2^{Bayes} x^{*Bayes}\right)}{w}$, where *EOBR* is the economically optimal Bayesian recommendation. We provide more details of the approach in the appendices.

2.2.2.2 Stochastic dominance comparisons

To make comparisons across different scenarios we use first order stochastic dominance, in particular "posterior stochastic dominance." Thus, different information sources are being combined probabilistically and stochastic dominance is being used to compare among them. Stochastic dominance is normally defined with respect to stochastic outcomes, which in the case of this study are profits. The study therefore concentrates on whether the fertilizer response parameters dominate each other across the entire measured range of fertilizer use when the prior is updated with additional information. Definition 1 below provides a description of posterior stochastic dominance.

Definition 1: First Order Posterior Stochastic Dominance- Let F(π) and G(π) be two cumulative distributions of outcomes (for example profits) based on different experimental priors. Drawing on Levy's (2000, p.56) definition, the distribution of outcomes F(π) will first order stochastically dominate the distribution of outcomes G(π) if and only if F(π) is less than or equal to G(π) for every π and there is at least one π for which a strong inequality holds.

Using the definition of stochastic dominance and interpretation of posterior parameter estimates as consisting of a prior and a data-based likelihood, two important claims follow when interpreting the prior scenarios. The first claim, based in the mean responses, is that if the mean for a prior is greater than the mean of the likelihood holding the variance or precision parameter constant, then the resulting posterior parameter is greater than the mean of the likelihood. Second, if the variance for a prior is greater (i.e., has lower precision) than the variance of the likelihood assuming the same mean, then the variance for the posterior is greater than the variance of the likelihood.

The stochastic dominance ordering is therefore an empirical question that depends on the relative magnitudes of the prior mean and precision versus the mean and the uncertainty of the likelihood. To illustrate the concept of stochastic dominance in comparing the prior scenarios, Figure 2-1 demonstrates three hypothetical cumulative distribution functions; F(.), G(.) and Q(.). In the figure, F(.) first order stochastically dominates G(.) since F(.) < G(.) across the entire measured range of profits. Higher order levels of stochastic dominance can be used to compare F(.) and Q(.) or G(.) and Q(.) (see Levy 2000).

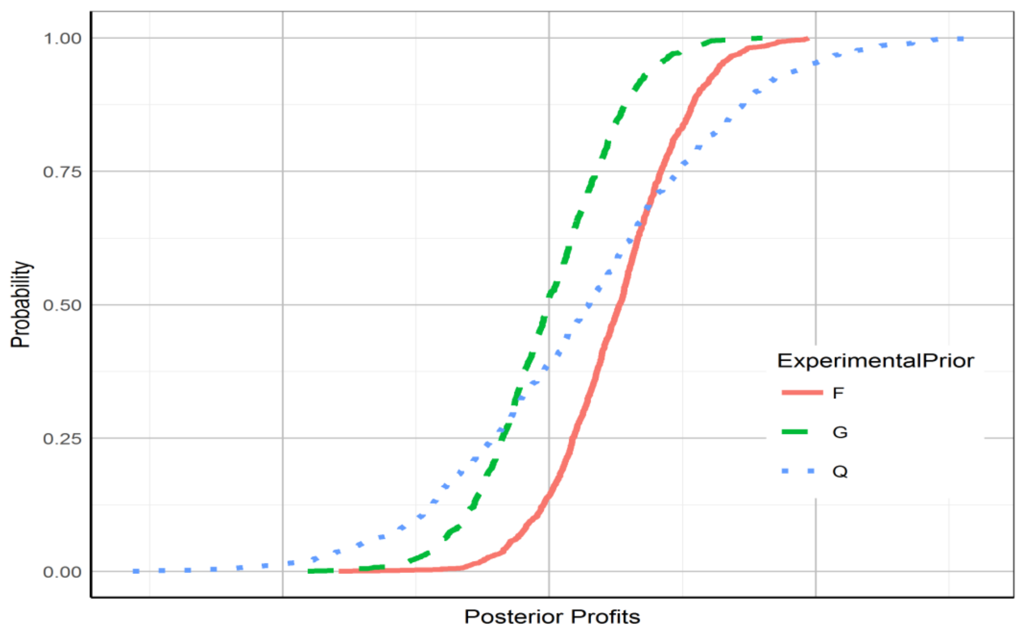


Figure 2-1:Stochastic dominance of hypothetical posterior outcomes given experimental priors

2.2.2 Empirical models

To incorporate the facets of the theory above, we use two estimation strategies, namely a Bayesian linear model and a Bayesian hierarchical model. All the models are quadratic in the crop response parameter.[6] The choice was made so that the results of the chapter are comparable with most of the experimental and observational estimates hitherto reported for Malawi, which have used this functional form (see, for example, Harou et al. 2017).

2.2.2.1 Estimating equation: Bayesian linear model

A Bayesian linear model is used to estimate the ray production function (see appendix A). The Bayesian linear model is equivalent to the ordinary least squares (OLS) regression model when a non-informative prior (e.g., zero mean and an arbitrarily large variance such as 10,000) is used. The Bayesian linear model for the ray production function approach is

$$\tilde{y}_i = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + \alpha z_{ij} + \sum_{j=1}^{J-1} \xi \lambda_{ij} + \epsilon_{ij} \qquad \textbf{(3)}$$

where $\tilde{y}_i$ is the output norm (hereafter referred to as total output index) defined as $\tilde{y} = \left[\sum_{j=1}^{p} y_j^2\right]^{0.5}$, and y_j is the yield (kg/ha) of crop j. When mono-cropped maize is considered, the total output index is equivalent to maize yield; $\mathrm{x_{ij}}, \mathrm{x_{ij}^2}$, $\mathrm{z_{ij}}$ and λ are vectors of nitrogen (N) fertilizer use (kg/ha), squared N fertilizer use, other explanatory variables (like seed use (kg/ha), rainfall etc.), and angular crop output coordinates (representing crop mix), respectively (see appendix A for details). The corresponding parameters are; β_1, β_2, α and ξ . Without loss of generality, we use the matrix notation for β to represent all the parameters and X the design matrix for all variables in the model in the derivations that follow. The disturbance term, ϵ has a multivariate normal distribution with mean 0 and covariance matrix $\sigma^2 I$, where I is an identity matrix, i.e., $\epsilon \sim^{iid} N(0, \sigma^2 I)$. In Bayesian econometric terminology, the variances σ^2 can be written as precision estimates, h where $h = \sigma^{-2}$ (Carlin and Louis 2015).[7] The multivariate normal density likelihood function is thus,

[6] There is a large body of literature that suggests that choice of a functional form affects the crop responses. The quadratic crop response is used in this paper because it is easy to estimate as compared with nonlinear options and also satisfies the basic properties of a production (or yield response) function.

[7] By definition, the posterior distribution function $\mathrm{p}(\theta|\mathrm{Y}, \mathrm{X})$ given crop yields and inputs is derived from a likelihood function $\mathrm{L}(\mathrm{Y}, \mathrm{X}|\theta)$ and a prior distribution function $\mathrm{p}(\theta)$ using the Bayes Rule as follows:

$$L(\tilde{y}\ |\beta, h) = \frac{h^{n/2}}{(2\pi)^{n/2}}\left\{\exp\left[-\frac{h}{2}(\|y\|\ - X\beta)'(\|y\|\ - X\beta)\right]\right\} \quad \textbf{(4)}$$

where n is the number of observations. This multivariate normal density can be written equivalently in terms of ordinary least squares (OLS) quantities (Chib 1995) where $v = n - k$, $\hat{\beta} = (X'X)^{-1}X'\tilde{y}$ and $s^2 = {(\tilde{y}\ - X\hat{\beta})'(\tilde{y}\ - X\hat{\beta})}/{v}$ to get:

$$L(\tilde{y}\ |\beta, h) = \frac{1}{(2\pi)^{n/2}}\left\{h^{n/2}\exp\left[-\frac{h}{2}(\beta - \hat{\beta})X'X(\beta - \hat{\beta})\right]\right\}\left\{h^{v/2}\exp\left[-\frac{hv}{2s^{-2}}\right]\right\}. \quad \textbf{(5)}$$

The normal-inverse gamma conjugate prior is used in which the prior for β is elicited conditional on h: $\beta|h \sim N(\underline{\beta}, h^{-1}\underline{V})$ where the arguments are prior for the mean estimate and its variance respectively. The prior for the model precision is, $h \sim G(\underline{s}^{-2}, \underline{v})$. The posterior is therefore a normal-inverse gamma, $NG(\underline{\beta}, h^{-1}\underline{V}, \underline{s}^{-2}, \underline{v})$:

$$p(\beta, h|\tilde{y}) \propto L(\tilde{y}|\beta, h) \times p(\beta) \times p(h). \quad \textbf{(6)}$$

Though exact sampling from the posterior is possible, the model was estimated using Markov Chain Monte Carlo using Gibbs Sampling. All the models were run with 11,000 MCMC iterations with 1,000 used as burn in and the remaining 10,000 for posterior analysis. Non-informative priors (0 prior mean and 0.001 prior precision) for the parameter estimates were assumed in the set of models presented in Table 4 in the appendix.

Selection of observables, partial identification strategy, and robust Bayesian analysis

The search for identification when estimating a production function—due to measurement errors and omitted variable bias—is an important area of research in production economics. Several strategies aimed at point identification of parameter estimates have been suggested

$p(\theta|Y, X) = \frac{L(Y,X|\theta)p(\theta)}{p(Y,X)}$,where $p(Y, X)$ is the normalizing constant which is a function of Y and X only, and can be defined by $p(Y, X) = \int_{\Omega} L(Y, X|\theta)p(\theta)$,where Ω is the parameter space. This implies that the posterior is proportional to $p(\theta|X, Y) \propto L(X, Y|\theta)\ p(\theta)$.

in the literature, including fixed effects estimation when using panel data or an instrumental variables approach. However, these strategies do not capture the uncertainty and unobserved heterogeneity associated with smallholder farming and the practical realities of observational data that is typically obtained via (farmer) recall. To achieve point identification, the ideal data set would be one with detailed data on day-to-day farm management practices, including certain details of prior crop management practices (e.g., crops grown before the present planting) and the farmers expectation of certain future factors (e.g., output prices, rainfall). In addition to these behavioral aspects, detailed soil quality, micro-scale precipitation and temperature measurements, as is done in precision agriculture, would be basic requirements. Such data are not available even for controlled experiments in most developing countries.

Unlike most of other micro-econometric work in which the reliance on randomized control trials to generate observational data is a comparatively recent phenomenon, crop response research by agronomists has relied on randomization for about a century. In the case of smallholder agriculture, it was long ago recognized that due to biased site selection and standardized management, achieving representativeness through randomization prevents analysis of factors that actually affect crop response in farmers' conditions (Coe et al. 2016). This chapter therefore argues that elusive attempts at point identification (i.e., estimating an unbiased causal effect) are inappropriate for crop response research when using currently available observational data on smallholder agriculture. Using district fixed effects in regressions for both experimental and observational data combined with a partial identification strategy, in which various assumptions are tested and a menu of estimates is assessed in terms of their plausibility, seems to be the most appropriate approach in this context. This chapter uses the prior robust Bayesian analysis for both sensitivity analysis and partial identification.

To make formal comparisons among the various yield response scenarios that we considered, we relied on the Bayes Factor. The Bayes Factor is the ratio of marginal likelihoods between two candidate models. Let model 1 be M_1 and model 2 be M_2. Their marginal likelihoods are simply the denominator of the Bayes rule

$$p(\tilde{y}|M_1) = \int p(\theta_1)L(\tilde{y}\ |\theta_1)\ d\theta_1 \text{ and} \tag{7}$$

$$p(\tilde{y}|M_2) = \int p(\theta_2)L(\tilde{y}\ |\theta_2)\ d\theta_2. \tag{8}$$

The Bayes Factor is therefore:

$$BF_{21} = \frac{p(\tilde{y}|M_2)}{p(\tilde{y}|M_1)}. \tag{9}$$

In complex models such as those being used in this study, the calculation of marginal likelihoods is infeasible. However, we used the Chib (1995) approach to approximate the marginal likelihoods. Since the aim is to maximize likelihood, a Bayes Factor (BF_{21}) of greater than unity implies that model 2 is a better (statistical) choice than model 1. The non-informative prior model (equivalent to using an ordinary least squares estimator) in Table A2 (see the appendix) is deemed improper, i.e., its marginal likelihood is not in closed form such that the Bayes Factors are undefined (Carlin and Louis 2015). The rest of the models which have informative priors are proper and are thus compared using Bayes Factors.

2.2.2.2 Heterogeneity in crop responses: Bayesian hierarchical model

The enormous heterogeneity in Malawi's smallholder farming systems implies that even crop response parameters that capture the uncertainty in associated model parameters may not be sufficient to characterize the different biophysical and socioeconomic circumstances faced by farmers. To address heterogeneity in the crop responses we deployed a Bayesian hierarchical modeling approach. According to Carlin and Louis (2015), a hierarchical modeling approach allows for a more explicit assessment of the heterogeneity both within and between groups. This modeling approach has been used extensively in the statistics and economics literature to model heterogeneity among individuals. For example, Cabrini et al. (2010) uses the Bayesian hierarchical approach to estimate market performance expectations (e.g., prospective prices) of individuals working in agricultural market advisory services. Chib and Carlin (1999) and Allenby and Rossi (1998) show how the hierarchical model can help in generating consumer and household specific parameters that are useful for marketers of consumer products.

Following Chib and Carlin (1999), consider the normal hierarchical model in matrix notation,

$$\tilde{y}_i = X_i\beta + W_i \ b_i + \epsilon_i \quad \textbf{(10)}$$

where each group i has k_i observations. The term "group" is being used generally here so that any type of heterogeneity may be considered. For instance, a group may constitute a location (region/district/village/agroecological zone), household, soil type or poverty status. X_i is $k_i \times p$ design matrix of p covariates. β is a corresponding $p \times 1$ vector of fixed effects. W_i is $k_i \times q$ design matrix. b_i is $q \times 1$ vector of subject-specific means and enable the model to capture marginal dependence among the observations on the i^{th} group. The group-specific random effects follow: $b_i \sim N_q\ (0, V_b)$. And the errors: $\epsilon_i \sim N(0, \sigma^2 I_{k_i})$. Assuming standard conjugate priors, $\beta \sim N_p\ (\mu_\beta, V_\beta)$ and $\sigma^2 \sim InvGamma\ (nu, \frac{1}{\delta})$ and $V_b \sim InvWishart\ (r, rR)$ where r is set to the number of parameters in the model and R is a diagonal matrix with values along the diagonal equal to the number of parameters (Chapman and Feit 2015). In their estimation, they used the MCMChregress function which implements the Gibbs sampling algorithm based on algorithm 2 in Chib and Carlin (1999).

2.3. Data sources and descriptive statistics

2.3.1 Data sources

The study uses both experimental and surveyed fertilizer response data for maize. In particular, the chapter uses evidence from a) the fertilizer verification experimental data collected and analyzed by the Malawi Maize Productivity Task Force in 1995/6-1997/8,[8] and b) the nationally representative Third Integrated Household Survey data, which were collected between 2010 and 2011 (and reflect production decisions for the 2008 and 2009 farming seasons) across all National Statistical Organization enumeration areas in Malawi.

---

[8] The author is indebted to Dr. Todd Benson at IFPRI who participated in the trials and kindly provided the data for the purposes of this study.

2.3.1.1 Experimental data

The study uses geo-referenced on-farm experimental data for the 1995/96 and 1997/8 growing seasons. The trials were carried out as experiments run on farmers' fields under the auspices of the Malawi Maize Productivity Task Force consisting of national and international experts. More than 1,500 trials were successfully implemented to evaluate six different inorganic fertilizer packages for hybrid maize grown by smallholders across the whole country (Government of Malawi 1997). The distribution of successful trials was unbalanced across the sites/regions and seasons due to statistical and administrative reasons. As reported in Table 2-1, all six treatments (A, B, C, D, E, F) were tested in the 1995/96 trials, while four (A, C, D, E) were tested in the 1997/98 trials. The structuring of treatments in the fertilizer trials suggests that the crop yield consequences of nitrogen and phosphorus may be confounding. That noted, agronomic studies on fertilizer use in Malawi (e.g., Government of Malawi, 1997) have argued that nitrogen is the most limiting macro-nutrient, and as such we focus on nitrogen responses.

In each of the two seasons, two hybrid maize varieties were planted; Malawi Hybrid 17 (MH17) was planted in upland sites with historically good rainfall conditions, and MH18 was supplied for trials in lowland areas and at those upland sites in rain-shadow areas. A few sites also tested composite varieties. The soil texture was recorded for each plot for each treatment plot per year, and a standard protocol was followed across all locations to ensure timely weeding, pest management, and other agronomic management activities. According to Benson et al. (1999, p. 12), one notable feature of the standardized protocol was to conduct the trials on farmer's field that had not received fertilizer or been planted to legumes in the previous two years. The plot size was 6.3m by 9m, consisting of seven ridges spaced 90cm apart. The net harvest plot size was five full ridge lengths, or 1/247ha (0.00405ha).

Table 2-1: Fertilizer treatments tested in 1995/96 and 1997/98

Treatment	**Nutrients**			**Fertilizer**	
Name Code	Nitrogen (kg/ha)	Phosphate (kg/ha)	Sulphur (kg/ ha)	Basal (50kg per ha)	Top dressing (50kg/ha)
A	0	0	0	0	0
B	35	0	0	0	1.5Urea
C	35	10	2	1 (23:21:0+4S)	1 Urea
D	69	21	4	2 (23:21:0+4S)	2 Urea
E	92	21	4	2 (23:21:0+4S)	3 Urea
F	96	40	0	1.75DAP	3.5Urea

Note: The nitrogen (phosphorus and sulphur) rates were computed based on the major nutrients composed in each of the basal and top dressing fertilizer. Consider for example for treatment C which required applying one 50kg bag of NPK or 23:21:0+4S and one 50kg of Urea. NPK has 23% of its composition in nitrogen while Urea has 46% of it composition in nitrogen. The total nitrogen applied for the C treatment is therefore $\mathrm{N} = 0.23 * 50 + 50 * 0.46 = 35$.

Table 2-2 includes descriptive statistics for the fertilizer trials in the two seasons. For each of the treatments, yields in 1995/96 were relatively higher than those obtained in 1997/98, reflecting less favorable weather during the 1997/98 season. In terms of treatments, the average yields were highest in treatment E, while the nil N treatment (Treatment A) had the lowest mean yield, which is expected considering that nitrogen fertilization is considered yield increasing, at least when moving from little or no N.

Table 2-2: Descriptive statistics of yields under different fertilizer treatments during 1995/96 and 1997/98 seasons

Season	Treatment	Mean	Median	Min	Max	Std.Dev	CV(Std.Dev/ Mean)*100
1995/96	A	1,410.47	1,261.18	0.00	7,245.40	873.26	61.91
1995/96	B	2,182.90	2,028.86	0.00	6,854.25	989.95	45.35
1995/96	C	2,358.06	2,284.75	182.78	8,577.87	985.00	41.77
1995/96	D	2,881.76	2,833.09	310.73	9,029.33	1020.50	35.41
1995/96	E	3,147.30	3,107.26	219.34	8,407.88	1086.49	34.52
1995/96	F	2,946.88	2,924.48	274.17	7,018.75	1079.05	36.62
1997/98	A	1,124.05	968.73	0.00	5,117.84	710.77	63.23
1997/98	C	1,996.54	1,919.19	109.67	5,940.35	927.44	46.45
1997/98	D	2,523.04	2,467.53	91.39	6,762.86	1029.17	40.79
1997/98	E	2,914.52	2,833.09	237.61	7,402.59	1157.81	39.73

Note: Total number of trials is 1,677 for 1995/96 and 1,407 for 1997/98.

A notable feature of the data summarized in this table, is the large variation in yield responses across each of the treatments. For both seasons, the coefficient of variation for the nil fertilizer treatment (A) are highest, with the lowest variation observed in treatments with the highest amount of nitrogen fertilizer applied (E and F). The variability observed can be attributed to interactions between fertilizer application and many other observed and unobserved factors including location, weather and topography. In this study, we explore the importance of understanding variability attributable to location effects.

2.3.1.2 Household survey data

The survey was conducted by the Malawi National Statistical Office. The data are analyzed at the crop-plot level to distinguish between input crop responses in single crop versus multi-crop farming systems. The observations pertain to rainy season plots that were owned and/or cultivated by the farm household and that were subject to Global Positioning System (GPS)-based land area measurement. The data files were merged first using the available plot geo-codes and then using household geo-variables (e.g., longitude, latitude, distance to road). The merging was done in a way that made sure that all the households in the final sample had consistent and identifiable household geo-coordinates. The geo-referenced data allow for the analysis of both agronomic and farmer behavioral responses. The use of this spatially explicit plot level data therefore implies that it is possible to estimate a structural model of multi-crop production enterprises (Fezzi and Bateman 2011). All plots not grown with either maize or a legume were excluded from the analysis. In the final data used for the analysis, there are 19,692 plot-crop observations for five key crops: maize, groundnuts, beans, pigeon peas and soybeans. This represents 70 percent of the plot crop observations in the data. These are the major crops for Malawi (accounting for 70 percent of the country's total cropped area in 2009-2013, Johnson (2016)) that are also featured in the integrated soil fertility management literature.

2.3.1.3 Administrative data

Administrative data were compiled from annual production estimates included in the Ministry of Agriculture and Food Security annual statistical bulletin for the period 1983-2015. These data are reported at the district level and consist of the total hectarage and production and average yield for each crop (i.e., maize, groundnuts, beans, pigeon peas and soybeans). This source does not report any fertilizer use data by district, and thus these administrative data were only used in calculating cross-district differentials in crop yield performance, a spatial dimension of yield gaps[9]. For maize, the data has varietal (local, composite and hybrid) specific yield, hectarage and production information.

2.3.2 Descriptive statistics

Table 2-3 presents selected descriptive statistics for the various variables characterizing the farm households and plots.

---

[9] The national per capita N fertilizer use in the current survey data is about 51 kg N/ha while in the period the experiments were conducted (1995/96 and 1997/98)it was about 38 kg of fertilizer perha (Minot et al. 2000, p.50). This implies that the observational application rates were between the 0 and 35 kg N/ha treatments in the experiments.

Table 2-3: Descriptive statistics for selected dependent and independent variables (n=19,692)

Variables	**Unit**	**Mean**	**Standard Deviation**
Dependent variables			
Euclidean norm of the yields	kg/ha	1,275.39	1,886.39
Maize yield	kg/ha	763.22	1264.7
Groundnut yield	kg/ha	163.6	1,122.86
Bean yield	kg/ha	45.98	397.13
Pigeon pea yield	kg/ha	107.49	501.61
Soybean yield	kg/ha	25.63	481.75
Maize dummy	Proportion	0.89	0.31
Groundnut dummy	Proportion	0.35	0.48
Bean dummy	Proportion	0.28	0.45
Pigeon pea dummy	Proportion	0.28	0.45
Soybean dummy	Proportion	0.29	0.45
Key independent variables			
Total inorganic fertilizer applied	kg/ha	162.75	205.98
Organic fertilizer use (Yes=1)	Proportion	0.12	0.33
Inorganic fertilizer use (Yes=1)	Proportion	0.69	0.46
Total N applied	Kg/ha	51.37	65.08

Note: The Euclidean norm of the crop output vector y is computed by $\tilde{y} = \left[\sum_{j=1}^{p} y_j^2\right]^{0.5}$

Table 2-3 shows that almost 90% of the plots in the sample were planted with maize followed by groundnuts (35%). The maize yields are within the range reported in most microeconomic studies. Inorganic fertilizer was used on almost 70% of the plots, while only 12% of the plots received organic fertilizers. The average fertilizer use is about 162kg/ha (corresponding to 51.37 kg N/ha), which is around the application rate reported for Malawi in other microeconomic studies.[10] This figure is higher than in other sub-Saharan African countries possibly because farmers are cultivating very small plots on especially small farms in the context of a generous farm input subsidy program.[11] Additional descriptive statistics are presented in Table A1 in the appendices. On average, the plots are 0.79 km away from the homestead, though with a huge variation across the sample (ranging from 0 to 10 km). The average plot size is 0.44ha. Most farmers perceive that their plots are either good (45%) or fair (43%) in response to a question about the perceived soil quality. Most of the plots (59%) have soils that are loam (i.e., between sand and clay) which are considered good soils for crop cultivation.

The majority of the households (75%) are male-headed with an average household size of 4.8 people. About 76 percent of the household heads have had no formal education. Almost 46 percent of these households are classified as poor, with average household incomes less than MK 37,002 per person per year based on the formal definition of the Malawi National Statistical Office (NSO). Most of the households live in remote rural areas, about 9km from a main road and 37km from the nearest trading center. Figure A1 in the appendices shows the number of plots planted with each of the crops. Most of the plots are planted with a pure stand of maize followed by a pigeon pea-maize intercrop.

2.3.3 Experimental and observational yield gaps

The challenge of combining observational and experimental evidence is that it is unlikely that one will find directly comparable treatments, that is, yield responses obtained using similar amounts (and types) of fertilizer grown in the same weather events and similar soil types. The nationally representative datasets available are almost 20 years apart (experiments in the 1990s and surveys in the 2010s). To demonstrate that experimental-

---

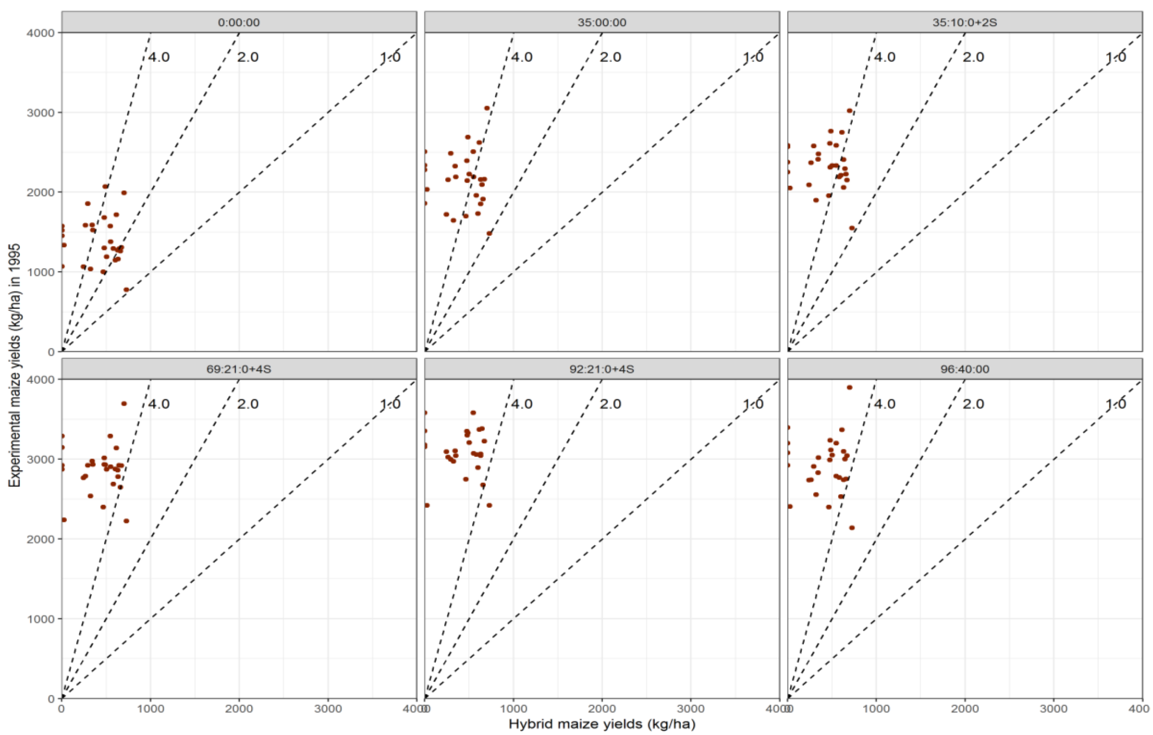
[10] For example, Sheahan and Barrett (2017) report averages of 146 kg/ha (which is equivalent to 53.1 kg/ha of nitrogen) for Malawi while Komarek et al (2017) reports 51 kg/ha nitrogen for central Malawi. The nitrogen application rate reported here, was derived by multiplying 0.23 to basal (23:21:0+4S) fertilizer amount applied and 0.46 to top dress (Urea) fertilizer applied, where 0.23 and 0.46 represent the proportion of nitrogen in the fertilizer.

[11] The subsidy program targets about 1.5 million farm households representing half of the farm households in Malawi with two 50 kg bags of fertilizer (Arndt et al. 2016).

observational yield gaps existed in the 1990s when the experiments were being conducted, we compared the district averages from the plot level experimental data with the corresponding hybrid varietal-specific administrative data for each of the two seasons, 1995/96 and 1997/98. Figure 2-2 shows scatterplots of district level averages of experimental hybrid maize yields for each of the fertilizer treatments (see Table 2-2) and the corresponding district averages of hybrid maize yields from administrative data in each of the respective seasons.[12]

There are six plots for the 1995/96 agricultural season and four plots for the 1997/98 agricultural season, with each of the plots representing the fertilizer treatments in the experimental evidence.[13] The rays indicate the ratio of experimental to observational yields.

A. 1995/96 agricultural season



[12] MH17 and MH18 were the main improved maize varieties planted during the years of the trial.

[13] While the experimental data were parsed into their respective fertilizer treatment cohorts, the same (albeit seasonal and varietal specific) observational data were used in each of the fertilizer cohorts.

B. 1997/98 agricultural season

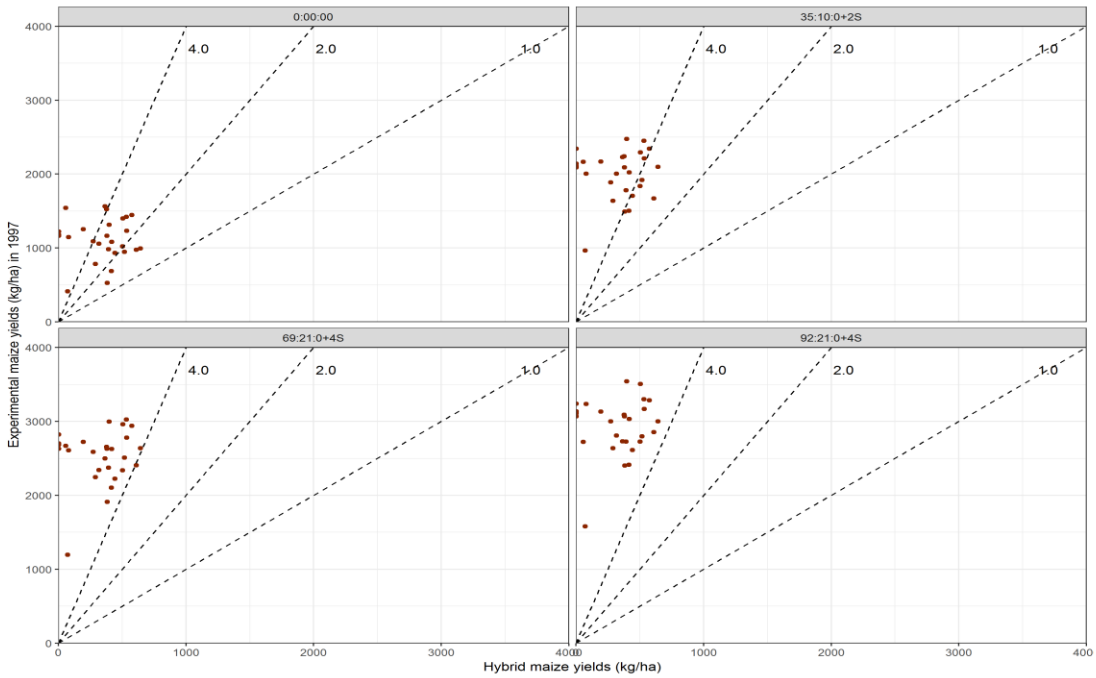


Figure 2-2: District level hybrid maize yields from experimental and administrative data for 1997/98 agricultural season

Across all the treatments, experimental yields are more than two times higher than the corresponding farm yields reported in the administrative database. As expected, the experimental-observational yield gaps increase as the amount of nitrogen applied in the experimental data increases. According to the Government of Malawi (1999), the 1997/98 agricultural season was a bad maize-growing year in that some districts experienced drought. This is especially evident in Figure 2B where the yield gaps for the no-fertilizer treatment are much lower. This highlights that the gap between observational and experimental maize yields are affected by environmental and climate conditions.

2.4. Results and discussion

2.4.1 Overview of the existing maize crop response literature for Malawi

The research on crop responses in Malawi dates back to at least the 1960s. Benson et al. (1998) reported estimates of experimental maize responses in studies conducted from the 1960s to 1998 ranging from 23.1 to 34 kilograms of maize per unit of additional applied nitrogen. Table 2-4 below taken from Arndt et al. (2016, supplemental material) reports the microeconomic evidence on the marginal returns to fertilizer use for selected types of maize seed. The mean maize responses range from 2.8 to 15 kg/ha for observational studies, much lower than the 23 to 34 kg/ha range reported in the experimental research. Informed by this prior evidence, below we use maize yield responses in the range of 0 to 30 kg/ha of maize for an additional kg of applied nitrogen as priors by which to anchor my assessment of estimates in prior studies.

Table 2-4: Marginal returns to nitrogen fertilizer use, by maize seed variety

	Dorward et al. (2008) (Survey of literature)	Harou et al. (2017) (Malawi field trials)	Chibwana et al. (2010) (Malawi FISP)	Ricker-Gilbert et al. (2011) (Malawi FISP)	Ricker-Gilbert and Jayne (2012) (Malawi FISP)
	Kilogram of maize yield for an additional unit of nitrogen				
Local varieties	10-12		12.0		
Composites	15				
Hybrids	18-20				
All improved varieties			9.6		
All maize seed	15	24-32			
Contemporaneous effect				6.1	
Enduring effect				11.7	
Measured at the 10th percentile					2.8
Measured at the median					7.6
Measured at the mean					9.0
Measured at the 90th percentile					9.7

Source: Adapted from Arndt et al. (2016).

In all the prior published assessments of both the experimental and observational maize yield response estimates for Malawi and sub-Saharan Africa, only mean responses were reported, absent any measures of the associated variation or uncertainty in these reported responses.[14] But to compare across studies and to make sense of these mean crop response parameter estimates, one cannot ignore the associated measures of precision.

[14] See a recent comprehensive review by Jayne and Rashid (2013, p. 533, Table 3).

2.4.2 Experimental, observational and Bayesian crop responses

In this section, we present the results from a Bayesian linear model (with results that are the same as using an ordinary least squares on equation 3). The set of results (see details in Appendix Tables A2 and A3) show the ray production functions for maize intercropped with either groundnuts, beans, pigeon peas or soybeans. The variables of interest in the production functions include N fertilizer, N fertilizer squared and coordinate angles, the latter representing the crop output mix. The coefficients for the polar coordinate angles are negative for all maize-legume combinations (see Table A3 in the appendices). This implies that an increase in the output mix reduces the total output index, meaning that the total output index is lower when maize is intercropped with a particular legume. We estimate that the mono-cropped maize response to N fertilizer application is about 10.56 kg/ha per kg of applied nitrogen, with a 95% credible interval of 9.78-11.36 kg/ha (see Table 2-5 and appendix Table A2). The experimental maize responses are about two times higher at 20.58 kg/ha per kg of applied nitrogen, consistent with finding of Anderson (1992) who observed that

> "There is a systematic overstatement of the extent of responsiveness of crops to applied fertilizer in Africa, relative to what is achievable under most farm conditions. The extent of overstatement is of the order of a factor of, say, two in terms of incremental response ratios." Anderson (1992, p. 393).

Table 2-5: Experimental and observational maize response function to nitrogen

	Experimental			**Observational**		
Parameter	2.50%	50%	97.50%	2.50%	50%	97.50%
(Intercept)	1026.8	1281.98	1535.99	181.52	1104	2040.73
N fertilizer amount	23.29	25.41	27.6	10.21	11.09	11.99
N fertilizer squared	-0.1	-0.09	-0.07	-0.01	-0.01	-0.01
Marginal Effect at N=55kg/ha	19.13	20.58	22.05	9.78	10.56	11.36

Note: Controls and district fixed effects are included in all specifications. For details, see Table A2 and Table A3 in the appendices. The marginal effects are calculated as $\beta_1 + \beta_2\bar{N}$, where β_1 and β_2 are estimated coefficients and $\bar{N}$ is the average nitrogen fertilizer evaluated at N = 55 kg/ha.

Given these results, we can combine the experimental coefficient and the observational coefficient by simply using the experimental estimate (25.41) and its standard deviation (0.41) as the prior in a regression using the observation data. Figure A3 shows that the resulting posterior distribution of the N coefficient (i.e., median: 12.01, 95% credible interval: 11.53, 12.50) is still closer to the distribution of N responses derived from the observational estimates (i.e., median: 11.09, 95% credible interval: 10.21, 11.99) that the distribution derived from the experimental results.

2.4.3 Robust Bayesian analysis with sensitivity testing

The foregoing analysis documents the crop response gaps and the hybrid crop responses when using particular experimental and observational data. It is justifiable to question the use of experimental data that were collected almost two decades before the observational data. A lot of biophysical factors (including varieties and soil quality) may have changed. Therefore, the following set of results uses a range of alternative priors that span the plausible range drawing on evidence gleaned from the prior published literature.

In particular, the Bayesian robustness results indicate changes in Bayesian estimates of the maize yield responses given changes in the prior distributions of crop responses at 55kg of nitrogen per ha. The robustness checks are in the changes to the prior on N fertilizer use on the mono-cropped maize response function. We considered a range of experimental crop response estimates reported for Malawi as summarized by Arndt et al. (2016) and Snapp et al. (2014) to assess if incorporating these priors leads to revisions in the crop responses that would warrant a change in the recommendations.[15] We considered six mean prior levels of the crop response coefficient; specifically values of 0, 6, 12, 18, 24, and 30.

A directed search for variance parameters across prior literature revealed that the estimates vary as well. For example, using different econometric specifications of a quadratic response function as we do, the standard error for the maize response to nitrogen ranges from about 0.3 to 0.5 in Harou et al. (2017). Using a quadratic production function, Darko

[15] The reported crop response rates are derived from different functional form specifications of the crop response functions (e.g., quadratic, linear), econometric procedures (e.g., ordinary least squares, quantile regression) and using different datasets (e.g., household surveys, demonstration trials).

(2016, p.92) estimated standard errors of the crop responses ranging from 1.6 to 3.2. In this study, we therefore consider three precision levels: 0.1, 1 and 10 corresponding to variances of 10, 1 and 0.1 respectively. Table 2-6 shows crop response quantiles for 18 different models for the various plausible mean and variance priors for the parameter corresponding to N fertilizer use. It is important to note that these are based on calculating marginal effects (which we also call crop response) not the direct coefficient of the N fertilizer. Marginal effects are calculated as $\beta_1 + \beta_2\bar{N}$ where β_1 and β_2 are coefficients for N and N squared terms, and $\bar{N}$ is the average nitrogen fertilizer rate at which the effect is evaluated at (i.e., 55kg N/ha).

The computed crop responses (Table 2-6, row 2 to row 13) are largely invariant to changes in the prior mean when the prior precision falls in the 0.1 and 1 range, but are indeed sensitive to the choice of prior means when a higher precision (10) is assumed. This is revealed by the extent of the overlapping 95% credible intervals when the different prior means are compared across the same low prior precision level (e.g., compare 5.92 upper quantile for 0 mean prior and 0.1 prior precision with 4.97 lower quantile for 30 mean prior and 0.1 prior precision). In terms of the effect of differences in the precision of the prior estimates, the table shows that when crop response posteriors at different prior precision are compared across the same prior mean, the pattern of crop response posterior results is fuzzy. The reason for this is that the prior precision of one parameter also affects the off-diagonal terms in the variance-covariance matrix, such that depending on the associations across all parameters in the regression, the Bayesian crop responses may not be an intuitive weighted average. In Bayesian terminology, this is called the effect of nuisance parameters. This also partly explains why the posterior estimates in the baseline case (with 0 prior mean and 0.001 for all parameters) are different from the sensitivity analyses (0 prior mean and 0.1 prior precision for all other variables except N and N squared terms).

2.4.2.1 Mean crop response scenarios

The scenarios in Table 2-6 are illustrated graphically in Figure 3 below using cumulative distribution functions (cdfs) of the crop response parameters subject to different priors. The comparisons of the cdfs can be interpreted as the posterior stochastic dominance. This is done by first holding the precision constant while varying the mean of the prior for the nitrogen coefficient (Figure 2-3A and 2-3B). In Figure 2-3A, as expected, holding precision

Table 2-6: Sensitivity analysis to various priors on the mean and precision

Row number	**Prior Precision**	**Prior Mean**	**Marginal effects of N fertilizer use**		
			2.50%	50%	97.50%
1	Baseline: 0.001	0	9.78	10.56	11.36
2	0.1	0	4.62	5.27	5.92
3	0.1	6	4.69	5.34	5.98
4	0.1	12	4.76	5.41	6.05
5	0.1	18	4.83	5.48	6.12
6	0.1	24	4.90	5.55	6.19
7	0.1	30	4.97	5.61	6.26
8	1	0	4.16	4.78	5.39
9	1	6	4.78	5.40	6.01
10	1	12	5.41	6.02	6.64
11	1	18	6.03	6.65	7.26
12	1	24	6.65	7.27	7.88
13	1	30	7.28	7.90	8.51
14	10	0	2.02	2.47	2.91
15	10	6	5.25	5.69	6.13
16	10	12	8.47	8.92	9.36
17	10	18	11.74	12.19	12.64
18	10	24	15.09	15.55	16.01
19	10	30	18.57	19.04	19.51

Note: The prior means for all the controls including N squared term were set to 0 and prior precision was set to 0.1 (so as to make the prior proper for the calculation of the marginal likelihood needed for Bayes Factor). Marginal effects are calculated as $\beta_1 + \beta_2\overline{N}$ where β_1 and β_2 are coefficients for N and N squared terms and $\overline{N}$ is the average nitrogen fertilizer the effect is evaluated at 55.kg/ha.

constant at the same low level, the model with the higher prior mean (30 maize kg/ha per kg of N) stochastically dominates all the other models. However, the differences between the posterior distributions of mean responses are small. At a higher precision level, the ordering of the mean response distributions is maintained, but now the differences between the posterior distributions are quite pronounced (Figure 2-3B). This implies that achieving lower variance (high precision) may be necessary in the development of recommendations as also suggested by recent agronomic research (Vanlauwe et al. 2016 and Coe et al. 2016). These results show that comparing mean experimental and observational estimates without considering the variation in responses around the mean can result in final combined recommendations that are a reflection of the scenario in Figure 2-3A or that in Figure 2-3B. Either way, the posterior responses are less than the experimental mean crop responses that are used to develop fertilizer application recommendations in Malawi.

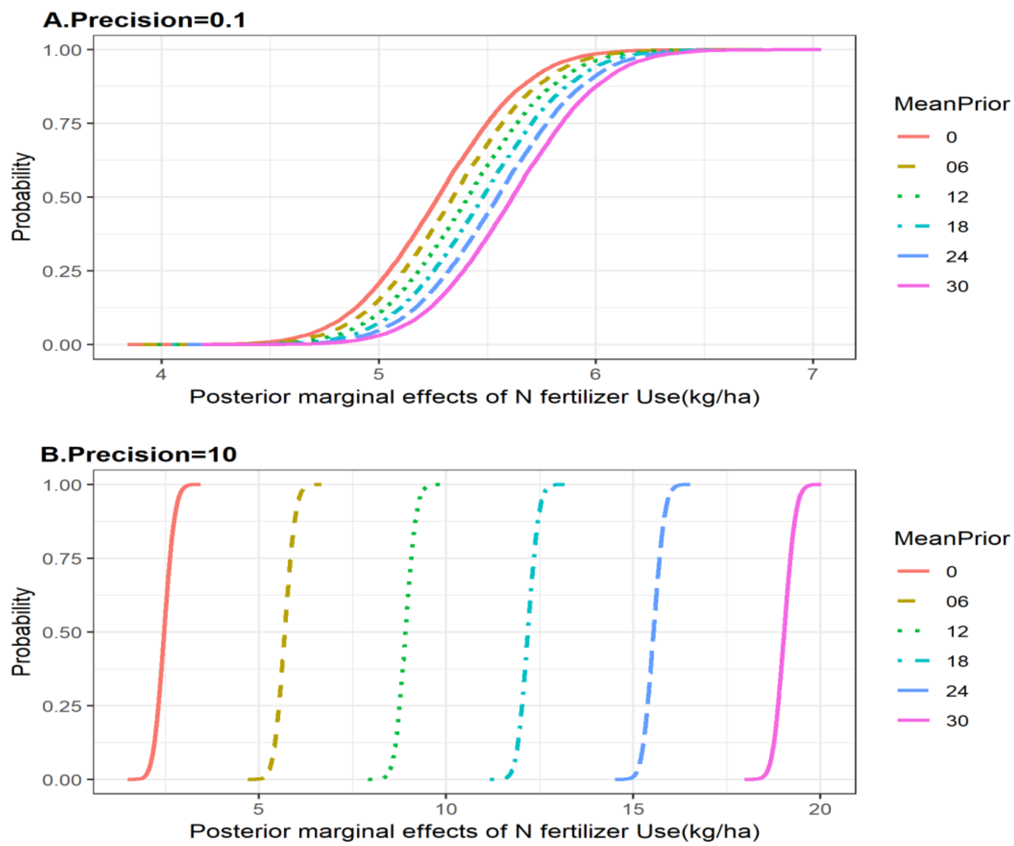


Figure 2-3: Posterior stochastic dominance at different prior mean values [prior precision = 0.1 and 10]

2.4.2.2 Precision of crop response scenarios

In Figure 2-4A and 2-4B, we show the cases where the precision priors are varied while the prior mean is fixed at some value. As expected from stochastic dominance, the higher the precision assumed for the prior, the more likely it is that the posterior will be similar to the prior distribution. Since the prior mean is 0 in figure 2-4A, the posterior parameter estimates for the model with a highly informative prior (i.e., a high precision prior = 10) is stochastically dominated by the ones with a weakly informative or effectively non informative priors (1 or 0.1). A contrasting case is presented in figure 2-4B where a prior mean of 30 is assumed. Here the model with a high precise prior mean stochastically dominates the models with low(er) precision.

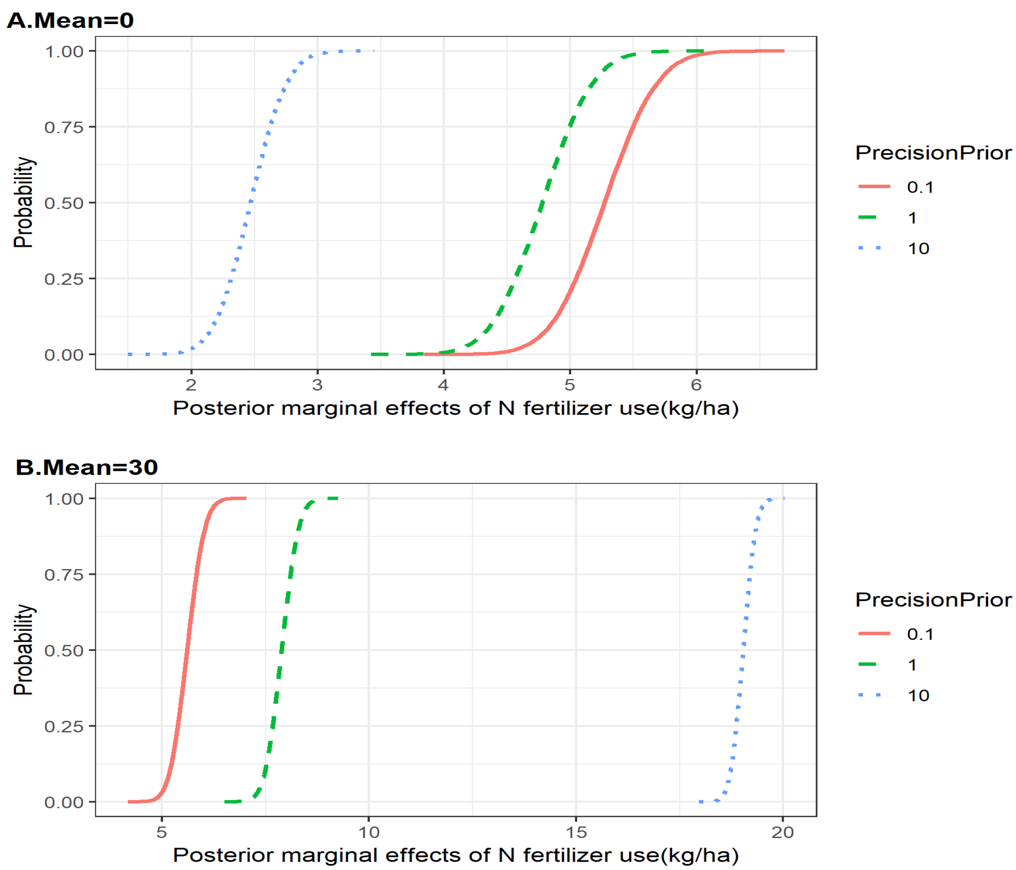


Figure 2-4: Posterior stochastic dominance at different prior precision values [prior mean = 30]

2.4.2.3 Combined mean and precision scenario

In the preceding stochastic dominance results, a clear and consistent ordering of either the prior mean or prior precision was assumed. But what of the case that has a higher prior mean but a lower precision relative to the opposite case (i.e., lower prior mean higher precision)? This is where Bayesian stochastic dominance becomes useful. Let's consider a case where the prior means and variances are different. In Figure 2-5, a low mean-high variance prior leads to posterior parameter estimates that stochastically dominate the high mean-low variance results. This implies that the debate between Jayne et al. (2015) and Dorward and Chirwa (2015) on whether a lower or higher mean crop yield response is appropriate for assessing the economic veracity of Malawi's farm input subsidy program is problematic when the precision (or variance) of the mean estimates are ignored.

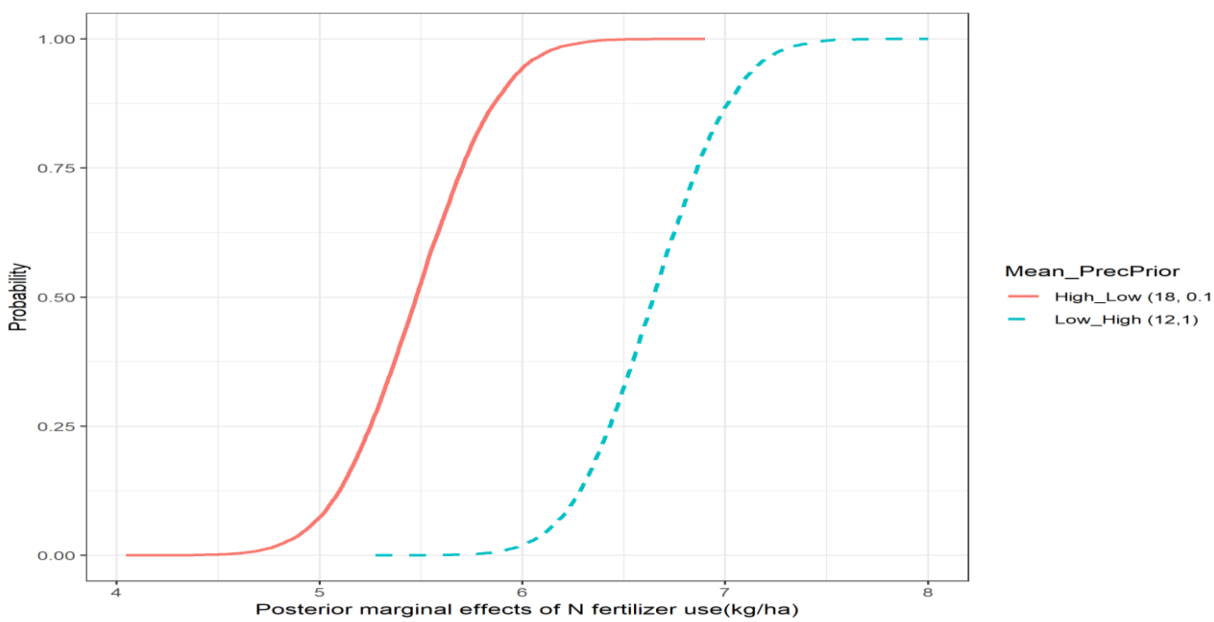


Figure 2-5: Posterior stochastic dominance at different experimental priors for mean and precision

When uncertainty is incorporated it is the case that for assumed precision levels of 0.1 to 1—which are typical of the precision levels in observational research—, the posterior crop response estimates that are likely relevant for commercial agriculture (observational) range from 4 to 9 kg of maize output for a unit of fertilizer per ha when the prior mean responses levels for experimental trials range from 0 to 30. While for precision levels of 1 to 10 which are prevalent in experimental research, the posterior crop estimates range from 9 to 19 for prior mean levels from 0 to 30. Based on our evidence, the likely fertilizer crop responses for Malawian agriculture are low and highly variable. Thus, any claims of substantial crop responses to fertilizer application in Malawian maize production are questionable. Therefore, when evaluating the efficacy of policies that depend on empirical estimates of crop responses, it would be advisable to err on the conservative side (and draw on all the plausible evidence about the mean responses and variations around this mean).

2.4.4 Heterogeneity in crop responses: Bayesian hierarchical model results

Beyond the question of variations around the mean crop responses, the crop response gaps may also be due to differences in locations where each of the studies were conducted within the country. We therefore need to understand the heterogeneity in the crop responses across locations. There are two extremes in the way heterogeneity is typically handled in econometric analysis. Most studies pool all the data and generate a single response parameter, assuming a homogenous response for the whole sample. At the other end of the spectrum, one may consider estimating the response parameters with specific (additive and multiplicative) fixed effects for individual cohorts of the data (e.g., individual districts), but this is generally inefficient due to data limitations. A Bayesian hierarchical modeling framework is an efficient (i.e., in terms of degrees of freedom) middle ground, which allows estimation of individual specific parameters as random parameters.

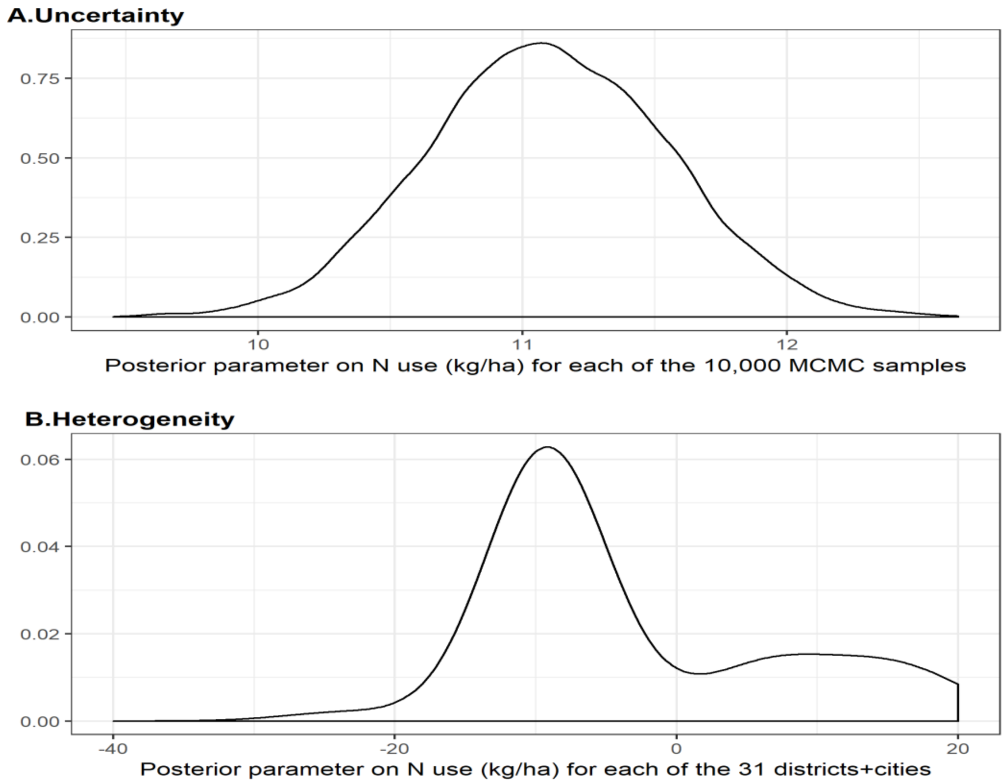


Figure 2-6: Uncertainty and heterogeneity in crop response parameter on fertilizer use (kg/ha)

Figure 2-6 shows the density plots of parameter uncertainty (panel A) and heterogeneity (panel B). Panel A is based on crop response parameter for the draws from the Bayesian linear model, and illustrates the parameter uncertainty under the maintained assumption of a spatially invariant response function. Panel B is based on a random parameter specification of the district-specific crop response parameter in a hierarchical Bayesian model, and represents the district-level heterogeneity in crop responses.

The results indicate that the model-based parameter uncertainty (ranging from about 8 to 14 additional kgs of maize for additional unit of fertilizer) plotted in Panel A is smaller than the district-to-district heterogeneity plotted in Panel B (-40 to 30 additional kgs of maize for additional unit of fertilizer). The negative responses imply that soils are not conditionally responsive to fertilizer application in these locations. While it is uncommon for agricultural economists and agronomists concerned with average responses to report negative responses, this can occur when the fertilizer applied scorches the seed, especially in relatively dry conditions (Vanlauwe et al. 2011). The findings on district-to-district heterogeneity may seem unrealistic when compared with previous analyses that assumed spatial homogeneity in responses. This result is nonetheless consistent with agronomic research that addresses individual plot heterogeneity. For example, a recent study by Vanlauwe et al. (2016) compared empirical distributions (heterogeneity) to model based distributions (measuring uncertainty) of crop responses from agronomic trial data related to maize in Western Kenya and beans in Eastern Rwanda. They concluded that model based distributions provide better precision in the extremes than empirical curves, but that model based distributions depend on the assumption that the model is unbiased. In terms of developing targeted crop response support, the heterogeneous model may be more appropriate as it can help identify districts and plots that are non-responsive.

Based on the observational data used in our analysis, the districts of Machinga, Nsanje and Chikwawa, for example, appear to be non-responsiveness to fertilizer application (see also Figure 2-7). This is in line with experimental evidence (Government of Malawi 1997) that reports lower crop responses in the shire valley districts (Nsanje and Chikwawa). This suggests a future research strategy that proceeds by answering two questions: 1) will maize in a given field respond to fertilizer; 2) if so, what is the optimum fertilizer rate? Answering question 2 is more difficult than answering question 1. At a minimum, being able to answer question 1 is really impactful. The hierarchical Bayesian model allows one to answer both questions in that we can identify unresponsive districts and the magnitude of the response for the responsive districts.

Figure 2-7 shows the scatterplot of the district level linear crop response parameters in a hierarchical model for experimental and survey data. For almost seven districts (specifically, Chikwawa, Mulanje, Machinga, Rumphi, Mangochi, Phalombe and Mulanje), the soils are not responsive to fertilizer application based on the observational evidence but are responsive in the experimental evidence.

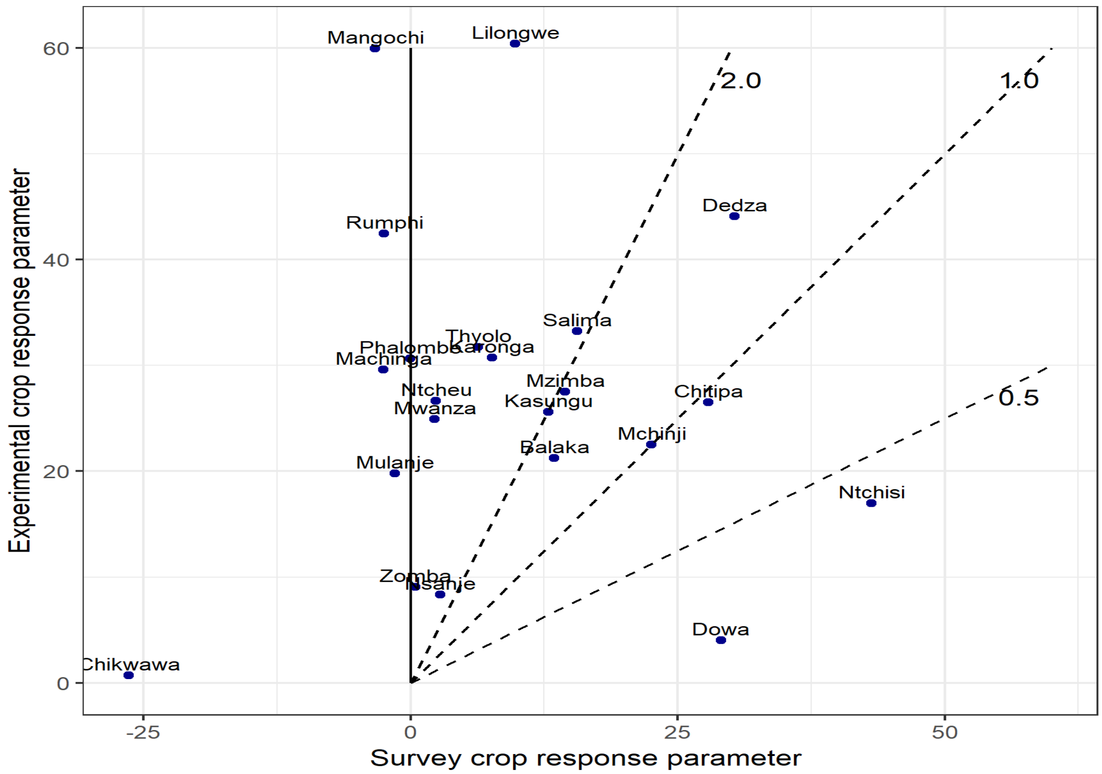


Figure 2-7: Scatterplot of hierarchical crop response coefficient by district from experimental and survey evidence.

Note: The x and y axes correspond to the hierarchical coefficient for linear term in a quadratic crop response function not the marginal effect.

As a research matter, this implies that there still more we need to learn about the biophysical and socio-economic aspects that distinguish these districts (i.e., water holding capacity or timing of fertilizer application). In terms of policy, it implies that well targeted extension services are required so that farmers do not waste fertilizer on unresponsive soils.

2.4.5 Limitations and future research

There are still several remaining limitations in addressing modeling challenges of parameter uncertainty, heterogeneity and disparate information sources in the estimation of crop responses to fertilizer application. The first limitation is that given the weaknesses of both experimental and observational studies, it is difficult to measure the quality of the subsequent posterior evidence.

The second limitation is that the scenarios on the effect of changes to the prior on the assumed posterior parameter estimates is based on the pooled model not the hierarchical model, which entails that heterogeneity is being treated separately from partial identification of the distribution of the parameters. This is inevitably the case because estimates treating fertilizer use parameters as random parameters across groups are not available. Other shrinkage models like empirical Bayes modeling and machine learning using ridge regression are potential candidates for future research. In addition, endogeneity concerns across both the experimental (due to self-selected master farmers) and observational (management bias and substantial measurement errors) evidence are areas of valid concern that future research could systematically address using quasi-experimental methods.

Finally, the models are not directly linked to any policy parameter like whether to subsidize fertilizer, which not only depend on the uncertainty and heterogeneity of crop response parameters, but also on other parameters (e.g., relative profitability of other crops) and the associated political economy considerations. Future research should consider the effect of incorporating multiple sources of information, uncertainty and heterogeneity on a policy decision and the analytical tools proposed in the chapter are the most appropriate.

2.5. Conclusion

This chapter has incorporated three aspects that are often ignored in the crop response literature, namely parameter uncertainty, multiple sources of information, and (spatial) heterogeneity in the response to fertilizer use. A Bayesian approach is employed to address each of these themes and close the measured gaps in the responses. This is an important goal for agricultural research because of the fairly constant trends of crop output/fertilizer price ratios across sub-Saharan Africa, which are indicative of the proposition that long-term trends in fertilizer profitability require improvements in farmer crop response rates (Jayne and Rashid 2013). The analysis has shown that using prior knowledge of crop response estimates adds insights to the assessment of crop responses using observational data. In particular we find that ignoring the precision parameter when using crop response estimates may lead to inconclusive policy prescriptions.

The debates on whether crop responses to fertilizer application are high or low are therefore questionable when uncertainty that appears to measurably affect the stochastic dominance ordering of crop response estimates is ignored. Unless uncertainty is considered, the arguments for or against the use of experimental and observation crop response estimates (e.g., Dorward and Chirwa (2015) and Jayne et al. (2015)) are inconclusive, thereby leading to questionable policy prescriptions. Moreover, while the debates have centered on means of crop responses, this chapter has shown that both the means and variances matter in these policy discussions. The results of incorporating heterogeneity in the estimation by way of using a hierarchical Bayesian modelling approach are quite revealing. We find that the degree of spatial heterogeneity in fertilizer responses varies markedly, with some districts being effectively non-responsive to the application of fertilizer (e.g., Chikwawa) while other districts are highly responsive (e.g., Dedza). Our results present a different and more nuanced picture relative to all the previous published research that completely ignores heterogeneity and relies on the assumption of homogeneity in assessing fertilizer crop responses.

3. Characteristics Space Analysis of Improved Maize Variety Adoption in Malawi

> *"…it is most important to distinguish between the lag in "**availability**" and the lag in "**acceptance.**" It does not make sense to blame…Southern farmers [in the United States] for being slow in acceptance, unless one has taken into account the fact that no satisfactory hybrids were available to them before the middle nineteen-forties".*
>
> Griliches (1957, p.507).

3.1 Introduction

In his seminal study of the time path of hybrid maize adoption in 20th century U.S. agriculture, Griliches (1957) elucidated (and empirically examined) the lags associated with delays in the market "availability" of relevant new varieties for each U.S. state, and the subsequent lags in the farmer uptake or "acceptance" of these new varieties within each state. However, most subsequent micro-econometric studies of varietal adoption—see, for example, the reviews by Feder et al. (1985) and Doss (2006)—focused exclusively on the acceptance problem by analyzing the fundamental constraints in remoteness, weak markets, distorted policies, low education, cultural and many other related constraints that affect farmer's willingness or ability to adopt new varietal technologies.

Focusing solely on these particular demand side determinates of adoption fails to provide operationally relevant information that can shape the supply of varietal innovations. In particular, past studies have been largely silent about the particular yield and crop quality traits or characteristics deemed desirable (or otherwise) by farmers facing particular (and often spatially variable) production problems.[16] An approach to assessing varietal adoption that improves our understanding about the nature of the (spatially variable) varietal attributes most valued by farmers can help shape the supply side behavior of crop innovators and thus help reduce the demand versus supply side information asymmetries that result in inefficient and underperforming varietal markets.

[16] Another limitation of many varietal adoption studies is that they also fail to make explicit the uptake implication of spatial and farmer heterogeneity (e.g., Suri 2011 and Duflo et al. 2011) and production risks and uncertainty. In the latter case, what may seem to be lack of adoption for unspecified reasons may simply reflect an optimal choice under risk and uncertainty (Hurley et al. 2018).

One approach to assessing varietal adoption using a varietal trait approach is to turn to the characteristics space model introduced by Lancaster (1966) to study the demand for goods (or in our case, farm inputs) where each of these goods (or inputs) is viewed as a "collection of characteristics." This model operationalizes Griliches' availability and acceptance delineation when analyzing micro adoption data. In his model, Lancaster differentiates between the characteristics of the product from the personal tastes of consumers. Characteristics were considered objective and universal properties of the good. In this chapter, we calculate the willingness to pay for varietal characteristics and then infer the predicted rate of adoption of new varieties using the characteristics space model.

An example of the effect that changes in the unit cost and characteristics of a technology have on the extent of adoption of that technology may help illustrate the importance of the approach. Although maize breeding research in Malawi began in the late 1950s, adoption of new hybrid maize varieties remained quite low for decades thereafter (Mkondiwa et al. 2019). It was not until the flint characteristic was introduced into hybrid maize, beginning around 1990, that the extent of varietal adoption began increasing from around 10% of the country's maize area in the late-1980s to roughly 30% by the mid-1990s. During the past decade, a subsidy on maize seed and fertilizer has spurred significant additional adoption of improved maize seeds to now roughly 70% of maize area (Figure 3-1). What changed here is the unit cost and characteristics of the seed, not necessarily the personal characteristics of the farmers.

In the context of Malawi, studying the specific characteristics of the varieties being adopted is of particular importance. Malawians generally reveal a distinct preference for certain varieties, like the flint varieties loosely categorized as local or "maize of the ancestors" (chimanga cha makolo) because of the efficiency in household processing to a fine white flour and the stored crop's resistance to infestation by weevils (Smale et al. 1995).

This study is motivated by the notion that when the characteristics bundled into a new varietal technology match the market and (spatially variable) production realities facing

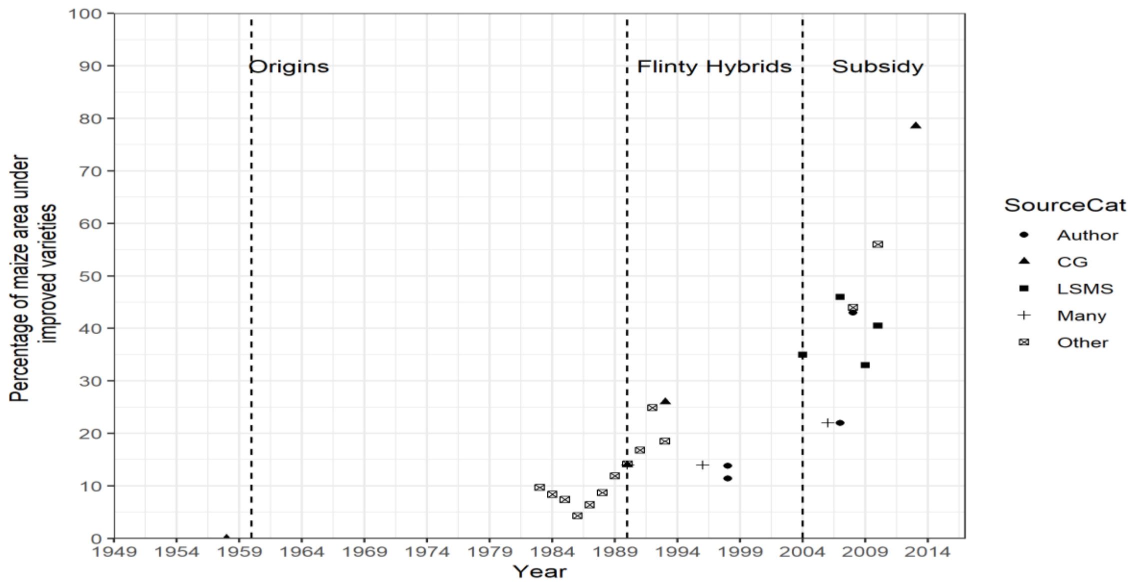


Figure *3-1*: Dynamics of improved maize *variety adoption in Malawi.*

Source: Mkondiwa et al. (forthcoming*).*

each farmer then that technology is more likely to be adopted. Varietal technologies that fail to be adopted tend to be those that are simply not known by farmers or those that contain characteristics (or in breeders' parlance, traits) that are not well aligned with the heterogeneous contexts faced by farmers (in terms of spatial concordance, market opportunities, and tastes). Investments in agricultural research and technology are made on the justification that the technologies developed will be used by the farmers. The challenge is in estimating the demand for these new technologies based on the observed adoption rates of similar technologies. Elucidating the specific characteristics deemed desirable by farmers helps inform crop breeders and those making public and private decisions regarding investments in R&D (research and development) and the bulking and marketing of new varieties

This chapter addresses the following two research questions: (i) how much are farmers willing to pay for various maize variety traits?, and (ii) what is the probability that a variety with a particular combination of characteristics will be adopted? We propose a characteristic space model that conceives of technologies (in this instance, improved maize varieties) to be a bundle of characteristics, and that adoption of these technologies depends on the heterogeneous preferences of farmers. This approach has had only limited use in technology adoption studies even though agricultural technologies (such as improved crop varieties) often constitute a set of quality differentiated inputs, each with its own set of bundled characteristics. Some of the few such studies in the crop variety demand literature include Edmeades and Smale (2006), Useche et al. (2013) and Ward et al. (2014) (see appendix table A1, for a summary of the studies on varietal adoption that use a characteristics approach).

However, even these prior adoption studies deploying a characteristics approach suffer from three important omissions. First, the studies ignore the supply side of the seed market, which is an integral part of the new industrial organization models for differentiated products (e.g., Nevo 2001) that we draw on for this study. Second, only a few of the models consider the fact that (subsistence or semi-subsistence) farmers are both consumers and producers such that varietal adoption involve both production and consumption decisions. Third, all the models ignore the objective characteristics conception that Lancaster described in the original framework he introduced in 1966. Most of the prior studies have farmers' scoring varietal traits, yet these are not objective measures of the technical and biological characteristics of the traits embedded in each variety. In essence most of what these studies measure and denote as traits are simply part of the heterogeneity in the location and conditions of the farmer. For instance, an early maturing variety does not suddenly become a late maturing variety simply because a farmer in a certain agroecological locale feels the growing season in that location is too short for the variety to mature. In this instance, the subjective notion of "late maturing" is affected by the variable agroecological realities faced by each farmer based on their particular locale and their perception of the duration of the growing season vis a` vis the maturity potential of the variety.

The use of farmer subjective ratings of various attributes of (varietal) technologies introduces endogeneity because the varietal ratings are correlated with unobserved (or unmeasured) household characteristics such as locale and taste. To circumvent endogeneity concerns in farmers' ratings of revealed varietal choices, some studies (e.g., Kassie et al. 2017 and Ward et al. 2014) use stated choices in choice experiments. For example, Ward et al. (2014) presented rice farmers in Bihar, India with four hypothetical rice varieties varying in characteristics which included duration (days to maturity), yield (under normal and low rainfall scenarios), re-usable seed, seed price and seed rate. While this approach provides an unbiased way to analyze adoption, just as any stated preference method it fails to capture the reality of the production aspects of varietal adoption, which is that the choice is made without all the constraints of access, weather patterns and credit constraints.

When we are explicit about spatial, topographic and weather-related factors when assessing the demand for varietal characteristics, we will be better able to account for that part of *non-adoption* that can be ascribed to a *"mismatch"* between the supply and demand of objective characteristics bundled in a particular variety, setting aside potentially confounding factors. Our model is estimated using frequentist and Bayesian random coefficient logit model (Jiang et al. 2009) to allow for parameter uncertainty when seeking to predict the adoption of new varieties. Using these procedures to identify an optimal set of characteristics to enhance adoption is a novel way of deploying economic analysis to direct crop breeding efforts towards those bundles of traits that are more likely to be adopted. In addition, by distinguishing between varietal characteristics and consumer specific tastes, it is possible to appreciate analytical methods like genotype by trait (GxT) and biplot analysis that breeders use when choosing a set of varieties to advance in a breeding program (see, for example, Yan 2014, chapter 9 for details of these approaches). We can also use production economic concepts of data envelopment analysis (DEA) or non-parametric revealed preference to describe the characteristics frontier as envisaged by Lancaster (see, for example, Fernandez-Castro and Smith 2002 and Blow et al. 2008).

We anchor this analysis on the salient features of the differentiated maize seed market in Malawi in which one of the varieties being considered, a dominant local variety, is saved from previous harvests and exchanged among farmers. The study allocates this variety to an imaginary seed company, "Ancestor's Seed Company" to imply the local name of "*Chimanga Cha Makolo*" (Maize of the Ancestors). This then allows an analysis of the market power features of this "Seed Company" and how government policy can help reduce the market share of this variety in favor of varieties with superior crop performance and other desired attributes.

3.2 Model

This section provides an overview of the theoretical model of adoption in the characteristics space and is divided into two parts, (1) theory of characteristic space adoption analysis, and (2) estimation and identification strategy.

3.2.1 Characteristic space adoption analysis: Theory

(a). What is a variety characteristic?

Lancaster (1971) defines a characteristic as an objective and universal property of a good. For example,"…beauty is not a characteristic because it is in the eyes of the beholder but such things that spring out the beauty of a good like color are characteristics [are deemed characteristics] (Lancaster 1971, p.114)". Because characteristics can be numerous, he suggested that only the more relevant ones be considered as part of a demand analysis. The theoretical rationale for this particular operational definition springs from the ability to isolate the technical-goods characteristic relationships from the people-goods characteristics. In the context of maize varieties, this delineation allows identification of key characteristics that breeders incorporate into the new varieties. By requiring that these characteristics be external from the farmers' subjective assessments, we are able to assure exogeneity of these characteristics relative to unobserved farmer characteristics.

For example, cookability is not a characteristic in a Lancasterian sense, but flint or dent are such characteristics. Cookability reflects a farmer's perception of a particular variety, whereas a farmers' opinion does not change the flinty or dent characteristics of a variety. This conception of characteristics has been used in the industrial organization literature to estimate the demand for goods with specific quality attributes, including, for example studies by Berry et al. (1994 and 2004) and Nevo (2001). Berry et al. (1994 and 2004) used the following characteristics when modeling the demand for cars: the horsepower to weight ratio of the modal car engine (designation acceleration), number of passengers (size), city miles per gallon, payload in thousands of pounds, plus a dummy variable to designate a minivan. When studying the demand for ready-to-eat cereals, Nevo (2001) used calories, sodium and fiber content as the set of defining characteristics for these goods.

The key challenge to the proposed characteristic approach of modeling adoption is that the differentiated product market model of Lancaster was specifically developed in the context of consumption goods not production goods (inputs) like varieties. There are, however, general models that capture both the production and consumption decisions of agricultural household behavior (e.g., Pollak 1989). Deploying an agricultural household (demand and supply) model in the context of a crop with varietal attributes draws on the same empirical modeling constructs as assessing supply and demand in a conventional differentiated products market model.

(b). Demand

We assume that a farmer i derives utility u_{ij} from the choice of a maize variety j. Let $j = 0, \ldots, J$ index the varieties competing in the market, where variety $j = 0$ is the outside input (in this case a local reference variety). Let k index the observed variety characteristics, including price, and r index the observed household attributes. Following Berry et al. (2004), the farmer's utility from choosing a particular variety is given by

$$u_{ij} = \sum_k x_{jk}\tilde{\beta}_{ik} + \xi_j + \epsilon_{ij} \quad \textbf{(11)}$$

$$\tilde{\beta}_{ik} = \overline{\beta_k} + \sum_r z_{ir}\beta^0_{kr} + \beta^u_k v_{ik} \quad \textbf{(12)}$$

where ξ_j are unmeasured aspects of maize variety quality, x_{jk}, the observed aspects of the maize variety quality, $\tilde{\beta}_{ik}$ is the taste of consumer i for product characteristic k, z_i are observed consumer attributes, v_i are unobserved consumer attributes, and ϵ_{ij} represents idiosyncratic individual preferences that are assumed independent of the variety characteristics and of each other. Farmers differ in terms of their varietal tastes by a vector $\boldsymbol{z}$ of observed demographic variables.

Substitute equation (12) into (11) to get:

$$u_{ij} = \delta_j + \sum_{kr} x_{jk} z_{ir}\, \beta_{kr}^{0} \ + \sum_{k} x_{jk} v_{ik}\, \beta_k^{u} + \epsilon_{ij} \qquad \textbf{(13)}$$

where, for $j = 0,1, \ldots, J$

$$\delta_j = \sum_{k} x_{jk}\, \bar{\beta}_k + \xi_j.$$

We estimate the equation using a discrete choice model where the dependent variable is a multinomial choice of the variety grown. Empirically, the model is estimated using frequentist and Bayesian approaches. The aim is to obtain willingness to pay measures, the estimated shares of each of the varieties that can then be used to compute the price-cost margins of the seed companies, and predictions on the adoption rates of new maize varieties.

Assuming the error term is from an identically, independently distributed (iid) extreme value distribution, a standard multinomial logit regression can be used for equation 13. However, the challenge with the standard multinomial logit regression approach is that the number of parameters to be estimated increases with the number of varieties. Again, assuming the error term is from an identically, independently distributed (iid) extreme value distribution, we can instead use a linearized version of a logit specification for equation 13 where the dependent variable is defined as $\log\left(\frac{s_j}{s_o}\right)$ such that s_j is the market share of variety j and s_0 is the share of the local variety.[17] This reduces the number of parameters to be estimated to the total number of characteristics and demographic attributes. The estimating equation for a linearized logit model is

$$\log\left(\frac{s_j}{s_0}\right) = \beta_0 + \beta x_j - \alpha p_j + z_i \tau + \xi_\mathrm{j} + v_i + \epsilon_{ij} \qquad \textbf{(14)}$$

[17] This is standard in industrial organization literature and for derivations, the interested reader is referred to Nevo and Rosen (2012, p.666).

The willingness to pay for variety traits is calculated as

$$WTP_k = -\frac{\beta_k}{\alpha} \qquad \textbf{(15)}$$

(c). Supply

To understand the effects of the demand for specific varietal characteristics on the development and pricing behavior of seed suppliers (market power), requires also developing the supply side of the model. The seed supply market can be categorized as oligopolistic with few companies that exhibit market power. Previous research (Chinsinga, 2010; Mason and Ricker-Gilbert 2013) has reported that although the Malawian seed industry is reasonably open to entry, it effectively has been dominated by a few multinational companies that have little incentive to develop and supply maize varieties that better meet smallholder needs.

We can therefore be confident to follow the tradition in the industrial organization literature of inferring price-cost margins using a pricing equation identity. Following Nevo's (2001) notation, consider a finite number of seed suppliers (seed companies), F, each firm (f) producing a subset of maize varieties ($j = 1, \dots, J$) adopted by farmers, $\mathcal{F}_f$, to maximize profits

$$\Pi_f = \sum_{j \in \mathcal{F}_f} \left(p_j - mc_j\right) M s_j(p) - C_f \qquad \textbf{(16)}$$

where $s_j(p)$ is the market share of variety j as a function of the prices of all varieties, M is the size of the market and C_f is the fixed cost of production. The term in the bracket is the price-cost margin, measuring the difference between the seed price and the marginal cost of seed production. After deriving the Betrand-Nash first-order condition, Nevo (2001) found that the price-cost margin can be deduced as a function of the market shares (s) of each variety and the ownership matrix (Ω) (consisting of dummies to delineate which company produces which variety) as follows:

$$p - mc = \Omega^{-1}s(p) \quad \textbf{(17)}$$

The price-cost margins can then be used to analyze whether the seed market is competitive.

3.2.2 Identification strategy

There are three potentially problematic identification issues when estimating any structural model (including this one) using observational data; specifically omitted variable bias (unobserved heterogeneity), simultaneity and measurement error. We now discuss each of these identification issues in turn.

(a). Threats to identification

First, omitted variable bias is a common threat to identification in differentiated product markets, not least because consumers are assumed to make choices based on complex unobserved tastes and preferences. In the variety adoption decision, there are several sources of omitted variables. These can be categorized into two groups: unobserved technical characteristics and unobserved farmer characteristics. Although we sought to collect a comprehensive set of key characteristics for each of the varieties using data obtained from the seed companies and breeders, it is possible that the companies and the scientists did not provide all the relevant information, especially if that information were the source of their comparative advantage in the seed market. We can control for some of this variation by including a company fixed effect and the date of release of the variety to signal the level of knowledge in the market, but this won't wipe away all unobserved characteristics. In addition, there are several straightforward unobserved factors like the temperature, humidity and moisture requirements for each variety. Essentially, we expect that the details on each variety's set of characteristics made available to members of the variety release committee are much more than that company makes available to the public or is published in government variety release documents. The instruments that have been considered in the literature include characteristics of other products (varieties) and cost shifters.

Second, simultaneity has been the main thrust of employing the structural models found in the industrial organization literature to capture both demand and supply factors. By following this conventional approach, we are less concerned with this threat to identification. In addition, we use variables that are deemed likely to affect varietal choices, but in turn are not affected by that choice based on our knowledge of agricultural production. For example, instead of the actual yield of the variety, which is affected by the choice of the variety, we use yield potential, which does affect the choice of the variety but does not have reverse causality.

Third, the problem of measurement error is common in variety adoption decisions based on recall information. First, it is difficult even for well-trained scientists to recognize some particular varieties, let alone open pollinated varieties whose observable traits might be affected by the varieties grown in close proximity. Second, in developing our database on varietal characteristics, different sources of data sometimes offered contradictory descriptions of a particular variety, thus introducing the possibility of measurement error when determine the characteristic list for a given variety. By restricting our empirical analysis to the better-known and widely used hybrid varieties, we reduced the potential of measurement error. In addition, triangulating information from several key sources provides some confidence for the second concern.

(b). Control function approach to addressing endogeneity

Much of the industrial organization literature uses the Berry et al. (1994) instrumental variables (IV) approach when dealing with endogeneity in this context. However, this approach is computationally heavy and prone to numerical errors. We thus considered alternatives to addressing the problem, and settled on using the control function (CF) as an alternative estimation strategy because it has several advantages in our particular case. It is easier to estimate and is applicable in cases where the Berry et al. (1994) approach is not valid, which includes our case where the number of products is small (Petrin and Train 2009, p.4). In addition, standard IV approaches are problematic in nonlinear models like the multinomial logit, thus it is advisable to use the control function in this circumstance (Zeng 2014, p.90). The two candidate groups of instruments are the price of the variety in other countries (e.g., Zambia), or the biological characteristics of the varieties that affect the cost of production but are not observed by the farmers.

The problem with the first set of instruments is that some of the varieties like *MH18* are only available in Malawi. We thus maintain the Malawi price in the instrument set. We specifically consider biological characteristics of the hybrid varieties that affect the cost of production by the seed company but potentially do not affect farmers' preferences. Some of the candidates of these biological characteristics include seed production method, sowing rate, multiplication factor, rate of deterioration of seed viability and frequency of purchase (Cromwell, Friis-Hansen and Turner, 1992; Maredia et al., 1999). Of these we considered the seed production method—i.e., a single cross, three-way cross, or double cross— as the most distinguishing characteristic that increases the cost of seed production.[18] According to studies cited above, in 1996, the grain/seed price ratios were: 1:5 for a single cross, 1:3 for a three-way cross and 1:2 for a double cross. A recent study (Mabaya et al. 2019, p.7) documents a hybrid maize grain/seed price ratio of 1:4.17 and 1:4.05 for OPVs. The reasons behind these OPV versus hybrid price differentials is perhaps succinctly stated by the Malawi seed company SeedCo, whose website states:

> *"SC 727 seed is a single cross hybrid, i.e., it is produced from crossing two inbred lines which are generally poor in terms of seed yield produced, hence it's very difficult and costly to produce. The price we give to SC 727 producers or seed growers is compensatory to cover up for the low productivity levels and the cost is transferred to the farmers/consumers hence the high price relative to three-way hybrids"* Seed co Malawi (www.seedcogroup.com/mw/media/faqs).

With these two instruments, we estimated the following first-stage regression:

$$P_{jt} = \alpha_0 + \alpha_1 ZambiaPrice_j + \alpha_2 Single\ Cross\ Dummy_j + \alpha_3 x + \tau_{jt}, \qquad \textbf{(18)}$$

[18] Single, double and three-way cross refer to the number of parents a hybrid has: two parental lines for a single cross; two single crosses for a double cross; and a pure inbred male parent and single cross female parent for a three-way cross (Maredia et al. 1999).

where P_{jt} is the calculated price of a particular hybrid seed j in market/district t. All the hybrids released in recent years in Malawi are three-way crosses, and as such we could not use this as an instrument in the first stage but would be relevant for multi-country varietal demand analysis. This also points to the changing nature of varietal supply and demand dynamics that have evolved overtime to affect the constellation of the available hybrid varieties.

Given there are no data on seed prices at the district level in Malawi, we computed the district-level seed price using the spatial price spreads in a related market, specifically the maize grain market, which has longitudinal price data at the district level. We also used the observed seed-grain price ratio to compute district level seed prices as follows:

$$P_{jt}^{Seed} = \frac{P_j^{Seed}}{P_{sc403}^{Seed}} \times \frac{P_{sc403}^{Seed}}{P_{grain}} \times P_{Grain,t}, \tag{19}$$

where P_{jt} is the seed price for variety j in district t, $P_{Grain,t}$ is the price of maize grain for each district, P_j^{Seed} is the national level price of the variety j and P_{sc403}^{Seed} is the seed price of the dominant hybrid seed. By using this approach, we assume that districts with high grain prices also register high seed prices and that the seed-grain price ratio is based on prices of the dominant variety. These assumptions are likely to hold because districts whose hinterland are hard to reach will have higher seed prices and because it is difficult to move grain to the center of the district, the grain prices are also very high[19]. And SC 403 is grown nationwide as such the price ratio can be assumed not to be district specific.

[19] This is also consistent with research that find that inputs are costly in the "last mile" (Minten et al. 2013).

3.3 Data and descriptive statistics

3.3.1 Data

We use two key sources of data to estimate the model. These are (i) household level data collected under the auspices of the adoption pathways project,[20] and (ii) varietal characteristics data compiled by the authors from various company websites, variety release documents, and maize breeding trial reports

Household Level Data

A key variable gleaned from the household level survey data is the name of any improved varieties the farmer planted in a particular reference year, the dependent variable in our model.[21] Other (right hand side) variables collected from this source of data include: i) household characteristics, including household size, household head education, household head age, and ii) farm characteristics, including farm size, fertilizer use, yield, whether the plot is mixed, and average maize yield. The data are compiled and analyzed at a plot level. The key assumption in using Berry et al. (1994) type models is that the consumer buys one product only. In the data used in this chapter, there is a tendency for individual farmers to grow multiple varieties. This however is controlled by plot level variables that allow a distinction across households based on plot characteristics. The results from this chapter should be interpreted with respect to the study population we are considering in the data. The study population includes smallholder farmers growing maize only within the sampled districts. In addition, because CIMMYT, via the SIMLESA project, has conducted several studies in these locations, it is expected that the adoption levels are likely to be higher than the average farmer in Malawi.

[20] These data can downloaded from http://data.cimmyt.org/dvn/dv/cimmytdatadvn/faces/StudyListingPage.xhtml?mode=1&collectionId=122.

[21] The survey covers 16 districts in Malawi, specifcally Mzimba, Dedza, Kasungu, Ntcheu, Dowa, Ntchisi, Salima, Mchinji, Balaka, Blantyre, Chiradzulu, Machinga, Mangochi, Mwanza, Thyolo and Lilongwe.

Varietal Characteristics Data

In constructing the variety characteristics database, we used multiple sources of information. The first step was to determine the attributes of the more than 40 varieties that were reported in the household survey using various sources. One was the DIIVA database (https://www.asti.cgiar.org/diiva), which contains information on the release date for each of the varieties and the name of the company that produced the variety. We searched online for information posted on the respective company websites regarding the different characteristics of each variety. For some varieties, we were able to cross-check the information provided by the seed companies against variety technology release documents published by the Ministry of Agriculture and Food Security in Malawi (Saka et al. 2006, Mviha et al. 2010, and Chisama et al. 2015) and a seed security assessment report by (Bokosi et al. 2011). Wherever possible we also verified the information with experimental field trials estimates drawn from multiple studies. Table 3-1 summarizes information on the main varieties, including their year of release and selected traits of each variety.

Table 3-1: Varietal characteristics

Variety name	V.age	Seed Company	Maturity (Days)	Flint	Yield (t/ha)	Price, MK/ kg	Drought Tolerance	MSV Resistance	GLS Resistance
SC403	14	SeedCo	100	Flint	4.5	320	Yes	Yes	No
SC627	13	SeedCo	125	Flint	9.0	320	No	Yes	Yes
DK8033	10	Monsanto	112.5	Dent	8.0	342	Yes	Yes	Yes
DK8053	5	Monsanto	125	Flint	10.0	342	No	Yes	Yes
MH18	22	NMBP*	125	Flint	6.0	365	No	No	No
SC719	4	SeedCo	130	Dent	11.5	400	Yes	Yes	Yes
PAN53	5	Pannar	137.5	Flint	9.0	365	Yes	Yes	Yes
DK9089	3	Monsanto	117.5	Flint	10.0	342	No	Yes	Yes
Local	> 50	Ancestor		Flint		33	-	-	-

Notes: * released by the National Maize Breeding Programme (NMBP) and currently produced and marketed by Pannar seed company. The disease resistance traits (MSV and GLS resistance) are similar across varieties, and as such we did not include them in the estimation. For local variety, the characteristics are assumed for exposition but were not used in the estimation. V.age is variety age, i.e. number of years between variety release and survey year.

There are more than 40 maize varieties reportedly used by the smallholder farmers in the sample, with the top nine accounting for more than 85 percent of the market share (delineated by the total number of farmers growing maize). With this level of varietal concentration, the sample sizes for the other varieties were small and as such were not used in the analysis.

3.3.2 Piecewise linear representation of the characteristics frontier

Figure 3-2 illustrates the characteristic efficiency frontier represented by variety 1, variety 2, and variety 3. If variety 2 is not available, then the characteristic frontier is composed of variety 1 and variety 3.

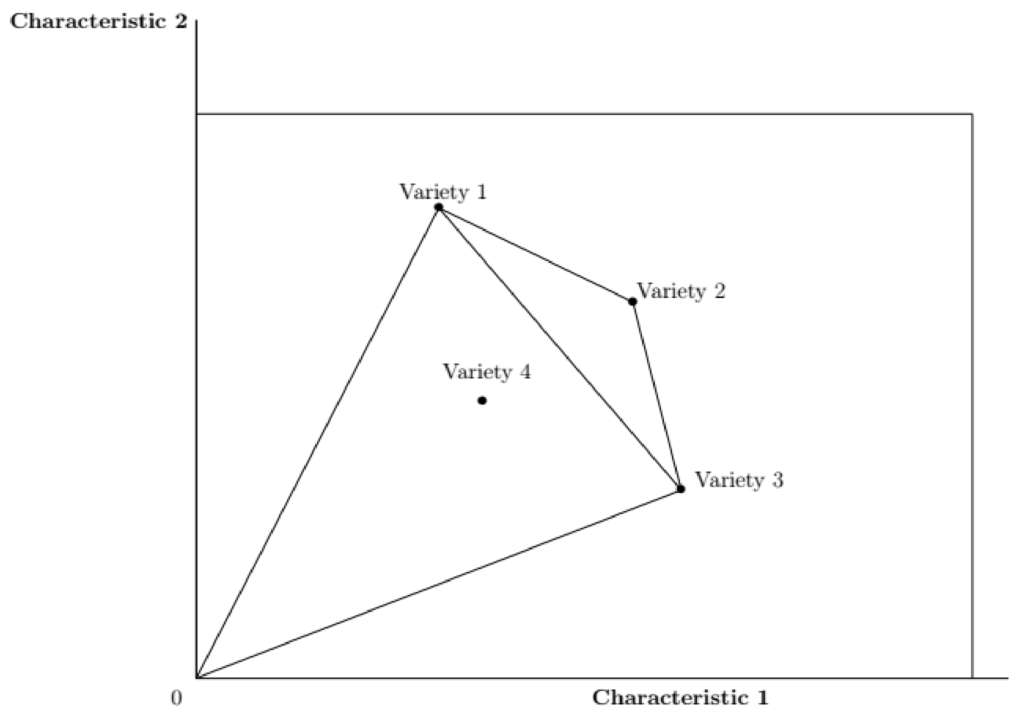


Figure 3-2: Varietal characteristic frontier

Note: The lines illustrate the discrete nature of the choices. That is, a farmer cannot choose a convex combination of the varieties.

Using the characteristics database, it is possible to graphically illustrate the efficient frontier varieties across selected two-dimensional characteristics (as in figure 3-2). For a two-dimensional approximation of a multi-dimensional frontier, a biplot analysis as commonly used by breeders may be adapted (Yan 2014). The graphical analysis can help ascertain the varieties that are likely to be adopted. Figure 3-3 shows a two-characteristic (potential yield and days to maturity) plot for the eight top hybrid varieties (see appendix Figure A1 for two-dimensional representations of multiple characteristics).

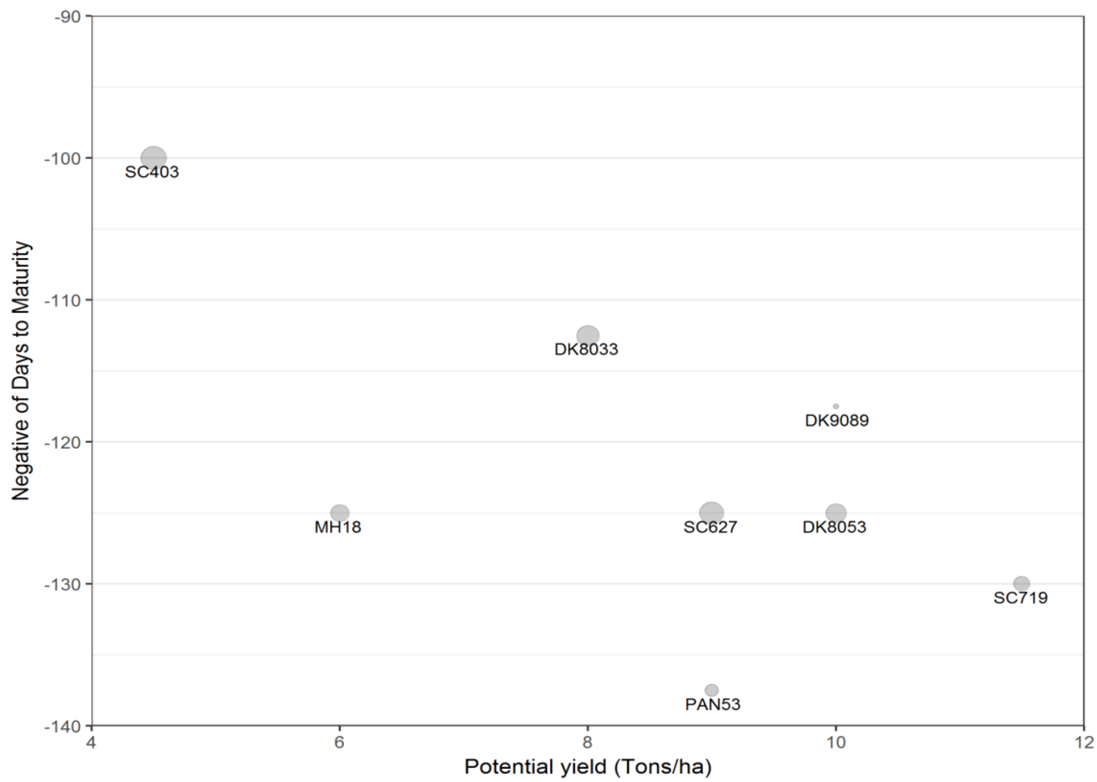


Figure 3-3: Maturity-Yield variety frontier, discrete case

Notes: The size of the circle represents the area extent of adoption, with SC403 being the most widely planted improved maize variety.

It is apparent that days to maturity matters for adoption because the variety SC403 with the lowest days to maturity (i.e., highest negative days to maturity in the plot) is also the most planted even though it has lower yields. Nonetheless, a high-yielding variety like DK9089 is planted less because it is a younger variety, as shown figure 3-4. However, DK9089 is better than most varieties on days to maturity and is flinty, as such using the variety frontier we can deduce that in the fullness of time this variety may take up much of the market share particularly from DK8053.

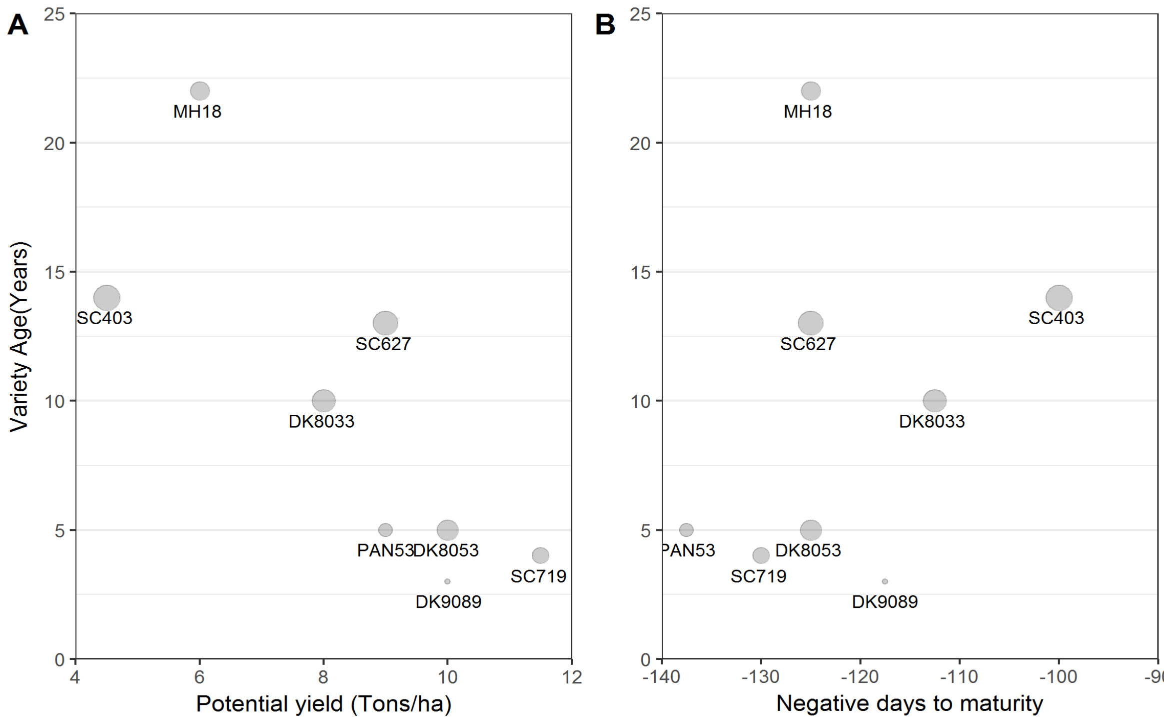


Figure 3-4: Maturity-Variety Age-Yield frontier, discrete case

3.3.3 Household characteristics

The main household characteristics in the literature include household size, age, education, and gender. In our sample, 88% of the household heads were male. This is consistent with national household surveys that find that due to cultural norms, most households identify the oldest male in the household as the household head. The average household size was 5 people, with the household head aged about 44 years. Years of schooling for the household head was about 6.68. The average plot size allocated to maize was about 0.5 ha. Over half of the households received subsidized seeds and fertilizers. Comparing across the different maize varieties, it is apparent that households adopting local varieties do not markedly differ in characteristics to those adopting improved varieties, except for slight differences in household sizes, plot size allocated to the variety, access to seed subsidy, and access to fertilizer subsidy. In general, those sticking with local varieties have smaller household sizes, larger plot sizes allocated to the variety, low access to seed subsidy and low access to fertilizer subsidy.

Table 3-2: Shares of each of the varieties studied and household characteristics

			Household size		Age		Plot Area		Years of Schooling		Gender(Male)		Fertilizer subsidy (Yes)		Seed Subsidy(Yes)	
Variety	Sample	Proportion	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Local	821	0.33	5.44	2.62	47.10	14.72	0.53	1.77	6.55	12.36	0.81	0.39	0.56	0.50	0.49	0.50
DK8033 (Mkangala)	191	0.08	5.64	2.08	42.81	10.38	0.41	0.27	7.54	15.96	0.90	0.31	0.67	0.47	0.64	0.48
DK8053 (Mapasa)	197	0.08	5.89	2.35	42.01	13.27	0.54	0.51	6.93	7.56	0.90	0.30	0.61	0.49	0.60	0.49
DK9089 (Fumba)	88	0.04	5.83	2.57	41.75	10.15	0.45	0.30	6.19	3.23	0.90	0.30	0.67	0.47	0.66	0.48
MH18 (Chokonoka)	157	0.06	6.35	2.96	44.75	14.11	0.45	0.28	5.83	3.49	0.90	0.30	0.48	0.50	0.44	0.50
PAN53	82	0.03	5.88	2.32	46.60	14.41	0.52	0.51	6.18	3.57	0.94	0.24	0.68	0.47	0.56	0.50
SC403 (Kanyani)	501	0.20	5.69	2.52	44.63	14.01	0.38	0.37	6.52	10.94	0.81	0.39	0.66	0.47	0.63	0.48
SC627 (Mkango)	332	0.13	5.62	2.74	44.57	14.10	0.41	0.25	6.70	10.05	0.83	0.38	0.62	0.49	0.56	0.50
SC719 (Njovu)	127	0.05	5.85	2.22	42.20	13.74	0.47	0.34	7.72	9.17	0.91	0.28	0.62	0.49	0.60	0.49
All	2496	1.00	5.80	2.49	44.05	13.21	0.46	0.51	6.68	8.48	0.88	0.32	0.62	0.48	0.58	0.49

Note: The share of farmer growing the variety is consistent with other studies like Holden and Mangisoni (2013).

3.4 Estimation results and discussion

3.4.1 Multinomial logit results

The commonly used econometric model both in the traditional adoption analyses and the recent trait-based analyses is the multinomial logit or probit models (or their binomial variants). We also use the multinomial model as the baseline model for the mixed logit. Table A2 in the appendices shows the results for a multinomial logit estimation without the varietal characteristics included. This is akin to using a binary choice model in keeping with most of the prior agricultural economics literature. The reference variety in this case is the local variety. In general, age of the household head is negatively related to the extent of adoption of most of these varieties. This is consistent with most prior binary choice literature in Malawi, which finds that age is negatively related to the adoption of improved varieties (e.g., Bezu, et al. 2014). While most of these studies find a positive effect of education on adoption, this is not consistent across all improved varieties as shown in the table. The differences in the parameters across the different varieties also cautions against using binary choice models to ascertain whether certain variables positively or negatively affect adoption. This is because, depending on the portfolio of varieties in the market, one may get different results. The characteristics approach essentially add variety level characteristics to this basic model.

Table 3-3 shows linearized multinomial logit results. We show only estimates on the variety characteristics as estimates of the demographic are similar to the baseline model. Firstly, as expected from consumer theory, the price coefficient is negative. On average, it appears that relative to the local variety, an increase in price decreases the adoption of that variety after controlling for important varietal characteristics. Nonetheless, the price coefficient may be biased for two reasons. First, it may imply that there is individual heterogeneity in preference we are not capturing with multinomial logit model. Second, for measurement reasons we may not be capturing all the relevant varietal quality characteristics for either the local variety or each of the hybrid varieties. We thus consider next models that include individual level heterogeneity in preferences and in Table 3-4, we show the results using control function approach (model 1 and 2 are first stage) with varietal prices in Zambia being used as instruments for variety prices in Malawi. Adding a control function term marginally changes the price coefficient for national and district level price regressions (model 4 and 6).

Table 3-3: Linearized logit results

Dependent variable: $log\ (\frac{s_j}{s_0})$ where s_j share of variety adopted and s_0 is share of local variety				
	Model 1	Model 2	Model 3	Model 4
(Intercept)	-0.5600***	1.3716***	1.3793***	1.3641***
	(0.0787)	(0.1282)	(0.1308)	(0.1536)
Age	0.0033***		-0.0002	-0.0002
	(0.0011)		(0.0004)	(0.0004)
Household size	-0.0139**		0.0004	0.0004
	(0.0059)		(0.0019)	(0.0019)
Education	-0.0164***		0.0004	0.0004
	(0.0043)		(0.0014)	(0.0014)
Sex	-0.0087		-0.0105	-0.0026
	(0.0668)		(0.0227)	(0.0404)
Plot area (ha)	0.0138		0.0185	0.0184
	(0.0136)		(0.0134)	(0.0134)
Fertilizer subsidy	0.0200		-0.0235	-0.0234
	(0.0624)		(0.0213)	(0.0213)
Seed subsidy	-0.1159*		0.0151	0.0310
	(0.0613)		(0.0210)	(0.1407)
Marital status	-0.1927***		0.0040	0.0042
	(0.0666)		(0.0227)	(0.0227)
Seed price		-0.0095***	-0.0095***	-0.0095***
		(0.0004)	(0.0004)	(0.0004)
Variety age (years)		0.1580***	0.1583***	0.1582***
		(0.0035)	(0.0036)	(0.0036)
Days to maturity		-0.0619***	-0.0619***	-0.0619***
		(0.0016)	(0.0016)	(0.0016)
Flint		1.2926***	1.2946***	1.3027***
		(0.0368)	(0.0370)	(0.0501)
Yield potential		0.0006***	0.0006***	0.0006***
		(0.0000)	(0.0000)	(0.0000)
Drought tolerance		1.2325***	1.2368***	1.2365***
		(0.0308)	(0.0310)	(0.0311)
Seed price :Fertilizer subsidy				-0.0000
				(0.0004)
Flint: Sex				-0.0094
				(0.0395)
R^2	0.0362	0.8943	0.8946	0.8946
Num. obs.	2458	1651	1651	1651

***p < 0.01, **p < 0.05, *p < 0.1

Table 3-4: Addressing endogeneity using the control function approach

	Control function first stage		Control function second stage			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dependent variable	Seed price	District seed price	$log\ (\frac{s_j}{s_0})$			
(Intercept)	387.3366***	377.1106***	2.8392***	2.6266***	0.9609***	2.0242***
	(0.5090)	(26.8578)	(0.1282)	(0.0000)	(0.1169)	(0.1177)
Zambia seed prices	6.2450***	7.4459***				
	(0.0100)	(0.5283)				
Variety age (years)	-13.0841***	-11.6286***	0.1580***	0.1613***	0.1898***	0.1895***
	(0.0211)	(1.1128)	(0.0035)	(0.0000)	(0.0038)	(0.0034)
Days to maturity	5.1252***	4.5679***	-0.062***	-0.065***	-0.087***	-0.076***
	(0.0064)	(0.3361)	(0.0016)	(0.0000)	(0.0014)	(0.0014)
Flint	-112.5436***	-103.4466***	1.2926***	1.3420***	1.7407***	1.5606***
	(0.1690)	(8.9153)	(0.0368)	(0.0000)	(0.0367)	(0.0342)
Yield potential	-0.0605***	-0.0534***	0.0006***	0.0006***	0.0007***	0.0007***
	(0.0001)	(0.0049)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Drought tolerance	-105.6402***	-91.5964***	1.2325***	1.2452***	1.3750***	1.4895***
	(0.1983)	(10.4622)	(0.0308)	(0.0000)	(0.0352)	(0.0321)
Seed price			-0.0095***	-0.0085***		
			(0.0004)	(0.0000)		
Control function (CF) term				-0.2321***		
				(0.0000)		
District seed price					-0.0010***	-0.0071***
					(0.0001)	(0.0003)
CF term with district seed prices						0.0068***
						(0.0003)
R^2	0.9988	0.2960	0.8943	1.0000	0.8590	0.8865
Num. obs.	1651	1651	1651	1651	1651	1651

***p < 0.01, **p < 0.05, *p < 0.1

3.4.2 Mixed multinomial logit and heterogeneity

Following Useche et al. (2009), we analyze a district-level heterogeneous model for the biophysical traits. This allows an investigation of whether the biophysical matching of the characteristics to the locations affects the adoption of these varieties. We also analyze an individual household mixed logit model. The results in appendix tables A3 and A4 align closely with those presented in section 3.4.1.

3.4.3 Are variety characteristics overrated? The shadow prices of characteristics

Using the coefficients in Table 3-3 (model 3), we calculate the ratio of the seed price coefficient and each of the variety characteristics to estimate the shadow prices of each varietal characteristic. We report willingness to pay measures using a Bayesian version of the logit specification because of the difficulties in computing moments for a random coefficient logit model (see Daly 2011 for details). When a goal of the analysis is to estimate farmers' willingness to pay (WTP), or to conduct welfare analysis, it is important that the price coefficient be negative and not overlap zero for all consumers (Hess and Train 2017). While the price coefficients are negative in our particular application of the multinomial logit model, this may not be the case in most applications, thus resulting in potential for substantial specification search and publication bias. The advantage of using a Bayesian approach, is that we don't need to worry about the significance of the price coefficients and the characteristic coefficients. We simply use the 10,000 MCMC draws to determine the shadow prices of the various traits. Table 3-5 shows the estimates of the median and 95% credible intervals for the willingness to pay for each of the varietal characteristics included in our analysis.

Table 3-5: Estimated shadow prices of varietal characteristics, Malawi Kwacha

WTP for	**Willingness to pay quantiles**		
	2.50%	50%	97.50%
Variety age	15.077	16.659	18.479
Days to maturity	-7.338	-6.52	-5.815
Flintiness	121.676	136.325	153.226
Yield	0.055	0.061	0.068
Drought tolerance	118.354	130.201	143.829

The signs on the trait specific WTP estimates indicate that farmers are unwilling to pay for a late maturing variety, or in other words are willing to pay for an early maturing variety. Notably farmers are willing to pay significantly more for complex traits like drought tolerance and flint texture when choosing hybrid varieties. To understand the magnitude of these effects, we can consider the WTP values in yield equivalents. The WTP for yield is 61 Malawi Kwacha for an additional yield of 1,000kg. our results implies that farmers are willing to pay about 2.13 times more for drought tolerance than they are willing to pay for an additional 1,000kg of yield. Similarly, farmers are willing to pay about 2.23 times more for a flint variety relative to a higher yielding (by 1,000 kg/ha) variety. Coincidentally, our Bayesian estimates of farmer's valuation of these traits in Malawi is similar to a discrete choice experimental valuation in Zimbabwe where farmers were paying a drought tolerance and flint texture premium of 2.56 and

5 times respectively relative to the amount they were willing to pay for a higher yielding (1,000 kg/ha) maize variety (Kassie et al. 2017).

3.4.4 Predicting adoption and market power implications of a new variety

There are several approaches that are used to predict the demand for a new product and the market power implications associated with that new product. These include predictions of market shares from the multinomial logit model (e.g., Berry et al. 2004) and calculation of the virtual prices for the differentiated product by computing prices at which demand for the new product is zero(Hausman 1996). Here, we introduce a new approach that uses the WTP measures and the characteristics frontier to predict the demand for new products. We begin by re-calibrating the characteristics frontier from a quantity to a value space. To do that, we simply multiply the WTP estimate by the quantity of the variety trait, X_{kj}. The sum of the tarit-specific value across all the measured characteristics gives the quasi-varietal value

$$Quasi - variety\ value_j = \hat{Q}_j = \sum_{k=1}^{K} WTP_k \times X_{kj} \quad \textbf{(20)}$$

Assuming farmers aim to maximize the value of this measure or minimize the sum of losses from over or under-planting particular varieties, then in aggregate we expect that the share of farmers adopting a variety will be proportional to this measure. Thus

$$Share\ adopting\ variety_j = s_j = \frac{Quasi - variety\ value_j}{Total\ Quasi - variety\ value} = \frac{\hat{Q}_j}{\sum_{j=0}^{J} \hat{Q}_j} \quad \textbf{(21)}$$

The way to think about this is to consider a social planner who wants to minimize the total national losses from over- or under-planting particular varieties throughout the country. This objective function is consistent with the goals of African governments (especially Malawi) that are aiming to increase the number of farmers planting improved varieties and reduce the number of farmers planting local variety. The relative penalty for over-planting a variety is given by s and the penalty for under-planting is

$1 - s$. Let the individual variety loss function be defined by the following piece-wise linear function

$$L_j(s, Q) = \begin{cases} s(Q - \hat{Q}_j) & if\ Q \geq \hat{Q}_j \\ (1 - s)(\hat{Q}_j - Q) & if\ Q < \hat{Q}_j \end{cases} \tag{22}$$

where Q is the desired level of quasi-variety value. If this level is equal to or more than some level $\hat{Q}_j$ (representing the over-planting of a variety), then the social planner loses $s(Q - \hat{Q}_j)$. If the level is less than this (representing under-planting a variety), the social planner loses $(1 - s)(\hat{Q}_j - Q)$. The social planner seeks to minimize the sum of losses across all varieties by planting varieties that give the maximum value Q. The objective function is therefore

$$\sum_{j=0}^{J} L_j(s, Q) = s \int_0^Q \left(Q - Q_j\right) dF\left(\hat{Q}\right) + (1 - s) \int_Q^\infty \left(Q - Q_j\right) dF\left(\hat{Q}\right) \tag{23}$$

We posit and provide a proof for the following theorem:[22]

Theorem 1: The predicted socially optimal share of farmers planting a variety is equivalent to its proportion of the quasi-variety value in the market.

Proof: Using the Leibniz rule,[23] we calculate the first order condition and solve for Q

$$\begin{aligned} \frac{\partial L(s, Q)}{\partial Q} = s \int_0^Q (1) dF\left(\hat{Q}\right) + (Q - Q)(1) - (0 - Q)0 + (1 \\ - s) \int_Q^\infty (-1) F\left(\hat{Q}\right) + (\infty - y)0 - (Q - Q)1 = 0 \end{aligned} \tag{24}$$

[22] This result has also been proved by Weitzman (2015) in an application concerning externalities. The general geometric methodology of this novel idea (use of weights or penalties in vertex and edge optimization) was also discussed at length by Lancaster (1971) in Chapter 5 and section 7.4. The proof we provide is motivated by a proof for the absolute error loss by Giles (2012).

[23] Let $\omega(\alpha) = \int_{u_1}^{u_2} f(x, \alpha) dx$ where $a \leq \alpha \leq b$. Then,

$\frac{\partial \omega}{\partial \alpha} = \int_{u_1}^{u_2} f(x, \alpha) dx + f(u_2, \alpha) \frac{\partial u_2}{\partial \alpha} - f(u_1, \alpha) \frac{\partial u_1}{\partial \alpha}$ (Giles,2012)

$$s\int_0^Q dF(\hat{Q}) = (1-s)\int_Q^\infty dF(\hat{Q}) \tag{25}$$

Note that $\int_0^Q dF(\hat{Q}) + \int_Q^\infty dF(\hat{Q}) = \int_0^\infty dF(\hat{Q}) = 1$. Thus, we get the following equality

$$s\int_0^Q dF(\hat{Q}) = (1-s)(1-\int_0^Q dF(\hat{Q}) \tag{26}$$

$$s\int_0^Q dF(\hat{Q}) = 1-\int_0^Q dF(\hat{Q}) - s + s\int_0^Q dF(\hat{Q}) \tag{27}$$

$$\int_0^Q dF(\hat{Q}) = 1-s \tag{28}$$

Therefore, by definition, $\Pr(\hat{Q}_j > Q) = s$. Thus, we expect that share of farmers adopting particular varieties will be equivalent to the share of the cumulative quasi-varietal value relative to the total quasi-variety value. QED.

This powerful idea is similar to Bonnet and Simioni (2001) who computed inverse demand functions as a function of WTP. When a new product is introduced into the market, we can simply recalibrate the total market value using the new set characteristics embodied in the variety and to determine how much of the share of overall market value will accrue to the new good. Table 3-6 shows the results of applying this heuristic in this instance. It is apparent that the varietal market shares predicted by this algorithm are almost exactly the same as the actual shares (e.g., the hybrid variety SC 403 has the exact share, 20%).

Assuming a new variety (Table 3-6, last row) is being proposed with all the desirable biological characteristics (early maturing, flint, highest yield potential, and drought tolerant), we can use our methodology to determine the ex-ante predicted rate of adoption. The last column in Table 3-6 shows the predicted shares of each of the varieties following the introduction of the hypothetical new variety. It is apparent that the new variety will take 37% of the market share, displacing the local variety (37% to

21%) and the variety SC403 (20% to 12%). Notice that this superior variety does not fully displace the local variety–an important finding considering Malawi's seed policies that aim to rapidly reduce the market presence of local varieties. This is the case because the perceived value of the local variety is still high and that farmers need many years to learn about the variety.

Table 3-6: Predicted shares of varieties after introduction of a new variety

Variety	Price	Variety age	Days to maturity	Flint	Yield, t/ha	DT	Quasi-variety value	Predicted shares	Actual shares	Diff.	New Share
SC403	320	14	100	1	4.5	1	122.25	0.20	0.20	0.00	0.12
SC627	320	13	125	1	9	0	86.89	0.14	0.13	0.01	0.09
DK8033	342	10	112.5	0	8	1	51.29	0.08	0.08	0.00	0.05
DK8053	342	5	125	1	10	0	14.62	0.02	0.08	-0.05	0.01
MH18	365	22	125	1	6	0	53.82	0.09	0.06	0.02	0.06
SC719	400	4	130	0	11.5	1	50.74	0.08	0.05	0.03	0.05
PAN53	365	5	137.5	1	9	1	2.32	0.00	0.03	-0.03	0.00
DK9089	342	3	117.5	1	10	0	30.20	0.05	0.04	0.01	0.03
Hybrid all							412.14	0.67	0.67	0.00	**0.42**
Local							202.99	0.33	0.33	0.00	0.21
New Variety		1	100	1	12000	1	363.19				**0.37**

Note: Flint and DT (drought tolerance) are dummies.

Given the predicted shares and ownership matrix, we can also determine the market power implications by computing the price-cost margins (PCM) for each seed company using equation 17 (Table 3-7). In the current market, Seed Co has market power almost two times that of the other two major competitors (Monsanto and Pannar)[24]. If Seed Co is the one introducing the new variety, its PCM increases from 0.42 to 0.64. If the new variety is introduced by Monsanto however, its PCM increases from 0.16 to 0.47 while that of the market leader, Seed Co, decreases from 0.42 to 0.27. Similarly, with Pannar introducing the variety its PCM increases from 0.09 to 0.43. These results are contingent on the sample of varieties and the characteristics used in this study and are thus only illustrative of the substantive real-world varietal market implications of introducing new crop varieties.

Table 3-7: Price-cost margins under introduction of a new hybrid variety.

		PCM if new variety is produced by		
Company	Current market PCM	SeedCo	Monsanto	Pannar
Seed Co	0.42	0.64	0.27	0.27
Monsanto	0.16	0.10	0.47	0.10
Pannar	0.09	0.06	0.06	0.43

3.4.5 Limitations and future research

The characteristics space analysis laid out in this chapter allows for the direct estimation of the willingness to pay for varietal traits, the prediction of demand for new varieties, and their market power and welfare implications. Efforts to investigate the adoption of new varieties that delve deeper than the simple improved versus local variety dichotomy raise concerns as to the extent to which farmers can really identify the actual variety they planted or whether the phenotypes reported by seed companies are legitimate. This has been documented in the emerging literature on field based DNA fingerprinting of maize varieties (see, for example, Wossen et al. 2019) where discrepancies between reported varieties and the actual genetic make-up of these varieties have been found.

[24] Sutcliffe (2014) reports that there nine seed companies that operated in 2011 and documents the nature of seed market competition.

Future research should address these concerns by reaching beyond the names of the varieties (and their derived traits) to the phenotypes arising from the actual genetic make-up of the varieties. The other concern, and possibly an opportunity for further research, is that most of the varieties are produced by multinational companies which operate under different sets of agricultural policies across countries. Understanding the varietal adoption and related welfare implications of varietal development and releases policies using the characteristics space model proposed opens up exciting and socially valuable lines or research regarding agricultural innovation policy. Lastly, the novel approach of predicting adoption of new varieties we have developed can be used at a more disaggregated spatial resolution, thereby better matching a new variety to the locale where its embodied agronomic and economic benefits will be more fully realized.

3.5 Conclusion

This chapter investigates the willingness to pay for maize varietal traits among smallholder farmers in Malawi. The maize seed industry is key to food security in Malawi, and the country's seed policies have huge impacts on the growth trajectory of the national economy. Given that the adoption of improved variety seeds is increasing particularly because of the generous Malawi farm input subsidy program, it is important to consider the quality attributes of the improved varieties and match these to the quality needs of farmers operating under heterogeneous production conditions. In addition, the multiplicity of varieties makes it highly questionable to treat improved varieties as if they were homogenous products. This implies that determinants of the adoption of improved varieties may be misleading if one relies on binary choice data (e.g., improved versus local varieties) since specific characteristics (or traits) may be accounting for uptake performance of a particular variety, traits that may be shared across certain varieties be they improved or local. A characteristic based approached to the analysis of varietal adoption provided more nuanced and operationally valuable insights into the determinants of adoption. Specifically, we find that farmers are willing to pay more than two times for a drought tolerant and flint variety for a higher yielding (1,000 kg of maize per ha) variety— an important result when yield benefits are a key goal of a breeding program.

The previous literature on heterogeneous varietal demands focused on farmer ratings of varietal characteristics. In do so they ignored the most fundamental aspect of the

Lancaster/characteristics space framework, namely that the characteristics are best defined in an objective, technical and invariant fashion, setting aside (likely variable) farmer centric perceptions of varietal characteristics. Doing so has the advantage that breeders and seed companies can focus on enhancing the development of a portfolio of varieties that embody a clearly defined set of varietal traits whose deployment can be targeted to particular production locales and market realities in ways that enhance the overall market value of these new varieties.

4. Calibrating Inter-District Food Crop Flows in Malawi

"... information on trade flows within a country is rarely available to researchers, yet the response of these trade flows to a transportation infrastructure improvement says a great deal about the potential for gains from trade."

Donaldson (2018, p.900).

4.1. Introduction

The quest for smart and well targeted agricultural development policies in sub-Saharan Africa has never been more important than now due to ever increasing financial demands on government budgets. In this quest, the effect that policies have on the flow of agricultural produce among districts within a country is usually neglected or poorly understood, with much of the emphasis being placed on the international trade implications of such policies. Understanding how food flows within a country is an important food policy issue, not least because of the likely spill-ins and spill-outs of impacts across regions/districts within a country as agricultural interventions are brought to scale. Absent effective spatial trade and price transmission information, substantial regional deficits or surpluses may emerge, even when nationally produced (or accessible) supply can ostensibly meet the needs of all the households within a country at prevailing average prices. Therefore, inter-district trade is an important part of an effective food security strategy, especially considering that most sub-Saharan African countries still protect their food sectors especially during deficit years (Myers 2013). Inter-district commodity trade data are, however, not collected on a systematic basis and thus unavailable in most sub-Saharan African (SSA) countries and many parts of world. In this chapter, we use a theoretical model introduced by von Thünen to guide the development of a mathematical programming model to calibrate food flows among districts in Malawi.

Almost two centuries ago, Johann Heinrich von Thünen, a German farmer-cum-economist developed a stylized theory of the effect of distance and transportation costs on the spatial structure of farm organization around a main city (Hall 1966). The theory predicts that with increasing distance from an isolated town, land will progressively be given up to products that are cheaper to transport in relation to their value. For this reason alone, von Thünen posits that sharply differentiated concentric rings will form around a town, with production in each ring being dominated by a particular agricultural product. Thus based on von Thünen’s insights, we know that locations for commodity production are not random. Thus because spatial mathematical programming models rely on spatially specific production and consumption data to calibrate flows, we can confidently use calibrated food flows in making policy decisions, e.g., fuel subsidies to reduce transport costs.

Beyond the theoretical justifications, Thünen like patterns in the spatial organization of economic activities have been observed in several developing countries. Jacoby and Minten (2009) report that for the least remote households in Madagascar, as much as 87% of crop sales by weight are either fruits, vegetables or tubers. Conversely, for remote households as much as 95% of crop sales by weight are high valued crops per kilogram of dry grains (mainly rice, maize and beans). Similar patterns have been reported for Nepal by Jacoby (2000) and Fafchamps and Shilpi (2003), and for the Democratic Republic of Congo by Minten and Kyle (1999). The benefits of reducing transport costs in sub-Saharan Africa on both agricultural productivity and welfare are well researched, mostly by development economists though without basing their analyses on von Thünen’s work (e.g. Fafchamps, Gabre-Madhin, and Minten, 2005 for Benin, Madagascar and Malawi; Jacoby and Minten, 2009 for Madagascar; Gollin and Rogerson, 2014 for Uganda; and Adam et al. 2018 for Tanzania).

These studies show how the spatial variation in transport costs affect different outcomes including incomes, prices, migration and structural transformation. For example, Fafchamps, Gabre-Madhin and Minten (2005) used trader survey data to investigate the presence of increasing returns to agricultural trade. Regarding transport costs, they concluded that for the Malawian and Madagascar cases that they studied, unit transport cost could be reduced simply by organizing larger loads. Jacoby and Minten (2009) use the canonical agricultural household model and novel cross-sectional data from Madagascar to examine the impact of a plausibly exogenous variation in transport costs on incomes. They find that a road that essentially eliminated transport costs would boost the incomes of the remotest households by nearly half, mostly by raising non-farm earnings. Gollin and Rogerson (2014) develop a multi-sector multi-region general equilibrium model to examine the effect of transport costs and changes in agricultural productivity on household welfare in Uganda. They find that high transportation costs lead to high food prices in cities, thereby restricting migration from rural to urban areas and exacerbating subsistence farming in rural areas.

Adam et al. (2018) expanded the spatial general equilibrium model in Gollin and Rogerson (2014) to analyze the effects of public investment programs—including those that reduce transport costs—on household welfare in Tanzania. In relation to transport costs, they find that interventions that directly reduce transport costs accelerate processes of structural transformation, making the economy more tradable in the process though this depends on market power of the transporters. They however rely on stylized assumptions— regions are defined functionally with no actual location on the map and urban regions do not produce any agricultural commodities— which are at best questionable since location matters in the distribution of economic activities and there is as much agricultural production in the peripherals of the urban areas as in the rural areas.

Globally, benefits of reducing transport costs mostly through transport infrastructure like roads and railroads have been reported in India (Donaldson 2018) and the United State (Costinot and Donaldson 2016), where the introduction of rail roads is reported to have reduced famine and increased agricultural production. Nonetheless, these studies did not analyze directly the effect of transport cost reduction on the volumes that are traded across districts understandably because data on flows across districts are not available in most countries. At the international level, transport costs have been found to be a key determinant in international trade flows and therefore gains to trade (Venables and Limao 2002).

The estimation of trade elasticities and share of home traded goods are at the pinnacle of the new international trade theories. Only a few studies (e.g., Donaldson 2018 for India) have applied these models to intra-national trade because these trade models require prior estimates of trade flows which are unavailable for most countries. Donaldson (2018, p.900) summarizes this food policy and data problem, *"information on trade flows within a country is rarely available to researchers, yet the response of these trade flows to a transportation infrastructure improvement says a great deal about the potential for gains from trade".* The analysis of intra-country trade flows is important for making food policies because of the likely spill-overs across space and time. For instance, to understand whether a farm subsidy for rural poor improves nutritional outcomes, we need to know the quantity of food that moves across the rural-urban space. In addition, trade flows may alleviate impacts of weather shocks like droughts and floods (see for example Burgess and Donaldson 2010 in India).

We relax some assumptions (e.g. uniform fertility and single town) in von Thünen's theory and propose that a price endogenous nonlinear mathematical programming model (following Chen and Onal, 2012) is a general modelling framework for implementing what von Thünen laid out. The advantage of this model is that the optimal solution includes calibrated volumes of produce traded. We implement the spatially specific mathematical programming model for six major food crops (maize, rice, cassava, potatoes, beans and groundnuts) in Malawi— a largely rural agricultural country with high domestic transport costs. The model is calibrated for each of the 27 districts using 10 years of crop mix data to avoid overspecialization with supply and demand schedules referenced for 2009/10 agricultural season. We implemented the model using the commercial version of GAMS (General Algebraic Modelling System).

The calibration results for the base year show that 7% of produced quantities of maize flow across districts as compared with 66% for rice, 74% for beans, 46% for groundnuts, and zero for cassava and potatoes. The results of the chapter show that instead of concentric rings, there exists "arrows" of product flows across the different separated but not isolated districts that reflect the spatial shadow price differences in relation to transport costs. A simulation experiment shows that a reduction in the per unit cost of transport nonlinearly increases by a small margin the share of production traded. This points to alternative investment and policy decisions mostly towards rural road infrastructure that enhances the urban connectedness of rural areas.

This chapter makes two main contributions to this literature. First, this chapter is the first in Malawi to calibrate trade flows of food crops within the country and thus provide a benchmark for future agricultural policy making in Malawi. Second, the chapter links a causal theory of the effects of distance (and transport costs) on the spatial organization of economic activity by von Thunen to mathematical programming models.

The rest of the chapter is organized as follows. The policy context motivating the modelling approach is discussed in subsection 4.1.1. The model is discussed in section 4.2 followed by a description of the data and agricultural statistics used in the calibration in section 4.3. The results of the calibration are presented in section 4.4.

Section 4.5 provides the discussion of the results with respect to von Thünen spatial economic theory and their implications for food policy followed by concluding remarks in section 4.6.

4.1.1 Food policy context

The pursuit of self-sufficiency in food is evident in both historical and current food policies of Malawi. The key priority of Malawi's post-independence era agricultural policy has been to attain food self-sufficiency through production of own staple foods. On the production side, Malawi has been implementing a farm input subsidy program for staple food production since 2005/06. On the trade side, Malawi occasionally invokes import and export bans whenever the country experiences crop production surplus and deficits respectively. In addition, the government annually recommends minimum farmgate prices, and supports a grain marketing board and food reserve agency (Pauw and Edelman 2015). While many economists espouse the principles of free trade and its benefits, these principles are often not heeded by governments and policy makers for both political and practical economic reasons. This position is unlikely to change in the foreseeable future. In this chapter, we take a neglected "*intra-country, inter-district*" view of the Malawi economy. Assuming a closed staple food sector economy in the case of Malawi is plausible because although Malawi is an occasional importer and exporter of maize (the main staple food), the internationally volumes are typically less than 6% of production. There is virtually no recorded trade in the other key staples like cassava and sweet potatoes because of low value-bulk ratio (Minot 2010a). A much lower percentage (2%) for traded maize is reported by Adam et al. (2018) for Tanzania.

We rely on the fact that heterogeneity within a country can generate free-trade-like benefits across the different districts or regions of the country that would result in a second-best food policy under self-sufficiency. Benson et al. (2016) similarly argues that the implication of farming diversity in Malawi is that the comparative advantage of different areas of the country for production of different crops, livestock and other agricultural products differs significantly from place to place. To achieve these district or regional comparative advantages, targeted policies are a necessity. Targeting of agricultural policies to the poor and across geography nonetheless has been a difficult task among policy makers and economists in developing countries including Malawi because agricultural statistics are either sparse or not timely. For the purposes of the sector model, Malawi is disaggregated into 27 districts as shown in the Figure 4-1 below. The district is a useful spatial unit of analysis for calibrating the model because most agricultural and development statistics are available at this level. In addition, Malawi's decentralization efforts are targeted at making the district as the central point of policy discourses with structures and powers that were initially in the purview of the national government.

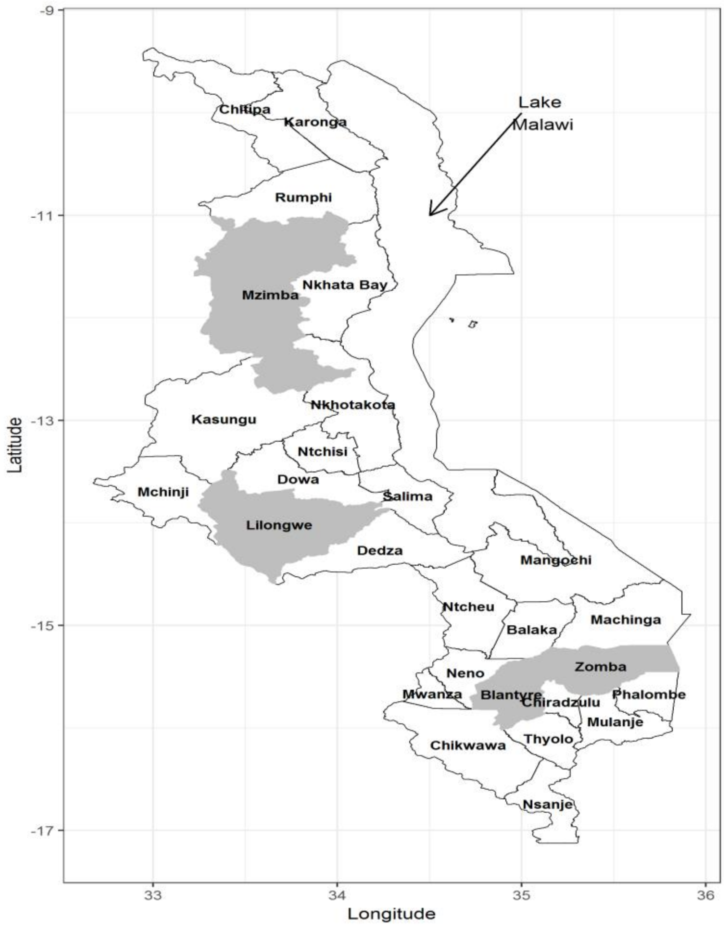


Figure 4-1: Districts of Malawi

Notes: Shaded districts (Lilongwe, Mzimba, Zomba and Blantyre) are city districts. Lilongwe is the capital city.

Maize is the most important food crop in the country, followed by cassava, potatoes and Sorghum (Minot, 2010a). Precisely, per capita consumption of maize is 133kg and it accounts for about 54% of caloric intake of households in Malawi (Table 4-1). Using nationally representative survey data for 2005/06, Ecker and Qaim (2011) report that on average, more than 60% of the total food quantity consists of staple foods, primarily maize. Maize accounts for 46% of total food quantities, more than 60% of energy, and almost half of protein consumption. It is also the source for 67% of total iron, 65% of total zinc, and almost 70% of total riboflavin consumed. In selecting the food crops to include in the model, we were guided by the long-term importance of the food crop to the Malawian population. Therefore, the major food crops- Maize, Rice, Cassava and Potatoes were chosen. We however included other food crops that offer other major nutrients apart from calories available through the staple food crops. Two pulses (i.e., Beans and groundnuts) were thus also included in the model.

Table 4-1: Food consumption (kg/person/year) in urban and rural areas

Food Crop	**Verduzco-Gallo, Ecker, Pauw (2014)**				**Minot (2010) based of FAO 2009 Food balance sheet**	
	Urban		Rural		National	Share of caloric intake (%)
	2004/05	2010/11	2004/05	2010/11	2009	
Maize	144	159	154	177	133	54
Rice	13	16	4	5		
Cassava	15	9	20	15	89	7
Potato	22	39	16	19	88	8
Beans	10	10	9	7		
Groundnuts	4	6	10	6		

Studies on market equilibrium models are comparatively sparse, especially in Malawi. Some of the previous market equilibrium models in Malawi include; Mapila, et al. (2013) which focused on the maize sector, and Simler (1997) and Kachulu (2018) on multiple crops. Though Simler looked at more than 10 major crops, the analysis was at the national level (without spatial consideration) and is now dated, with major changes in the Malawian economy since this work was published over two decades ago. Another related strand of research analyses the spatial flows of agricultural commodities. There are also only few articles as shown in Table 4-2 that attempted to estimate spatial flows usually for maize. The reason for this is the lack of data and the difficulty in collecting data from traders who for practical and strategic reasons may not be willing to publicly share their volume of operations.

Specific efforts that collected spatial flows in the context of Malawi include Gabre-madhin et al. (2001) and FEWSNET (2014, 2018).[25] These efforts relied on trader surveys and expert opinion to assess the direction of flows and in the case of Gabre-madhin et al. (2001) both the direction and volume of flows (Table 4-2). The estimates from these studies are not updated and though required are not given attention in the data collection efforts by the government or its development partners. This lack of knowledge affects decision making in that it is difficult to target production and consumption policies to where they would be most effective.

Table 4-2: Spatial flows and market analysis studies in Malawi

Study	Period studied	Data and Methods	Crops and total volumes	Share of traded produce (national)
		Panel a: Spatial flows studies		
Gabre-madhin et al. (2001)	2001	Trader survey	Maize, rice, beans and pulses, soybeans	Not given
FEWSNET (2014)	2009	Expert opinion	Maize	Not given
FEWSNET (2018)	2018	Expert opinion	Maize, pulses in Southern region	
Jayne et al. (2010)	2009	Trader survey	Maize	12.9%
Haggblade, Longabaugh, and Tschirley (2009)	2009	Mapping of administrative and survey data	Food staples (Maize, Cassava)	12.2%
Myers (2013)	2001-2008	Spatial cointegration models	Maize	Not given
		Panel b: Market analysis studies		
Mapila, et al. (2013)	2010	Agricultural Statistics and linear programming	Maize	Not given
Kachulu (2018)	2010	Malawi Agricultural Sector Model	Cassava, Cotton, Groundnuts, Maize, Paprika, Rice, Sorghum, Soybean, Sugarcane, Tobacco	Not given

[25] Famine Early Warning Systems Network, an initiative led by the United States Agency for International Development (USAID).

4.2. Model

4.2.1 Von Thünen spatial economic theory

To estimate inter-district trade flows without having to observe any prior estimates- because they are unavailable in Malawi anyway; we need both a coherent theory and empirical model that link observable variables like production, consumption and factor endowments to trade flows across locations. In this section, we discuss the proposition of using von Thünen spatial economic theory and mathematical programming as its empirical apparatus. Several studies have questioned the economic theory behind the use of spatial sector programming models because these models aggregate data at sector and regional levels yet much of the economic theory is at the level of an individual economic agent (Wiborg et al. 2005). However, this criticism applies to all other modeling alternatives, including computable general equilibrium (CGE) and econometrically estimated gravity models and so we set it aside.

We propose that it is useful to think of a spatial sector programming model as an empirical procedure for relaxing and implementing von Thünen's spatial economic theory. Figure 4-2 shows this theory where the net land rent, $\pi_i = P_i - R_i(TC)$ is a

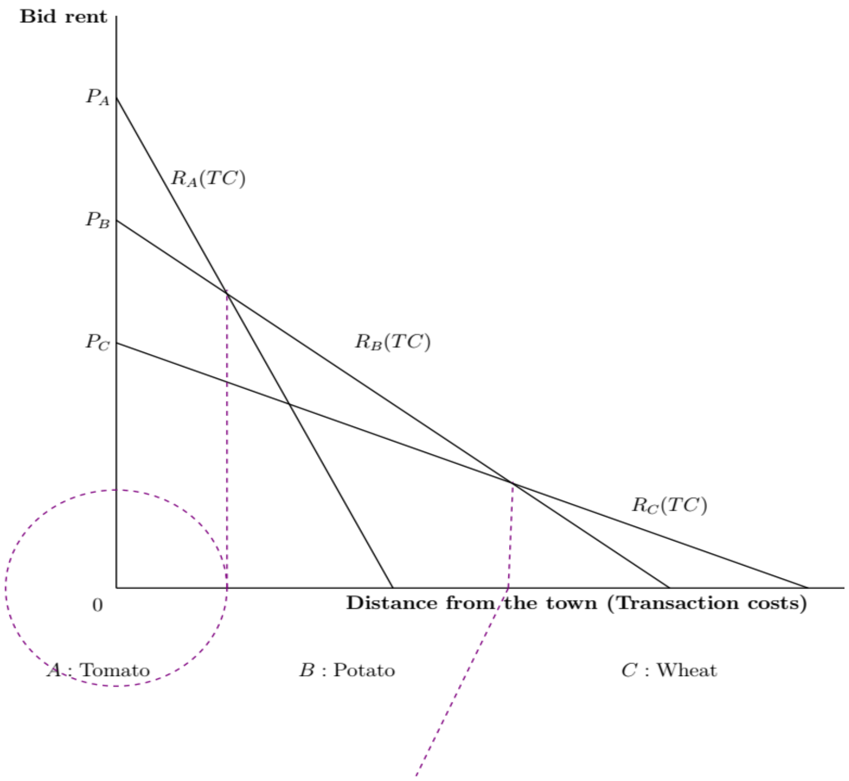


Figure 4-2: Land rent profile and von Thünen rings (Source: adapted from Fujita 2012)

function of land value (price) for each land profile, P_i (note: $i \in \{A, B, C\}$) and land profile transport costs given by $R_i(TC)$. At the center of the town, crops difficult to transport (e.g., tomato represented as A) will be grown because $P_A > P_B > P_C$. As we move away from the main town, the transport costs will be higher and it will be valuable to grow crop B.

The link between von Thünen theory and mathematical programming was first proposed by Stevens (1968). Stevens suggested that a formal "one-shot" model of von Thünen's theory should be of the general mathematical programming format since such a format would explicitly determine which activities (crops) appear in an optimal solution and which of the large set of inequalities forms the set of equations to be solved for equilibrium values. Hartwick (1972) later showed that von Thünen theory can be solved using an endogenous price equilibrium model. Paul Samuelson's essay "Thünen at Two Hundred" eloquently pointed in this direction, *"when Thünen perceived what spatial pricing and specialization patterns are admissible under competition, he was anticipating the methods and results of Kuhn-Tucker nonlinear programming"* (Samuelson 1983). These suggestions were however not followed through with empirical data. Drawing from von Thünen's assumptions, we make the necessary changes in assumption to make it the underlying theory for a spatial sector programming model—a regional model rich in structure regarding production and consumption (see Table 4-3 for a comparison of assumptions made by von Thünen and those used in spatial sector programming models).

This approach allows the interpretation of the results in light with the theory and thereby leading to consistent policy recommendations on trade flows. Using heroic assumptions, von Thünen showed analytically and with examples that with increasing distance from the isolated town, the land will progressively be given up to products cheap to transport in relation to their value. For this reason alone, sharply differentiated concentric rings or belts will form around the town, each with its own staple product. It will be shown— with non-linear, multi-district food sector model— that arrows instead of concentric rings will point to the cities with distance and hence transport cost being the key factor in dictating the type of crop being traded.

Table 4-3: Von Thünen spatial economic theory and spatial sector programming models

Assumptions	von Thünen Theory	Spatial sector programming model
1) **Soil fertility**	Uniform across plain leading to constant crop yields	Variation by soil type, region, topography etc. leading to differences in crop yields
2) **Market center**	Single	Multiple markets
3) **Unit transportation cost**	Uniform	Inter-region transport costs can vary by crop. Intra-region transportation costs usually ignored (Hall, et al. 1975)
4) **Production costs**	Uniform	Vary by region and crops
5) **Demand**	Infinitely elastic	Empirical demand estimates

4.2.2 Non-linear, multi-district food sector model

This study is based on three chronologically related economic models- the spatial economic theory proposed by von Thünen almost two centuries ago, the spatial equilibrium models of Samuelson (1952) and Takayama and Judge (1971), and sector programming models as introduced by McCarl (1982). Mathematical programming sector models have been widely used in developed countries to predict the impacts of changes in public policy, technology and infrastructure as well as in the general economic conditions on an agricultural economy and to evaluate alternative policy choices (Apland and Andersson 1996). Alternative approaches to estimating trade flows usually used in international trade include gravity models (Limao and Venables 2001), co-integration models (Myers 2013), input-output regional models (Uribe, de Leeuw and Theil 1966) and trader surveys. The advantage for using spatial sector programming models is that these models can help generate some nonexistent estimates that can then be used in future for improving the collection of agricultural statistics. This advantage motivates the use of the model since Malawi lacks agricultural statistics on inter-district trade flows of food crops. The major challenge of using these models is that they are data intensive, some of which may not be readily available in a developing country. In this study, we show how compromises can be made to make the model operational for policy making with readily available data

We assume more than two regions (or districts) trading in more than two homogenous commodities. Each district constitutes a single and distinct market which is separated but not isolated by a transportation cost. In addition, districts within the sector model allow for differences in available technology, resource supplies, and product demands (Apland and Andersson 1996). The key assumptions for the model include; competitive behavior for the participants and districts; and no legal restrictions to limit the actions of arbitragers in each district. These assumptions are plausible for Malawi because as reported by Myers (2013), spatial price transmission and seasonal price patterns in private sector maize markets in Malawi are generally consistent with long-run competitive inter-regional trade.

Using the assumptions stated, we can set up the net benefit function or net quasi-welfare maximization problem for a static, multi-region, multi-crop, non–linear programming model of the food sector in Malawi. The model captures the market equilibrium by maximizing economic surplus subject to market clearing and land allocation constraints. Consider the following nonlinear programming model of a closed food sector with a set of regions, Ω_G; set of multiple products, Ω_Y; set of multiple variable inputs,Ω_{ZV}; and a set of production activities in region g, Ω_{XG}:

$$\begin{aligned}\text{Maximize: } W = & \sum_{g\in\Omega_G}\sum_{i\in\Omega_Y}[a_{gi}Y_{gi} + 0.5b_{gi}Y_{gi}^2] \\ & - \sum_{g\in\Omega_G}\sum_{k\in\Omega_{ZV}}\left[c_{gk}Z_{gk} + 0.5d_{gk}Z_{gk}^2\right] - \sum_{g\in\Omega_G}\sum_{h\in\Omega_G,h\neq g}\sum_{i\in\Omega_Y} ty_{ghi}TY_{ghi} \\ & - \sum_{g\in\Omega_G}\sum_{h\in\Omega_G,h\neq g}\sum_{k\in\Omega_{ZV}} tz_{ghk}TZ_{ghk}\end{aligned} \quad \textbf{(29)}$$

Subject to:

$$Y_{gi} - \sum_{j\in\Omega_{XG}} e_{gij}X_{gj} + \sum_{g\in\Omega_G,h\neq g} TY_{ghi} - \sum_{g\in\Omega_G,h\neq g} TY_{hgi} \leq 0 \quad \forall\, \mathrm{g} \in \Omega_\mathrm{G}, \mathrm{i} \in \Omega_\mathrm{Y} \quad \textbf{(30)}$$

$$\sum_{j\in\Omega_{XG}} v_{gkj}X_{gj} - Z_{gk} + \sum_{g\in\Omega_G,h\neq g} TZ_{ghk} - \sum_{g\in\Omega_G,h\neq g} TZ_{hgk} \leq 0; \quad \forall\, \mathrm{g} \in \Omega_\mathrm{G}, \mathrm{k} \in \Omega_\mathrm{ZV} \quad \textbf{(31)}$$

$$Y_{gi}, X_{gj}, Z_{gk}, TY_{ghi}, TY_{ghk} \geq 0; \forall\ g \in \Omega_G; i \in \Omega_Y, j \in \Omega_X, k \in \Omega_Z; h \in \Omega_G, h \neq g \quad \textbf{(32)}$$

$$X_{gj} \leq \sum_{t=2000}^{2009} \Phi_{\mathrm{gt}} X_{gjt} \quad \textbf{(33)}$$

$$\sum_{t=2000}^{2009} \Phi_{\mathrm{gt}} \leq 1 \tag{34}$$

where

g is the region/district;

Y_{gi} is the quantity demanded of product i in district g;

Z_{gk} is the quantity supplied of variable input k in district g;

X_{gj} is the level of production activity j (area of land under crop j) in district g;

TY_{ghi} is the quantity of product i shipped from district g to district h;

TZ_{ghk} is the quantity of variable input k shipped from district g to district h;

e_{gij} is the output of product i per unit of production activity j in district g (or yield coefficient);

v_{gkj} is requirement of variable input k per unit of production activity j in district g;

ty_{ghi} is transport cost from district g to district h per unit of product i;

tz_{ghk} is transport cost from district g to district h per unit of variable input k;

Φ_{gt} is the endogenous weight for historical crop mixes.

The market demand function for product i in region g, in quantity dependent form is, $P_{gi} = a_{gi} + b_{gi}Y_{gi}$. The related terms in the objective function, $a_{gi}Y_{gi} + 0.5b_{gi}Y_{gi}^2$, are demand function integrals. The market supply function for variable input k in region g, in quantity dependent form is, $R_{gk} = c_{gk} + d_{gk}Z_{gk}$. The related terms in the objective function, $c_{gk}Z_{gk} + 0.5d_{gk}Z_{gk}^2$, are input supply function integrals. The constraint in equation 30 represents the product balance where total use of each product is restricted to its total supply. The constraint in equation 31 is the input balance constraint which restricts the use of input k in region i to its availability. The constraint in equation 32 is the usual non-negativity requirement constraint for all the endogenous variables. The constraint in equation 33 and 34 represent convexity restrictions on historical crop mixes. This is discussed in the next section.

4.2.3 Crop mixes in the food sector model

The use of aggregate level supply responses instead of individual supply response functions in the sector model has its caveats. There may be discrepancies for the following reasons: (i) details on production are typically much less in a sector model than in individual farm models, (ii) sector models typically ignore market factors like product differentiation and quality, and (iii) transaction costs are often omitted (Wiborg et al. 2005). In addition to these sources of aggregation bias, extreme specialization in mathematical programming models is also not consistent with observed production patterns. There are two main approaches of dealing with the aggregation bias and extreme specialization: positive mathematical programming (PMP) and crop mix approach. For trade flows calibration, the PMP approach requires prior knowledge of crop flows (Paris et al. 2011) which are not available in the case of Malawi.

The crop mix approach was introduced by McCarl and Spreen (1980) to reduce the potential aggregation biases. The crop mix approach also prevents extreme crop specialization (Apland & Andersson, 1996). This approach restricts the crop mix to the space spanned by a convex combination of historical crop mixes. The main assumption when using this approach is that there is a duality between solving an aggregate model with the full detail of all the farm firm models included on the one hand, and on the other building an aggregate model without the farm firm models which is constrained to the production possibility set spanned by a convex combination of all possible optimal solutions of the farm firm models (Wilborg, et al. 2005; Merel and Howitt 2014).

There are two important deficiencies when using the historical crop mixes. Firstly, the use of historical crop mixes does not constitute as rich a production possibility set, as one would have the full detail in a model. Historical crop mixes are reflections of producer decisions in the face of prevailing prices. Thus, the crop mixes will not be an accurate representation either if the expected prices confronted by the model are outside the historical range or if the situation to be examined substantially revises the production possibilities. Second, the approach does not take account of changes in production costs, inputs and yields when crop mixes change. Several extensions considering these have been made. These include; supplementing the historical crop mixes with expert information or survey information and in a recent study, Chen and Onal (2012) suggested combining historical crop mixes with synthetic crop mixes that are based on acreage response elasticity. The justification for the modification is that though historical crop mixes may be valid when simulating farmer's planting decisions under normal conditions, they may be too restrictive for future land uses.

In this chapter, we use historical crop mixes[26] because during this decade Malawi experienced both worst droughts in 2001/02, 2004/05 and 2007/08 with extreme price spikes, normal, and bumper harvests particularly after the 2005/06 agricultural season. For example, due to poor harvest, maize prices rose in 2001/02, 2004/05 and 2007/08 by 354%, 218% and 395% respectively (Ellis and Manda 2012). Despite these price spikes, the allocation of land to the various food crops has remained stable as can be seen in figure 4-3.

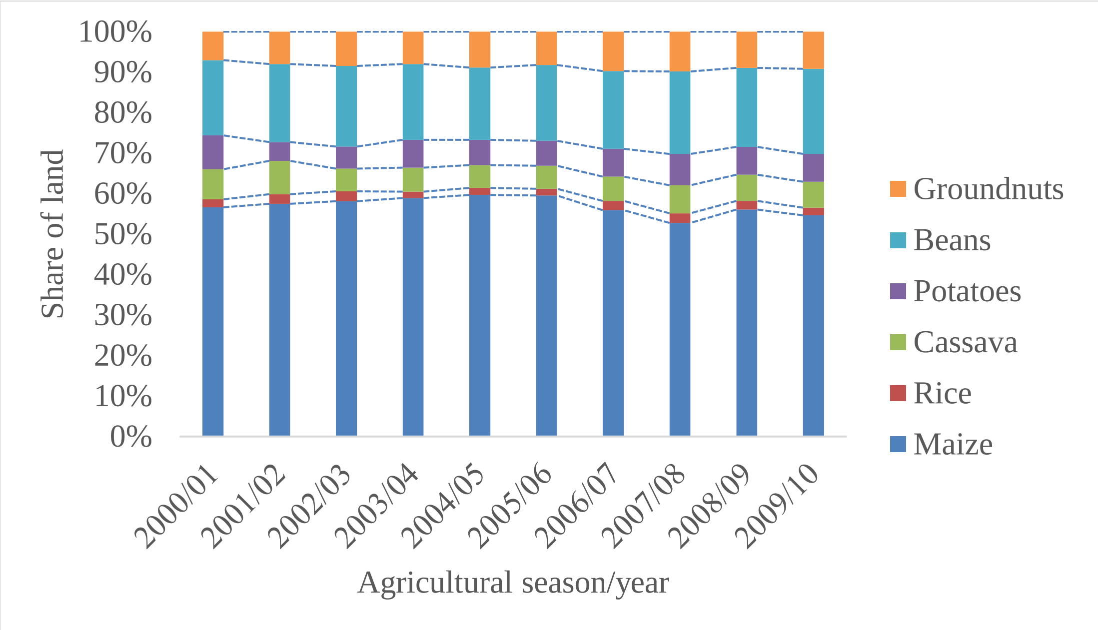


Figure 4-3: National crop mix, 2000/01-2009/10

4.2.4 Analytical results

The endogenous variable that lies at the core of this study is TY_{ghi}, the quantity of a crop i that is traded across regions g and h. The study seeks to determine its share to total production under different perturbations of per unit transport cost, ty_{ghi}. To understand how inter-district trade flows are estimated using the model, we show the Kuhn-Tucker-Karush (K-T-K) conditions corresponding to the inter-district product trade flow variable,TY_{ghi}. Let the Lagrange multiplier associated with the product

[26] The crop mixes are incorporated in the model using convexity constraint which is part of the input constraints as shown in the partial tableau in the appendices.

balance constraint (2) for product i in region g be λ_{gi}. Then K-T-K conditions for the product variables are:

$$\frac{\partial L}{\partial TY_{ghi}} = -ty_{ghj} + \lambda_{hi} - \lambda_{gi} \leq 0; \forall g \in \Omega_G, h \in \Omega_G, h \neq g, i \in \Omega_Y \tag{35}$$

$$\begin{aligned} &\sum_{g\in\Omega_G} \sum_{h\in\Omega_G, h\neq g} \sum_{i\in\Omega_Y} \left[\frac{\partial L}{\partial TY_{ghi}}\right] TY_{ghi} \\ &\qquad = \sum_{g\in\Omega_G} \sum_{h\in\Omega_G, h\neq g} \sum_{i\in\Omega_Y} [-ty_{ghi} + \lambda_{hi} - \lambda_{gi}] TY_{ghi} = 0 \end{aligned} \tag{36}$$

$$TY_{ghi} \geq 0; \forall\ g \in \Omega_G, h \in \Omega_G, h \neq g, i \in \Omega_Y \tag{37}$$

$$\frac{\partial L}{\partial \lambda_{gi}} = -Y_{gi} + \sum_{j\in\Omega_{XG}} e_{gij} X_{gj} + \sum_{h\in\Omega_G, h\neq g} TY_{hgi} - \sum_{h\in\Omega_G, h\neq g} TY_{ghi} \geq 0; g \in \Omega_G, i \in \Omega_Y \tag{38}$$

$$\begin{aligned} &\sum_{g\in\Omega_G} \sum_{i\in\Omega_Y} \left[\frac{\partial L}{\partial \lambda_{gi}}\right] \lambda_{gi} \\ &\qquad = \sum_{g\in\Omega_G} \sum_{i\in\Omega_Y} [-Y_{gi} + \sum_{j\in\Omega_{XG}} e_{gij} X_{gj} + \sum_{h\in\Omega_G, h\neq g} TY_{hgi} \\ &\qquad - \sum_{h\in\Omega_G, h\neq g} TY_{ghi}]\, \lambda_{gi} = 0 \end{aligned} \tag{39}$$

$$\lambda_{gi} \geq 0; g \in \Omega_G, i \in \Omega_Y \tag{40}$$

The set of K-T-K conditions can be interpreted based on whether inter-district trade occurs. If no trade takes place, by conditions [38-40], consumption within the region, Y_{gi}^*, is equal to endogenous crop supply or production, $\sum_{j\in\Omega_{XG}} e_{gij} X_{gj}^*$. If district g engages in inter-district trade of food crop i, then $\sum_{h\in\Omega_G, h\neq g} TY_{hgi} > 0$ or $\sum_{h\in\Omega_G, h\neq g} TY_{ghi} > 0$, indicating excess demand or excess supply, respectively. The arbitrage condition holds, so that, by conditions [35-37], the prices of the food crop in region g and h will differ by at most the unit cost of transportation between the two regions, $\lambda_{gi}^* - \lambda_{hi}^* \leq ty_{ghi}$.

In the simulations, we demonstrate how varying the unit cost of transportation affects the ratio of inter-district trade volumes to total production. This value will differ across crops and regions thereby resulting in a flurry of von Thünen arrows. Following Donaldson (2018), we present three qualitative results that a simple inter-district model would produce.

Result 1 [Von Thunen Arrows of Flows]: Reducing transport costs across all districts will increase total flows for all commodities that were flowing under the baseline. The flows will increase the most for crops that use the mobile factors the most intensively. Mathematically, we are claiming that $\frac{\mathrm{dTY}}{\mathrm{dty}} < 0$.

Result 2: Share of traded volumes and trade elasticity are sufficient statistics for making welfare comparisons of the effect of transport costs.

Result 3: Reducing transport costs increases total welfare.

The complexity of the model makes it difficult to derive envelope and implicit function properties for an analytic solution. We thus depend on a mathematical simulation to determine the effects of transport cost changes on food flows.

4.2.5 Computation

The model is calibrated in the commercial version of GAMS (General Algebraic Modelling System) following a structure of Forest and Agricultural Sector Optimization Model for the U.S (Adams et al. 1996) and Minnesota Sector Model documented in Moon et al. (2014). The GAMS code is provided in appendix C and the data matrices are available from the author upon request.

4.3. Data and model inputs

In this section, we provide a detailed description of the data used in the calibration. A sector model is as accurate as the data used for the calibration and careful attention is made to explicitly explain the data assumptions made. The data inputs for the model include the raw food crop prices and quantity demanded, historical crop mixes (hectares under each crop), crop yields, production and marketing costs for each of 27 districts. The model is calibrated and validated using 2009/10 as the reference year. The data are summarized in four broad categories: demand data, production data, transportation data and crop budgets.

4.3.1 Demand

The demand data used in the model included rural and urban own price elasticities for each of the food crops, district level quantities consumed per capita for each of the food crops, district level population and prices of the food crops for 2009/10 agricultural season. The demand functions we use in this study were obtained from a quadratic almost ideal demand system (QUAIDS) elasticities estimated by Ecker and Qaim

(2011). Instead of working out the inverse of the Quadratic Almost Ideal System, we use the own price elasticities, price data and quantity consumed in each district to derive the coefficients for the demand system. The slope coefficients for the district demand equations for each food crop are therefore calculated as: $b_{gi} = \frac{\delta_{gi}\ \bar{Y}_{gi}}{\bar{P}_{gi}}$ where δ_{gi} is the own price elasticity (different for rural and urban districts) and the bars on the variables represent the observed values for prices, P and quantity demanded,Y in each of the districts. The elasticity refers to the percentage change in Y with respect to change in P, but the slope is defined from the inverse demand function so it is $\Delta\text{P}/\Delta\text{Y}$ (Hazell and Norton 1986). The treatment of demand functions in this way implicitly assumes that the demand system in each district is proportional to the rural and urban disaggregated demand systems. According to Hall, et al. (1975), this does not imply that the quantities demanded in a particular district will be proportional to national quantities; price variations between areas will prevent that. Thus, this treatment ignores intra-rural or intra-urban district differences in preferences. The own price and expenditure elasticities were obtained from a 2009/10 study by Ecker and Qaim (2011). Table 4-4 summarizes the elasticities used in the study. The urban elasticities were used for the city districts of Lilongwe, Blantyre, Zomba and Mzuzu (in Mzimba District).

Table 4-4: Expenditure and Marshallian own-price elasticities of food demand among rural and urban households

	Expenditure elasticities		**Own-price elasticities**	
Crop	Rural	Urban	Rural	Urban
Maize	0.948	0.628	-0.877	-0.722
Rice	0.892	0.904	-0.816	-0.959
Cassava	-0.665	0.076	0.618	-1.152
Potatoes	0.712	1.004	-0.770	-1.248
Beans	1.365	0.197	-0.952	0.415
Groundnuts	0.744	0.413	-0.821	-0.013

Source: Ecker and Qaim (2011)

These values were considered acceptable since it is generally known that demand for staple crops is usually inelastic (The World Bank 2008). The positive elasticities for cassava and potatoes are inconsistent with demand theory and therefore have implications on the results for these two crops[27]. The demand quantities were calculated by multiplying the per capital food crop consumption per year as reported in Verduzco-Gallo, Ecker, and Pauw (2014) by the population size in each district from the 2008 Malawi Population Census. In the case of cities, the district and city population were summed. The income coefficient was calculated from the income elasticities reported in Table 4-4, and expenditure per capita calculated by the author from the Integrated Household Survey III data.

4.3.2 Production and input supply parameters

The Malawi food sector model has the following inputs: seeds for each crop (i.e., maize, rice, potatoes, cassava, beans, groundnuts), basal fertilizer, top-dressing fertilizer, pesticides, transport, packaging materials, labor and land. These inputs can be divided into three groups of inputs: (i) exogenously-priced inputs (e.g., seeds, fertilizer, pesticides, packaging, and transport), (ii) available in fixed supply (e.g., land), and endogenously determined (e.g., labor). For exogenously-priced inputs, a unit cost entry is made directly in the objective function. For such inputs, the implicit supply function is infinitely elastic and the supply function integral is linear (Apland and Anderson 1996). We thus included prices of the inputs as the intercept and a zero slope in the input supply equation. We obtained the price information from crop budgets provided by the Malawi's Ministry of Agriculture and Food Security.

Labor was assumed to be endogenously priced because labor use in smallholder agriculture in Malawi is largely family labor with under 10% of the total labor use being hired labor (Takane 2008); casual labor is very common. The data used on wage rates for labor use, labor requirement and available labor were obtained from Ministry of Agriculture and Food Security Crop Budgets for 2010 and were consistent with survey evidence from Takane (2008). The estimated labor supply elasticity was assumed to be (0.15) using the experimental results reported by Goldberg (2016). The only input available in fixed supply is food crop land. Crop land was therefore mapped to specific districts as land types and restricted using convexity restrictions. The crop production quantities, area cultivated and yields in each of the districts were collected from the

[27] Note that even using plausible values for demand elasticities the input and output data for cassava and potatoes are of poor quality and there is still no consensus on productivity levels. In Malawi for example, there are discrepancies among international databases, ministerial data sources and the national household surveys on the production statistics of roots and tubers (Kilic et al. 2018).

Agricultural Statistics Bulletin compiled by the Malawi Ministry of Agriculture and Food Security.

4.3.3 Transportation

Transportation costs from one district to another were computed from district-district distances and per unit per km transportation costs from the literature. We calculated geodesic distances in kilometres from the centroid of one district to another using geosphere package in R (specifically using the distm function). But these did not quite reflect the travel distances. Because geodesic distances do not capture the road infrastructure and practical realities of traveling across district centers, we use distances calculated from road networks instead and also supplemented these with google map API distances for three districts not in the database. The google map distances are the same as the distances calculated from the road networks.

This implies that transport costs take the form of iceberg costs as is standard in trade literature. For domestic routes, Tchale and Keyser (2010) estimate transport cost to be 18.00 Malawi Kwacha (0.129 USD) per ton per kilometre (or equivalently 0.018 Malawi Kwacha per kg per kilometre). This takes a value of 0.0252 USD using the 2016 exchange rate of 1 MWK/0.0014 USD (Guo and Hawkins 2016). Fafchamps and Gabremadhin (2006) using a trader survey estimated transport costs within Malawi as $0.70 per ton per km. Another study by Lall, Wang, and Munthali (2009) estimated using a survey of tobacco truckers that the average unit transport price (per ton, per km) is 228.4 Malawi kwacha from rural areas to the country’s main cities in comparison to 10 and 12 Malawi kwacha per ton per km on routes linking the country to international markets. In the analysis, we assumed the value by Tchale and Keyser (2010) to be the base transport cost. Though, a spatial sector programming does not include intra-district transportation cost, we got estimates of transportation cost for each of the crops from the crop budgets which were added to the production costs.

4.3.4 Crop budgets

The crop budgets used in the study are based on the 2010 Ministry of Agriculture and Food Security gross margin analyses. We verify the input requirements by comparing to the official guide to agricultural production by the Ministry of Agriculture and Food Security. The crop budgets are used to define the input and output coefficients for the model. The crop budgets are at national level but the crop yields are at district level which allows an approximation of agro-ecological comparative advantage of each district to produce a particular crop. We use a 10 year average (2000-2009) as baseline yields in each district.

The flow chart in figure 4-4 provides a list of the data inputs and outputs of the food sector model.

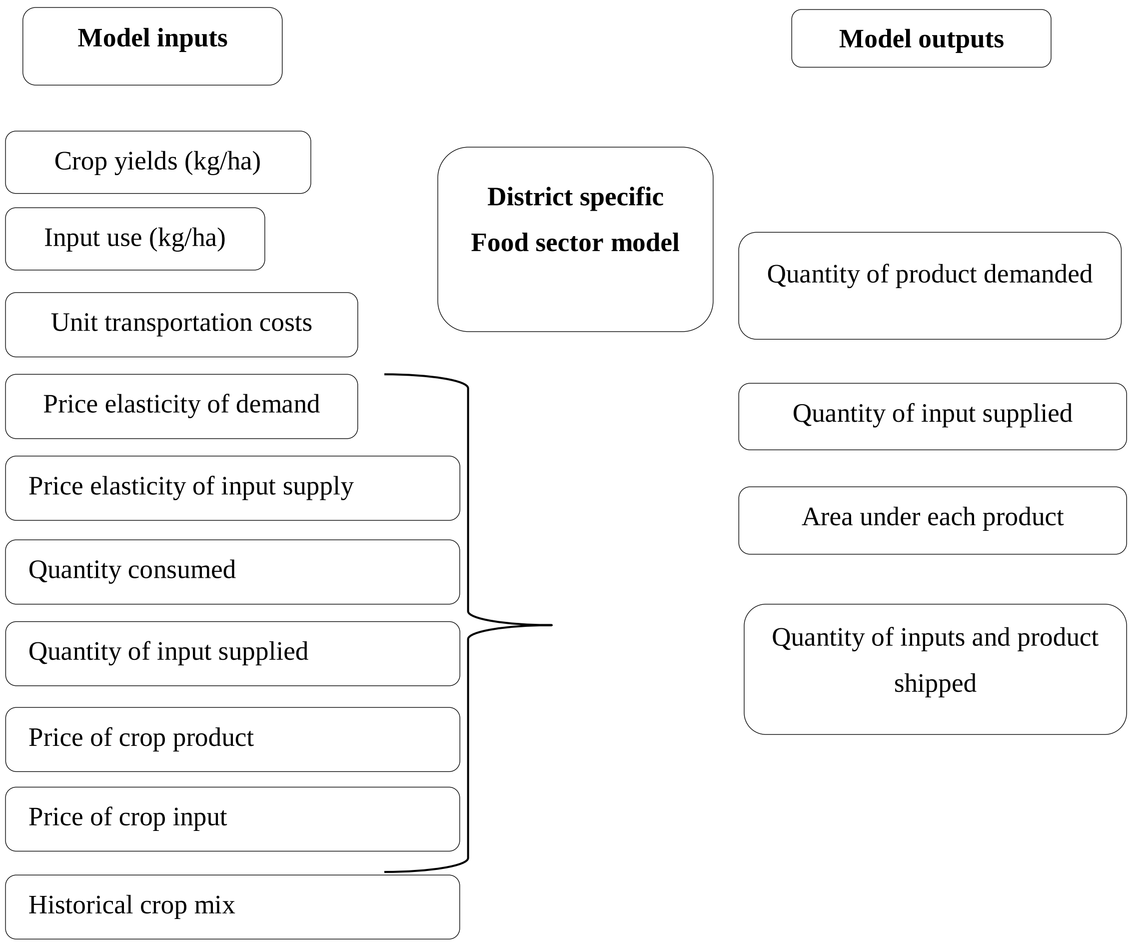


Figure 4-4: Model flow chart

4.4. Calibration results

In the previous two sections, we presented the model and associated data inputs. We now turn to the calibration results starting with the baseline results then transport cost simulations.

4.4.1 Baseline scenario (Scenario 1) and model validation

The common approach of validating a sector model is to compare the model results to observed values of interest. Since we do not have any data on food flows, the model is validated by comparing the simulated land allocation and crop production levels of the 6 food crops with the observed levels in the reference year (2009/10). We also compare the resultant map for maize flows to the FEWSNET market flow map for the corresponding year. In general, the base results are quite different from the observed levels. The table 4-5 shows the calibrated and actual land uses, yields and traded volumes in Malawi.

Table 4-5: Calibrated yields, production and consumption in Malawi (Baseline)

		Units	**Without Restrictions**	**with crop mix**
Maize	Area	Hectares	978,955.18	1059652.85
Maize	Prod	Kilograms	1,472,178,469.24	1557311930.30
Maize	Yield	Kilograms/hectare	1,503.83	1469.64
Maize	Ship-In	Kilograms	128,948,844.85	105367460.87
Maize	Demand	Kilograms	1,472,178,469.24	1557311930.30
Rice	Area	Hectares	98,773.37	59601.97
Rice	Prod	Kilograms	78,411,951.00	78411951.00
Rice	Yield	Kilograms/hectare	793.86	1315.59
Rice	Ship-In	Kilograms	64,351,594.30	51869157.49
Rice	Demand	Kilograms	78,411,951.00	78411951.00
Beans	Area	Hectares	140,137.80	136627.17
Beans	Prod	Kilograms	65,348,009.00	65348009.00
Beans	Yield	Kilograms/hectare	466.31	478.29
Beans	Ship-In	Kilograms	54,485,587.87	48439699.55
Beans	Demand	Kilograms	65,348,009.00	65348009.00
Groundnuts	Area	Hectares	141,539.74	115121.78
Groundnuts	Prod	Kilograms	78,400,476.00	78400476.00
Groundnuts	Yield	Kilograms/hectare	553.91	681.02
Groundnuts	Ship-In	Kilograms	66,278,838.00	35722477.70
Groundnuts	Demand	Kilograms	78,400,476.00	78400476.00

Table 4-6: Calibrated yields, production and consumption in Malawi (Baseline)

Crop	Transport cost Scenario	Share intra-district trade	Share intra-district (with crop restrictions)
Maize	0	10.19	23.19
Rice	0	76.52	81.76
Beans	0	110.27	96.77
Groundnuts	0	91.94	53.71
Maize	9	9.88	8.29
Rice	9	89.51	66.29
Beans	9	93.28	78.40
Groundnuts	9	86.60	46.01
Maize	18	8.76	6.77
Rice	18	82.07	66.14
Beans	18	83.38	74.13
Groundnuts	18	84.54	45.56
Maize	36	6.81	5.85
Rice	36	81.92	65.83
Beans	36	75.59	65.20
Groundnuts	36	79.74	46.05

For the baseline calibration year (2009/10), almost 7% of maize production is traded. This is close to estimates reported in other studies. Because rice, groundnuts and beans are demanded more in urban districts, the share of traded volumes is over 40% of production. The deviances imply a potential problem in the prediction of the economy. The differences between observed and model baseline values can be explained by the data regularities used for building the model. These include the cost of fertilizer. For instance, while the market cost of fertilizer was about 5500 Malawi Kwacha per 50kg in 2009/10, during this period a subsidy program (seeds and fertilizer) was provided to about half of the farming population to Maize implying the cost of producing maize was much lower due to the subsidy as compared to the market rate.

In terms of land allocation, we compared the results from the model to crop suitability maps reported by Benson, Mabiso, and Nankhuni (2016). The suitability maps are consistent with the results of the chapter for all the crops grown.

4.4.2 Transport cost experiments

We consider four transport cost scenarios. These are (i) Scenario 1: Baseline- current per unit transport cost of 18 Malawi kwacha/MT/Km (0.0252 USD), (ii) Scenario 2: Double transportation costs from 18 to 36 Malawi Kwacha/MT/Km, (iii) Scenario 3:

Half transportation costs from 18 to 9 Malawi Kwacha/MT/Km and (iv) Scenario 4: Reduce unit transportation costs to close to zero kwacha per MT per Km. These scenarios capture a range of transport cost changes that may occur under exogenous improvements in infrastructure and changes in fuel costs. Note that for scenario 4, the factor and output price equalization theorem across districts of Samuelson holds.

The figure 4-5 shows that a reduction in transport costs increases the share of traded volumes for all the crops. The increase in the share is however smaller as compared to the rate of transport cost reduction and exhibits nonlinearity. For instance, reducing transport costs by half increases the share of traded volumes of maize from 6.77 to 8.29 percent. Cassava and potatoes do not move across districts in the optimal solution even at this lower transport cost.

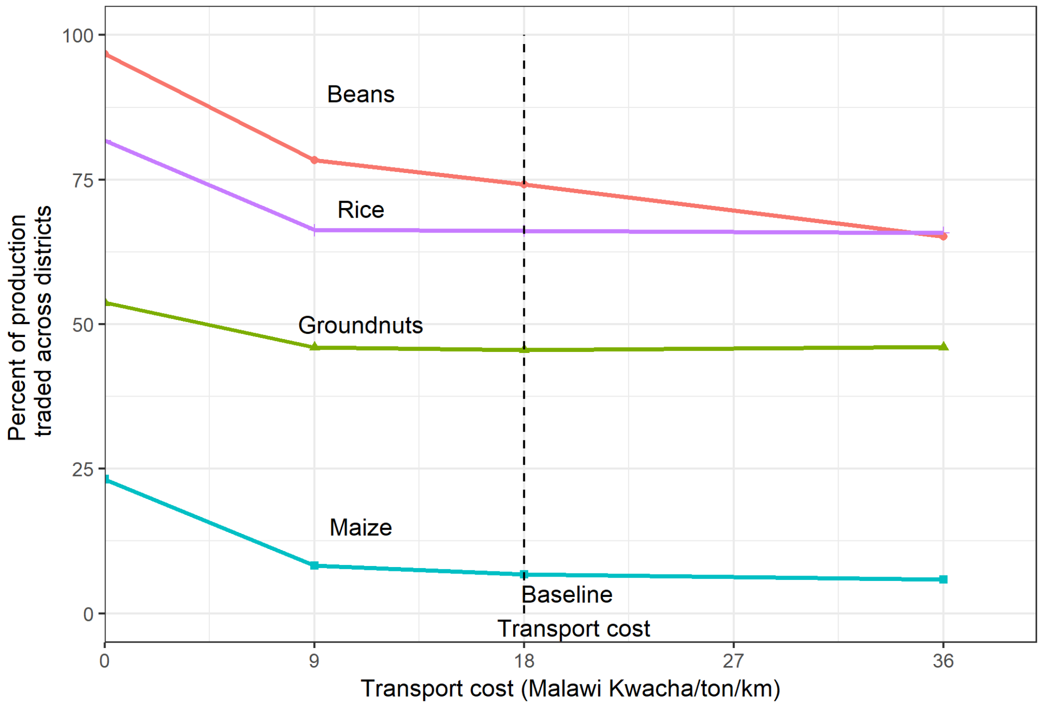


Figure 4-5: Share of production traded across districts under different transport cost scenarios (without cropping restrictions)

According to Minot (2010b), cassava is considered as a poor person's crop as such it is usually consumed within the producing districts. It is therefore expected that in the optimal solution, the share of cassava traded is equal to zero even under low transport costs. The results in the chapter are somewhat different from prior studies on effects of transport costs on internationally traded volumes. According to Minten and Kyle (1999), doubling transport costs can reduce trade flows by around 80%.

In Uganda, Gollin and Rogerson (2014) find that higher transport costs drive up the size of the agricultural workforce and the fraction in subsistence. Donaldson (2018) for India and Donaldson (2019) for USA found that the intra-national estimate of the elasticity of trade flows with respect to distance is close to minus one.

4.5. Discussion

4.5.1 Spatial trade flows in Malawi: a von Thünen interpretation

The map (figure 4-6) shows the baseline year (2009/10) direction but not the volumes of maize trade flows among the different markets in Malawi reported by FEWSNET (Famine Early Warning Systems Network) using expert opinions. To validate the model, we compare the direction of the flows to those shown in figure 4-6. We then discuss the quantities predicted using the model. There are several important distinctions between these flows and what is expected from the model. First, the figure shows bi-directional flows which may be due to differences in the timing of the availability of harvests. In the calibrated model, we only observe the net inter-district flows.

Figure 4-6: Map of observed maize trade flows in a normal year (2009)

Source: FEWSNET (2014). Note: Production and Market Flow Maps provide a summary of experience based knowledge of market networks significant to food security. Maps are produced by USGS in collaboration with other FEWS NET staff, local government ministries, market information systems, NGOs, and network and private sector partners.

It is evident in the figure 4-6 that most southern region districts and districts along the Lake Malawi are maize crop deficit districts. To illustrate this, consider districts that are wholly food insecure in the southern part of the country. These include, Machinga, Thyolo, Chikwawa and Nsanje.

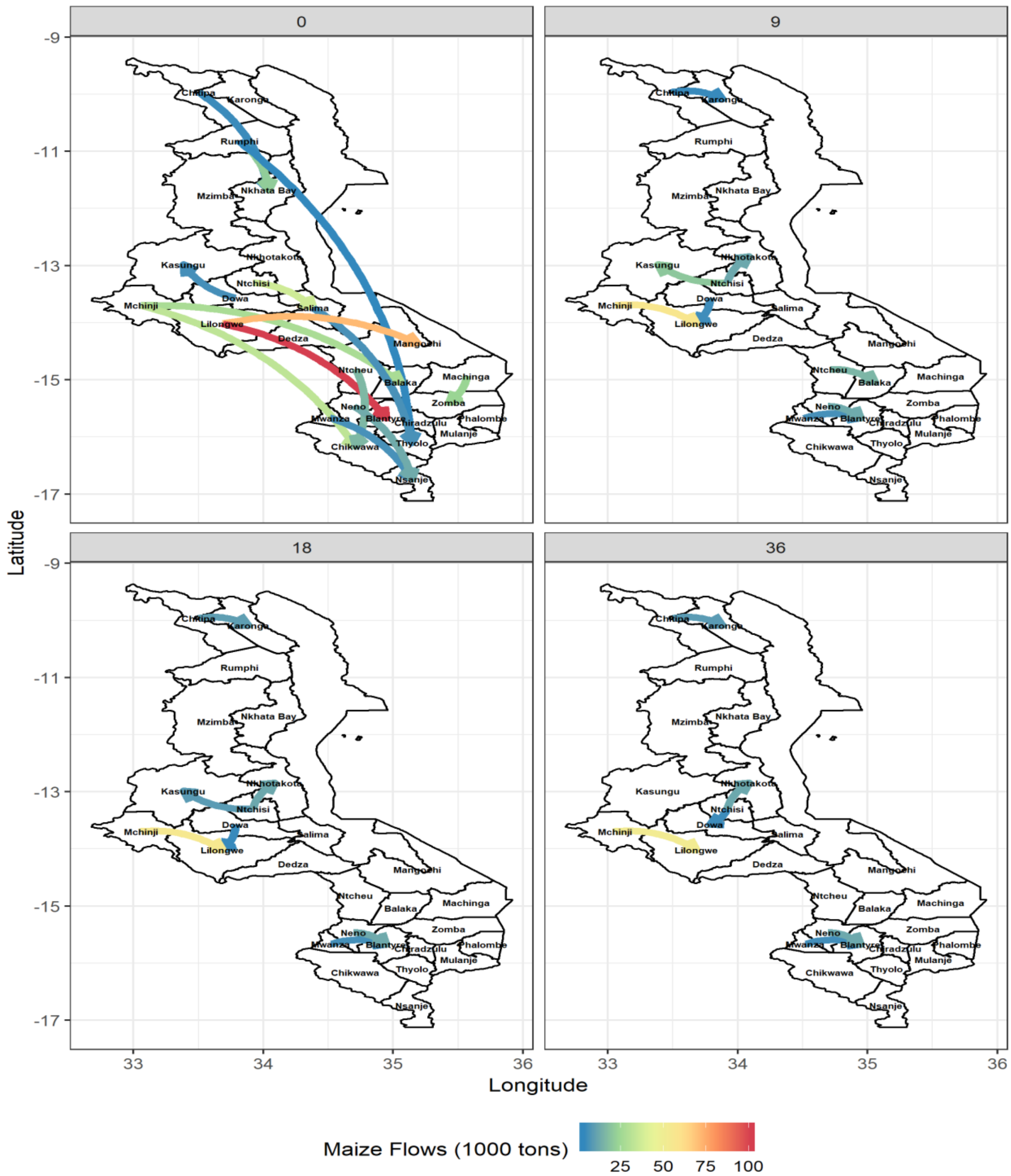


Figure 4-7: Calibrated inter-district maize trade in Malawi showing von Thunen arrows

Note: The flows are in metric tons. The flows for maize are consistent with expert opinion and literature on marketing of maize

The figures 4-6 and 4-7 also illustrate the intuitive and modified von Thünen prediction of flows of agricultural outputs into the main cities of Lilongwe and Blantyre. The closer by districts are acting more as service districts providing maize to these urban areas. It is evident from the maps that in each region there are central district markets that import large flows of food crop commodities consistent with an observation by Mapila et al. (2013). In the case of central region, Lilongwe is the main maize market—

serviced by Mchinji, Kasungu, Dowa and Dedza based on FEWSNET estimates in figure 4-6 and only Mchinji and Dowa based on our calibration results.

In the southern region, Blantyre District is the main market— serviced by Phalombe, Mulanje, Thyolo and Mwanza. In terms of inter-regional flows, the results are consistent with the analysis by Myers (2013) who asserted that major inter-regional maize flows are from the centre to the south, with intermittent flows in both directions between the centre and the north, depending on weather patterns and the season. The maps are generally similar when considered at district level. For instance, Mchinji ships maize to Lilongwe, Ntcheu ships maize to Balaka. It is however difficult to make comparisons in cases where some parts of district are maize insecure while others are not in the FEWSNET map. In the figure 4-6 and 4-7, this is resolved using a district as the main analysis unit. Further research should consider the economics of disaggregating the sector programming model to finer spatial units and thus downscaling any policy interventions to such levels. The disparity between the model results and the map may also be due to the assumption of competitive markets and closed economy. Thus, a model that allows for alternative market structures and international trade may be most appropriate to reproduce the observed levels.

The relative values of traded volumes for the food crops including maize (figures in the appendices) are higher than those reported by Gabre-madhin et al. (2001) for 1998-99 season. They concluded that maize was traded in amounts ranging from 400 to 8000 tons. Rice volumes were smaller ranging from 50 to 1000 tons. Beans/pulses trade ranges from 100 tons to 3000 tons whereas soybeans range anywhere from 10 tons to 5000 tons. This may imply that overtime; the level of traded volumes has been increasing. The flows for rice seem off because it requires a lot of water. A model with a water balance constraint would be appropriate to pin down the optimal rice flows.

The other limitation of the study is that it doesn’t have a nutrition module. The six major crops analyzed provide all the required calories but not enough of the other macro and micronutrients. This should be a key element of future modelling work in this research agenda.

4.5.2 Food and nutrition policy implications

The model presented in this chapter allows the estimation of food crop flows across the districts in Malawi. The policy simulations on reducing per unit transport costs show that cassava is not traded even under much lower transport costs. It is apparent that a reduction in transportation costs alone allows achieving food self -sufficiency for all crops while reducing total land area to food production and at modest levels of current crop yields. Under all transport scenarios, legumes (groundnuts and beans) are traded substantially across districts in optimum. This implies that encouraging rural farmers to grow more legumes does not necessary imply that the farmers will consume the legumes which has implications on nutrition since legumes are the cheaper source of protein in these areas. The demand scenarios are such that the grains are traded from surplus to deficit areas. One important challenge in Malawi's food sufficiency drive is ensuring that food commodities can move from food surplus districts to food deficit districts. To facilitate such economic activities, government and policy makers can make concerted investments in the transportation sector that reduce transport costs with the aim of increasing the flows. The transport costs also capture other related trade costs like bribes that transporters pay to move with good across police road blocks (these inevitably increase the charges the traders have to pay to move good across districts) and other implicit restrictions on volumes of inputs and products.

The cost of transportation is generally an important determinant of farm production decisions and thus, of aggregate land use, which in turn influences trade flows and aggregate welfare. The study found an important realization in terms of maize policy as producers and consumers in each district trade within the district more as compared to inter-district trade thus increasing the connectedness of districts through road infrastructural development cannot necessarily increase the trade flows between districts. Thus, government policy can be more effective if investments are made in rural road infrastructure that makes the rural areas connected to the district capital.

The results in the chapter also provide a caution to researchers who assume large, uniform and linear effects of transport costs reduction on trade. The Malawi case as presented in this chapter shows that the effects are small, non-linear and vary considerably by crop. This chapter has also provided a prototype model using readily available subnational data to guide targeted agricultural development planning in

Malawi and other data sparse countries in sub-Saharan Africa. In addition, the use of a spatial sector programming model can show the improvements required in the collection of agricultural statistics relevant for policy making. Given the importance of trade flows in making credible food policy decisions, it is important that statistical agencies in SSA introduce commodity flow surveys within countries. These can be implemented together with well-established agricultural surveys like the Living Standards Measurement Surveys (LSMS).

4.6. Limitations and future research

As with any other calibration exercise, our results depend on the quality of the underlying data. In this chapter, we have been explicit about the nature and sources of the data used, which we made clear so that our results and conclusions are interpreted within the limitations of the data realities we faced. First, food flows data are not collected and thus not available in Malawi— no one knows how much maize, rice, cassava, potatoes, beans, and groundnuts is traded across districts. This chapter estimates such data under the assumption of perfect competition. In addition, the results that even under low transport costs, cassava and potatoes are not traded across districts runs counter to the common sight of trucks carrying these commodities across different parts of the country. This may be largely due to poor quality of production and consumption data for these crops (see Kilic et al. 2018 for a description of the data quality issues for roots and tubers). Second, for the transport cost data, we assumed a constant per unit cost across crops and used distances based on centroids of districts. Both these assumptions have the potential to be relaxed, with possible important empirical implications. And finally, we did not include water balance and nutrition modules in the current version of the model. In general, these latter two limitations do not affect the qualitative conclusions of the study but remain important in this modeling work. Even with all these limitations, the potential for using spatial programming models in guiding data collection efforts for policy analysis remains huge. Future research agenda on generating "policy relevant" agricultural statistics in Malawi and other SSA countries should focus on introducing intra-national commodity flow surveys. The credibility of policy recommendations regarding the prospective benefits from freer international trade in food trade likely resonate with policy makers if the flow through benefits from enhanced intra-national trade are quantified and reported as national food sector policies are formulated.

References

Adam, C., Bevan, D., and Gollin, D. 2018. "Rural – Urban Linkages , Public Investment and Transport Costs : The Case of Tanzania." *World Development* 109:497-510. http://dx.doi.org/10.1016/j.worlddev.2016.08.013.

Adams, D M., Alig, R J., Callaway, J M., McCarl, B and Winnett, S M. 1996. "The Forest and Agricultural Sector Optimization Model (FASOM): Model Structure and Policy Applications." Research Paper. PNW-RP 495. Portland. https://doi.org/10.2737/PNW-RP-495.

Adesina, A.A., and Zinnah, M.M. 1993. "Technology characteristics , farmers perceptions and adoption decisions : A Tobit model application in Sierra Leone." *Agricultural Economics* 9: 297–311.

Allenby, G.M., and Rossi, P.E.1998. "Marketing models of consumer heterogeneity." *Journal of Econometrics* 89(1–2): 57–78. http://www.sciencedirect.com/science/article/pii/S0304407698000554.

Anderson, Jock R. 1992. "Difficulties in African agricultural systems enhancement? Ten hypotheses." *Agricultural Systems* 38: 387-409.

Angrist, Joshua D., Hull, Peter D., Pathak, Parag A., and Walters, Christopher R. 2017. "Leveraging lotteries for school value-added: testing and estimation." *Quarterly Journal of Economics* :871–919. doi:10.1093/qje/qjx001.

Apland, Jeffrey, and Andersson, Hans. 1996. "Optimal Locations of Processing Plants: Sector Modeling Considerations and an Example." *Review of Agricultural Economics* 18 (3): 491–504.

Arndt, C., Pauw, K., and Thurlow, J. 2016. The Economy-wide Impacts and Risks of Malawi's Farm Input Subsidy Program. *American Journal of Agricultural Economics* 98(3): 962–980. http://ajae.oxfordjournals.org/lookup/doi/10.1093/ajae/aav048.

Barrett, C., and Hogset, H. 2003. *Estimating Multiple-Output Production Functions for the CLASSES Model*. Working Paper.m http://crsps.net/wp-content/downloads/BASIS/Inventoried%203.16/13-2003-8-502.pdf

Beddow, Jason M., Hurley, Terry M., Pardey, Philip G., and Alston, Julian. "Food Security: Yield Gaps." Chapter in Neal K. Van Alfen, ed. *Encyclopedia of Agriculture and Food Systems*. New York: Elsevier, 2014.

Benson, Todd, Conroy, Anne, Gilbert, Rob, Kanyama-Phiri, George, Kumwenda, John, Mann, Charles, Mughogho, Spider, and Phiri, Alexander.1998. *Malawi: Soil Fertility Issues and Options*. Discussion paper.

Benson, Todd, Mabiso, Athur, and Nankhuni, Flora. 2016. "Detailed Crop Suitability Maps and an Agricultural Zonation Scheme for Malawi: Spatial Information for Agricultural Planning Purposes." http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/130494.

Berry, S., and Haile, P. 2016. "Identification in Differentiated Products Markets." *Annual Review of Economics* 8(1): 27–52. Available at: http://www.annualreviews.org/doi/10.1146/annurev-economics-080315-015050.

Berry, S., Levinsohn, J., and Pakes, A.2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of*

Political Economy 112(1):68–105. Available at: http://www.journals.uchicago.edu/doi/10.1086/379939.

Bezu, S., Kassie, G.T., Shiferaw, B., and Ricker-Gilbert, J. 2014. "Impact of Improved Maize Adoption on Welfare on Farm Households in Malawi: A Panel Data Analysis." *World Development* 59: 120-131.

Bhattacharyya, A., and Mandal, R. 2016. *A generalized Stochastic Production Frontier Analysis of Technical Efficiency of Rice Farming : A Case Study from Assam , India*. Sam Houston State University, Department of Economics and International Business. Working paper No. 16-03.

Blow, Laura, Browning, Martin, and Crawford, Ian. 2008. "Revealed preference analysis of characteristics models." *The Review of Economic Studies* 75: 371-389.

Bokosi, J., Musopole, E., and Kabambe, V.H. 2011. *Seed system security assessment: Southern Malawi*, Lilongwe. Available at: http://seedsystem.org/wp-content/uploads/2014/03/SSSA-Southern-Malawi-Final-Report-.pdf.

Bullock, D.S., and Bullock, D.G.2000. "From agronomic research to farm management guidelines: A primer on the economics of information and precision technology." *Precision Agriculture* 2(1): 71–101.

Burgess, R, and Donaldson, D. 2010. "Can Openness Mitigate the Effects of Weather Shocks ? Evidence from India's Famine Era." *American Economic Review: Papers and Proceedings* 100(May): 449–453.

Cabrini, S.M., Irwin, S.H., and Good, D.L. 2010. "Should farmers follow the recommendations of market advisory services? Ahierarchical Bayesian approach to estimation of expected performance." *American Journal of Agricultural Economics* 92(3): 622–637.

Carlin, B.P., and Louis, T.A. 2015. *Bayesian Methods for Data Analysis* Third., London: Chapman and Hall/CRC Press.

Chapman, C., and Feit, E.M., 2015. *R for marketing research and analytics*. Springer, New York.

Chen, Xiaoguang, and Onal, Hayri. 2012. "Modelling agricultural supply response using mathematical programming and crop mixes." *American Journal of Agricultural Economics* 94 (3): 674–686. doi:10.1093/ajae/aar143.

Chib, S., and Carlin, B.P. 1999. "On MCMC Sampling in Hierarchical Longitudinal Models." *Statistics and Computing* 9(1): 17–26.

Chib, S. 1995. "Marginal Likelihood from the Gibbs Output." *Journal of the American Statistical Association* 90(432): 1,313–1,321. Available at: http://www.jstor.org/stable/2291521%5Cnpapers2://publication/uuid/0B20DBD5-06CE-4DC9-91B8-C22D660DBD2A.

Chisama, B.F., Malaidza, H.M., and Bota, S. 2015. "List of agricultural technologies for use in Malawi (approved:2013-15." Department of Agricultural Research Services. Ministry of Agriculture, Irrigation and Water Development.

Chirwa, E., and Dorward, A. 2013. *Agricultural Input Subsidies: The Recent Malawi Experience*. Oxford University Press.

Choi, H., and Oh, I. 2010. "Analysis of product efficiency of hybrid vehicles and promotion policies." *Energy Policy* 38: 2,262-2,271.

Coe, Ric, Njoloma, Joyce, and Sinclair, Fergus. 2016. "To control or not to control: how do we learn more about how agronomic innovations perform on farms?" *Experimental Agriculture* 55 (S1): 303-309.

Costinot,A, and Donaldson, D. 2016. "How Large Are the Gains from Economic Integration? Theory and Evidence from U.S. Agriculture, 1880-1997". NBER Working Paper No. 22946.

Daly, Andrew, Hess, Stephane, and Train, Kenneth. 2011. "Assuring finite moments for willingness to pay in random coefficient models." *Transportation* 39(1): 19-31. Doi:10.1007/s11116-011-9331-3.

Darko, Francis Addeah. 2016. "Essays on Malawian agriculture: Micro-level welfare impacts of agricultural productivity; profitability of fertilizer use; and targeting of subsidy programs." PhD dissertation. Purude University.

Donaldson, D. 2018. "Railroads of the Raj : Estimating the Impact of Transportation Infrastructure". *American Economic Review* 108, 4-5: 899-934.

Dorward, A.R., and Chirwa, E.W. 2015. Crowding out, diversion, and benefit/cost assessments in fertilizer subsidy programs in sub-Saharan Africa: a comment on Jayne, T.S., Mather, D., Mason, N., Ricker-Gilbert, J., 2013. "How do fertilizer subsidy programs affect total fertilizer use in sub-Saharan Africa? Crowding out, diversion, and benefit/cost assessments." *Agricultural Economics* 46: 739–744.

Doss, C.R. 2006. "Analyzing technology adoption using microstudies: limitations, challenges, and opportunities for improvement." *Agricultural Economics* 34: 207-219.

Duflo, E., Kremer, M., and Robinson, J., 2011. Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *American Economic Review* 101(6): 2350–2390. Available at: http://pubs.aeaweb.org/doi/abs/10.1257/aer.101.6.2350.

Ecker, Olivier, and Matin Qaim. 2011. "Analyzing Nutritional Impacts of Policies: An Empirical Study for Malawi." *World Development* 39 (3): 412–28. doi:10.1016/j.worlddev.2010.08.002.

Edmeades, S., and Smale, M. 2006. "A trait-based model of the potential demand for a genetically engineered food crop in a developing economy." *Agricultural Economics* 35(3): 351–361.

Edmeades, S. 2007. "A hedonic approach to estimating the supply of variety attributes of a subsistence crop." *Agricultural Economics* 37(1): 19–28.

Edmeades, S., Phaneuf, D.J., and Smale, M. 2008. "Modelling the crop variety demand of semi-subsistence households: Bananas in Uganda." *Journal of Agricultural Economics* 59(2):329–349.

Ellis, Frank, and Manda, Elizabeth. 2012. "Seasonal Food Crises and Policy Responses: A Narrative Account of Three Food Security Crises in Malawi." *World Development* 40 (7): 1407–17. doi:10.1016/j.worlddev.2012.03.005.

Fafchamps, Marcel, and Shilpi, Forhad. 2003. The spatial division of labour in Nepal. *The Journal of Development Studies* 39(6): 23-66. doi.org/10.1080/00220380312331293577.

Fafchamps, Marcel, Eleni Gabre-Madhin, and Bart Minten. 2005. "Increasing Returns and Market Efficiency in Agricultural Trade." *Journal of Development Economics* 78 (2): 406–42. doi:10.1016/j.jdeveco.2004.10.001.

Feder, G., Just, R.E., and Zilberman,D. 1985. Adoption of agricultural innovations in developing countries: a survey. *Economic Development and Cultural Change* 33(2): 255-298.

Fernandez-Castro, Angel, S, and Smith, Peter, C. 2002. "Lancaster's characteristics approach revisited: product selection using non-parametric methods". *Managerial and decision economics* 23:83-91. DOI: 10.1002/mde.1048.

Fessler, Pirmin, and Kasy, Maximilian. 2018. "How to use economic theory to improve estimators: Shrinking toward theoretical restrictions." *Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_00795.

FEWSNET. 2014. "Production and Trade Flows Maps: Malawi." Famine Early Warning Systems Network. December 22. http://www.fews.net/southern-africa/malawi/production-and-trade-flow-maps/february-2009.

Fezzi, C., and Bateman, I.J. 2011. "Structural agricultural land use modeling for spatial agro-environmental policy analysis." *American Journal of Agricultural Economics* 93(4):1168–1188.

Fousekis, P. 2002. "Distance vs . Ray Functions : An Application to the Inshore Fishery of Greece." *Marine Resource Economics* 17: 251–267.

Fujita, Masahis. 2012. "Thünen and the New Economic Geography." *Regional Science and Urban Economics* 42: 907-912.

Gabre-madhin, Eleni, Marcel Fafchamps, Richard Kachule, Bio Goura Soule, and Zahia Kahn. 2001. "Impact of Agricultural Market Reforms on Smallholder Farmers in Benin and Malawi." IFPRI Collaborative Research Project. Final Report. Volume 2.

Giles, Dave. 2012. "Bayes estimators, loss functions, and J.M. Keynes". Econometrics Beat: Dave Giles Blog, May 11. Accessed June 23, 2019. https://davegiles.blogspot.com/2012/05/bayes-estimators-loss-functions-and-j-m.html.

Goldberg, Jessica. 2016. "Kwacha gonna do? Experimental evidence about labor supply in Malawi." *American Economic Journal: Applied Economics* 8 (1):129-149.

Gollin, Douglas, and Richard Rogerson. 2014. "Productivity, Transport Costs and Subsistence Agriculture." *Journal of Development Economics* 107: 38–48. doi:10.1016/j.jdeveco.2013.10.007.

Government of Malawi, 1997a. *Maize fertilizer recommendations demonstration 1997/98,* Lilongwe: Government of Malawi.

Government of Malawi, 1997b. *.Area-specific fertilizer recommendations for hybrid maize grown by Malawian smallholders: a manual for field assistants,* Lilongwe: Government of Malawi.

Griliches, Zvi. 1957. "Hybrid corn: An exploration in the economics of technological chanage." *Econometrica* 25 (4): 501-522.

Guo, Z., and Hawkins, J. 2016. Assessing travel time and transportation costs in sub-Saharan Africa: A case study in Malawi. , (August). Available at:

http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/130682.

Haggblade, Steven, Steven Longabaugh, and David Tschirley. 2009. “Spatial Patterns of Food Staple Production and Marketing in South East Africa: Implications for Trade Policy and Emergency Response.” East Lansing.

Hall, H.H., Heady, E.O., Stoecker, A., and Sposito, V.A. 1975. "Spatial Equilibrium in US. Agriculture: A Quadratic Programming Analysis." *SIAM Review* 17 (2): 323-338. http://www.jstor.org/stable/2029179.

Hall, Peter. 1966. *von Thunen's isolated state*. 1st edition. Edited by Peter Hall. London: Pergamon Press Ltd.

Harou, A.P., Liu, Y., and Barrett, C.B.2017. “Variable Returns to Fertiliser Use and the Geography of Poverty: Experimental and Simulation Evidence from Malawi.” *Journal of African Economies* 26(3): 342-371.

Hartley, M .J. 1983. “Econometric Methods for Agricultural Supply under Uncertainty: Fertilizer Use and Crop Response.” *Journal of Mathematical Analysis and Applications* 94: 575-601.

Hartwick, J.M.1972. “The pricing of goods and agricultural land in multi-regional general equilibrium.” *The Southern Economic Journal* 39(1):31–45.

Hazell, Peter B.R., and Norton, Roger D. 1986. Mathematical programming for economic analysis in agriculture. Washington D.C: International Food Policy Research Institute.

Heisey, Paul W., and Smale, Melinda. 1995. *Maize technology in Malawi: A Green Revolution in the Making?* CIMMYT Research Report No. 4. Mexico, D.F.: CIMMYT.

Henningsen, G., Henningsen, A., and Jensen, U. 2015. “A Monte Carlo study on multiple output stochastic frontiers: a comparison of two approaches.” *Journal of Productivity Analysis* 44(3): 309–320.

Hess, Stephane, and Train, Kenneth. 2017. “Correlation and scale in mixed logit models.” *Journal of Choice Modelling* 23:1-8.

Holden, Stein, and Mangisoni, Julius. 2013. Input subsidies and improved maize varieties in Malawi: What can we learn from the impacts in a drought year? Norwegian University of Life Sciences (UMB), Centre for Land Tenure Studies Working Paper 07/13.

Hurley, Terrace., Koo, Jawoo., and Tesfaye, Kindie. 2018. “Weather risk: how does it change the yield benefits of nitrogen fertilizer and improved maize varieties in sub-Saharan Africa.” *Agricultural Economics* 49:711-723.

Jacoby, Hanan G., and Minten, Bart. 2009. “On Measuring the Benefits of Lower Transport Costs.” *Journal of Development Economics* 89 (1): 28–38. doi:10.1016/j.jdeveco.2008.06.004.

Jayne, T.S, Nicholas Sitko, Jacob Ricker-Gilbert, and J Mangisoni. 2010. “Malawi’s Maize Marketing System.” Michigan State University. http://fsg.afre.msu.edu/malawi/Malawi_maize_markets_Report_to-DFID-SOAS.pdf.

Jayne, Thom S, Mather, David, Mason, Nicole M, Ricker-Gilbert, Jacob, Crawford, Eric W. 2015. Rejoinder to the comment by Andrew Dorward and Ephraim

Chirwa on Jayne, T. S., D. Mather, N. Mason, and J. Ricker-Gilbert. 2013. "How do fertilizer subsidy program affect total fertilizer use in sub-Saharan Africa? Crowding out, diversion, and benefit/cost." *Agricultural Economics* 46: 745-755. Available at: http://doi.wiley.com/10.1111/agec.12190.

Jayne, Thom S, Mason, Nicole M, Burke, William J, and Ariga, Joshua. 2018. "Review: Taking stock of Africa's second-generation agricultural input subsidy programs." *Food Policy* 75:1-4.

Jayne, Thom S, and Rashid, Shahidur. 2013. "Input subsidy programs in sub-Saharan Africa: a synthesis of recent evidence." *Agricultural Economics* 44: 547-562. Doi: 10.1111/agec.12073.

Jiang, R., Manchanda, P., and Rossi, P.E., 2009. "Bayesian analysis of random coefficient logit models using aggregate data." *Journal of Econometrics* 149(2): 136–148. Available at: http://dx.doi.org/10.1016/j.jeconom.2008.12.010.

Johnson, Michael, E, Edelman, Brent, and Kazembe, Cynthia. 2016. "A farm-level perspective of the policy challenges for export diversification in Malawi: Example of the oilseeds and maize sectors." IFPRI Discussion Paper 01549.

Just, R.E., Zilberman, D., and Hochman, E., 1983. "Estimation of Multicrop Production Functions." *American Journal of Agricultural Economics* 65(4): 770–780. Available at: http://ajae.oxfordjournals.org/content/65/4/770.abstract http://ajae.oxfordjournals.org/content/65/4/770%5Cnhttp://ajae.oxfordjournals.org/content/65/4/770.

Kassie, Girma T, Abdulai, Awudu, Greene, William H, Shiferaw, Bekele, Abate, Tsedeke, Tarekedgne, Amsal, and Sutcliffe, Chloe. 2017. "Modeling preference and willingness to pay for drought tolerance (DT) in Maize in Rural Zimbabwe. *World Development* 94: 465-477.

Kachulu, Mutisungilire. 2018. "Climate change effects on crop productivity and welfare sensitivity analysis for smallholder farmers in Malawi." *African Journal of Agricultural and Resource Economics* 3(1): 58-77.\

Kilic, Talip, Moylan, Heather, Ilukor, John, Mtengula, Clement, and Pangapanga-Phiri, Innocent. 2018. "Root for the tubers: Extended-harvest crop production and productivity Measurement in Surveys." Development Economics, Development Data Group, World Bank. Washington D.C.

Komarek, Adam, M., Drogue, Sophie, Chenoune, Roza, Hawkins, James, Msangi, Siwa, Belhouchette, Hatem, Flichman, Guillermo. 2017. "Agricultural household effects of fertilizer price changes for smallholder farmers in central Malawi." *Agricultural Systems* 154:168-178.

Koop, G. 2003. *Bayesian Econometrics*, West Sussex: John Wiley and Sons Ltd.

Lall, Somik V., Hyoung Wang, and Thomas Munthali. 2009. "Explaining High Transport Costs within Malawi: Bad Roads or Lack of Trucking Competition?" Washington D.C. http://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-5133.

Lancaster, Kelvin J. 1966. "A new approach to consumer theory." *Journal of Political Economy* 74 (2): 132-157.

Limao, N., and Venables, J. 2001. "Infrastructure, Geographical Disadvantage, Transport Costs, and Trade." *The World Bank Economic Review* 15 (3): 451–79.

doi:10.1093/wber/15.3.451.

Löthgren, M. 1997. “Generalized stochastic frontier production models.” *Economics Letters* 57(3): 255–259.

Lothgren, M. 2000. “Specification and estimation of stochastic multiple-output production and technical inefficiency.” *Applied Economics* 32: 1533–1540.

Lusk, Jayson, L, Featherstone, Allen, M, Marsh, Thomas, L, Abdulkadri, Abdullahi, O. 2002. “Empirical Properties of Duality Theory.” *The Australian Journal of Agricultural and Resource Economics*, 46(1): 45–68. Available at: http://www.blackwell-synergy.com/doi/abs/10.1111/1467-8489.00166.

Manski, C.F. 2010. Policy Analaysis with Incredible Certitude. *NBER Working Paper Series*, 121(16207): 261–289. Available at: http://papers.nber.org/papers/W16207.

Mapila, Mariam A.T.J, Kirsten, Johann F, Meyer, Ferdinand, and Kamkwamba, Henry. 2013. A partial equilibrium model of the Malawi maize commodity market. IFPRI Discussion Paper 1254, Washington D.C: International Food Policy Research Institute.

Mason, Nicole, M., and Ricker-Gilbert, Jacob. 2013. “Disrupting demand for commercial seed: Input subsidies in Malawi and Zambia.” *World Development* 45: 75-91. http://dx.doi.org/10.1016/j.worlddev.2012.11.006.

McCarl, Bruce A. 1982. "Cropping activities in agricultural sector models: a methodlogical proposal." *American Journal of Agricultural Economics* 62: 768-772.

Meager, R. 2019. “Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Sevem Randomized Experiments.” *American Economic Journal: Applied Economics* 11 (1):57-91.

Merel, Pierre, and Howitt, Richard. 2014. “Theory and application of positive mathematical programming in agriculture and the environment.” *Annual Review of Resource Economics* 6:451-470.

Minot, Nicholas. 2010a. "Staple food prices in Malawi." Variation in staple food prices: causes, consequence and policy options. Maputo, Mozambique: IFPRI.

Minot, Nicholas. 2010b. “Variation in Staple Food Prices in Eastern and Southern Africa : A Synthesis.” Variation in Staple Food Prices: Causes, Consequence, and Policy Options, no. January: 25–26.

Minot, Nicholas, Kherallah, Mylene, Berry, Philippe. 2000. “Fertilizer market reform and the determinants of fertilizer use in Benin and Malawi.” International Food Policy Research Institute, Markets and Structural Studies Division, Discussion paper No. 40.

Minten, Bart, Koru, Bethlehem, Stifel, David. 2013. “The last mile(s) in modern input distribution: Pricing, profitability and adoption.” *Agricultural Economics* 44: 629-646.

Mkondiwa, Maxwell, Pardey, Philip G, and Chan-Kang. Forthcoming. “A meta review of the conventional crop varietal use evidence for sub-Saharan Africa, 1981-2016.” University of Minnesota.

Moon, Jin-Young, Apland, Jeffrey, Folle, Solomon, and Mulla, David. 2014. "A watershed level economic analysis of cellulosic biofuel feedstock production with consideration of water quality." Staff Paper Series, January: P14-02.

Mviha, P.J.Z., Mtukuso, A.P., and Banda, M.H.P.2010. “A catalogue of agricultural technologies used by farmers in Malawi.” Department of Agricultural Research Services, Ministry of Agriculture and Food Security.

Myers, Robert J. 2013. “Evaluating the Effectiveness of Inter-Regional Trade and Storage in Malawi’s Private Sector Maize Markets.” *Food Policy* 41: 75–84. doi:10.1016/j.foodpol.2013.04.007.

Nevo, Aviv. 2001. “Measuring Market Power in the Ready-to-Eat Cereal Industry.” *Econometrica* 69(2):307–342.

Nevo, Aviv, Rosen, Adam, M. 2012. “Identification with imperfect instruments.” *Review of Economics and Statistics* 94(3): 659-671.

Paris, Quirino, Drogue, Sophie, and Anania, Giovanni. 2011. “Calibrating spatial models of trade.” *Economic Modelling* 28:2509-2516.

Pauw, K., and Edelman, B., 2015. “Is Malawi’s mix of maize market policies ultimately harming food security?”. IFPRI. Malawi Strategy Support Program policy note 22.

Peet, Richard. 1970. "Von Thunen theory and the dynamics of agricultural expansion." *Explorations in Economic History* 181-201. doi:10.1016/0014-4983(70)90009-4.

Pollak, Robert, A. 1989. “*The theory of the cost-of-living index*”. Oxford University Press, Inc.

Rosas, Francisco, Lence, Sergio , Hayes, Dermot J. 2018. “Crop Yield Responses to Prices: a Bayesian Approach to Blend Experimental and Market Data.” *European Review of Agricultural Economics* 1-27. doi:10.1093/erae/jby032.

Rossi, P.E., and Allenby, G.M. 1993. “A Bayesian-Approach to Estimating Household Parameters.” *Journal of Marketing Research* 30(2):171–182.

Saka, A.R., Mtukuso, A.P., Daudi, A.T., Banda, M.H.P., and Phiri, I.M.G. 2006. *Agricultural Technologies Released by the Ministry of Agriculture and Food Security-2000-2005*, Lilongwe. Available at: http://www.dphu.org/uploads/attachements/books/books_2215_0.pdf.

Samuelson, Paul, A. 1952. “Spatial price equilibrium and linear programming.” *The American Economic Review* 42(3):283-303.

Samuelson, Paul, A. 1983. “Thunen at Two Hundred.” *Journal of Economic Literature* 21(4):1468-1488.

Scott, James, C. 1998. *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press.

Sheahan, Megan, and Barrett, Christopher, B. 2017. “Ten striking facts about agricultural input use in Sub-Saharan Africa.” *Food Policy* 67: 12-25.

Simler, Kenneth. 1997. The transition to a market based agricultural economy in Malawi: A multi-market analysis. Discussion Paper, New York: Cornell Food and Nutrition Policy Program.

Smale, M. et al., 2001. “Maize Diversity , Variety Attributes , and Farmers ’ Choices in Southeastern Guanajuato , Mexico.” *Economic Development and Cultural Change* 50(1):201–225.

Smale, M., Heisey, P.W., and Leathers, H.D., 1995. “Maize of the Ancestors and Modern Varieties: The Microeconomics of High-Yielding Variety Adoption in Malawi.” *Economic Development and Cultural Change* 43(2):351-368.

Snapp, S. et al., 2014. *Maize Yield Response to Nitrogen in Malawi ’ s Smallholder Production Systems*, Washington D.C.

Spreen, Thomas H, and Takashi Takayama. 1980. "A Theoretical Note on Aggregation of Linear Programming Models of Production." *American Journal of Agricultural Economics* 62:146-151.

Stevens, Benjamin H. 1968. "Location theory and programming models: the Von Thunen case." *Papers of the Regional Science Association* XXI: 19-34.

Suri, T, 2011. “Selection and Comparative Advantage in Technology Adoption.” *Econometrica* 79(1):159–209.

Sutcliffe, Chloe Angelica Jane. 2014. “Adoption of improved maize cultivars for climate vulnerability reduction in Malawi.” PhD dissertation. University of Leeds.

Takane, Tsutomu. 2008. "Labor use in smallholder agriculture in Malawi: six village case studies." *African Study Monographs* 29 (4): 183-200.

Takayama, Takashi, and George G Judge. 1971. “*Spatial and temporal price and allocation models*.” Amsterdam: North-Holland Publishing Company.

Tchale, Hadwick, and Keyser, John. 2010. Quantitative value chain analysis: an application of Malawi. Policy Research Working Paper, Washington D.C: The World Bank.

Uribe, Pedro., de Leeuw, C.G., and Theil, H. 1966. “The information approach to the prediction of interregional trade flows.” *The Review of Economic Studies* 33(3):209-220.

Useche, P., Barham, B L., and Foltz, J D. 2009. “Integrating technology traits and producer heterogeneity: A mixed-multinomial model of genetically modified corn adoption.” *American Journal of Agricultural Economics* 91(2): 444–461.

Useche, P., Barham, B L., and Foltz, J D.2013. “Trait-based adoption models using ex-ante and ex-post approaches.” *American Journal of Agricultural Economics* 95(2): 332–338.

Vanlauwe, B., Kihara, J., Chivenge, P., Pypers, P., Coe,R., and Six,J. 2011. “ Agronomic use efficiency of N fertilizer in maize-based systems in sub-Saharan Africa within the context of integrated soil fertility management.” *Plant Soil* 339: 35-50.

Vanlauwe, B., Coe, R., and Giller, K.E. 2016. “Beyond Averages: New Approaches To Understand Heterogeneity and Risk of Technology Success or Failure in Smallholder Farming.” *Experimental Agriculture* 55(1): 1–23. Available at: http://www.journals.cambridge.org/abstract_S0014479716000193.

Venables, A J., and Limao, N. 2002. “Geographical disadvantage: a Heckscher-Ohlin-von-Thunen model of international specialisation.” *Journal of International Economics* 58: 239–263.

Verduzco-Gallo, Inigo, Ecker, Olivier, and Pauw, Karl. 2014. “Changes in Food and Nutrition Security in Malawi Analysis of Recent Survey Evidence.” Malawi Strategy Support Program. Lilongwe.

http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/128205.

Ward, Patrick S., Ortega, David L., Spielman, David J., and Singh, Vartika. 2014. “Heterogeneous demand for drought-tolerant rice: Evidence from Bihar, India.” *World Development* 64: 125–139. Available at: http://dx.doi.org/10.1016/j.worlddev.2014.05.017.

Weitzman, Martin.2015. “A Voting Architecture for the Governance of Free-Driver Externalities, with Application to Geoengineering.” *Scandinavian Journal of Economics* 117:1049-

Wiborg, Torben, McCarl, Bruce A, Rasmussen, Svend, and Schneider, Uwe A. 2005. “Aggregation and Calibration of Agricultural Sector Models through Crop Mix Restrictions and Marginal Profit Adjustments.” 2005 International Congress, August 23-27, European Association of Agricultural Economists. Copenhagen, Denmark. http://ideas.repec.org/p/ags/eaae05/24567.html.

Wossen, T., Abdoulaye,T., Alene, A., Nguimkeu, P., Shiferaw, F., Rabbi,I.Y., Haile,M.G., and Manyong,V. 2019. “Estimating the Productivity Impacts of Technology Adoption in the Presence of Misclassification.” *American Journal of Agricultural Economics* 101(1): 1–16. doi:10.1093/ajae/aay017.

Yan, Weikai. 2014. “*Crop Variety Trials: Data Management and Analysis*.” Sussex: John Wiley & Sons.

Zeng, D. 2014. “Three Essays on Adoption and Impacts of Improved Maize Varieties in Ethiopia.” Available at: https://vtechworks.lib.vt.edu/handle/10919/64332.

Zhu, Jessica. 2019. “Heterogeneous farmers' technology adoption decisions: Good on average is not good enough.” Job market paper. University of Wisconsin, Madisson.

Chapter 1 Appendices

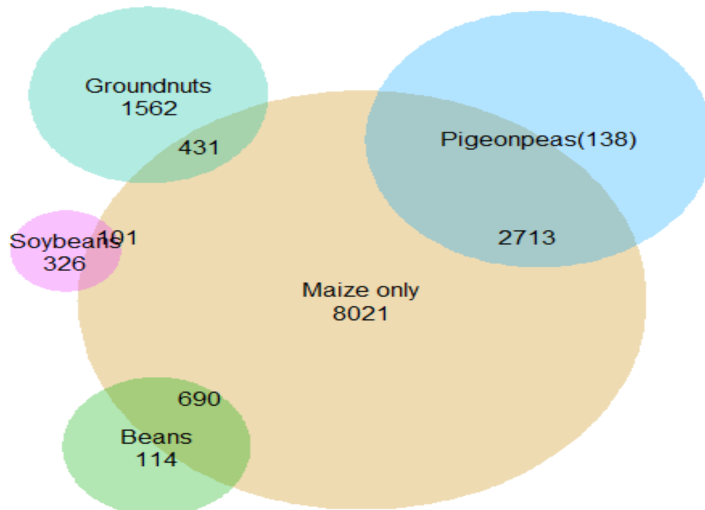


Figure A1: Venn-diagram of maize-legume intercropping in Malawi (numbers are number of plot observations)

Additional Tables

Table A1: Additional descriptive statistics

	Unit	Proportion/mean	Std. Dev
Plot characteristics			
Distance from plot to homestead	Kilometers	0.79	1.20
Plot area (GPS measured)	Hectares	0.44	3.08
Slope	Percentage	5.66	4.96
Rainfall	Millimeters	833.21	80.78
Elevation	Meters	876.33	312.25
Crop stand dummy (Pure stand=1)	Proportion	0.43	0.50
Soil type (Sandy=1)	Proportion	0.20	0.40
Soil type (Between sandy and clay=1)	Proportion	0.59	0.49
Soil type (Clay=1)	Proportion	0.17	0.38
soil type (Other=1)	Proportion	0.04	0.19
soil quality (Good=1)	Proportion	0.45	0.5
soil quality (Fair=1)	Proportion	0.43	0.5
soil quality (Poor=1)	Proportion	0.12	0.32
Household characteristics			
Family labor	Hours/ha/year	856.25	1225.33
Age of household head	Years	43.53	16.29
Household size	Number	4.79	2.18
Poverty quintile	1-10	3.17	1.35
Distance to road	Kilometers	9.86	10.32
Distance to trading center	Kilometers	37.91	20.81
Distance to nearest border	Kilometers	22.82	17.43
Distance to marketing parastatal depot	Kilometers	7.90	5.29
Total expenditure per capita	Malawi kwacha	53154.35	53692.46
Poverty (Poor=1)	Proportion	0.46	0.50
Gender of the household head (Female=1)	Proportion	0.25	0.43
Education level of household head (None=1)	Proportion	0.76	0.43
Education level of household head (Primary=1)	Proportion	0.10	0.30
Education level of household head (Secondary=1)	Proportion	0.13	0.33
Education level of household head (Tertiary=1)	Proportion	0.01	0.11

Table A2: Non-informative prior Bayesian linear model results for mono-cropped maize plots

Panel A: Survey evidence

	Dependent variable: Maize yield (mono-cropped plots only)			
	Posterior median	Posterior Std.Dev	2.5%	97.5%
(Intercept)	1104.00	475.44	181.52	2040.73
Total inorganic N fertilizer use	11.09	0.46	10.21	11.99
Squared (Total inorganic N fertilizer use)	-0.01	0.00	-0.01	-0.01
Organic fertilizer use(No=1)	-154.22	45.22	-243.59	-65.43
Total family labor per ha	0.25	0.01	0.22	0.28
Crop stand(mixed=1)	-238.47	40.42	-315.94	-157.90
Seed use	45.88	40.73	-35.20	124.82
Soil type (Between sandy and clay=1)	135.93	50.68	37.05	236.18
Soil type (Clay=1)	31.15	93.58	-148.73	217.06
Soil type (Other=1)	-125.43	33.01	-189.92	-59.55
Soil quality (Fair=1)	-201.16	51.48	-301.14	-100.92
Soil quality (Poor=1)	-0.32	0.43	-1.17	0.52
Rainfall	0.18	0.10	-0.02	0.38
Elevation	-1.31	3.95	-8.92	6.42
Slope	-7.31	3.72	-14.63	-0.11
Area planted (GPS recorded)	-87.28	37.29	-161.16	-14.83
Household head gender (Female=1)	-0.02	0.95	-1.88	1.80
Household head age	13.66	7.08	-0.21	27.64
Household size	-181.83	33.59	-247.01	-116.28
Poverty(poor=1)	93.42	49.82	-3.71	191.65
Education (Primary=1)	47.56	48.20	-45.14	142.55
Education (Secondary=1)	441.03	115.94	209.87	662.94
Education (Tertiary=1)	-19.55	12.92	-44.57	6.22
Distance from plot to house	1104.00	475.44	181.52	2040.73
District fixed effects	Yes	Yes	Yes	Yes

Panel B. Experimental Evidence

Maize yield				
	Posterior median	**Posterior Std.Dev**	**2.5%**	**97.5%**
(Intercept)	1281.98	130.17	1026.80	1535.99
N fertilizer applied	25.41	1.09	23.29	27.60
N fertilizer applied squared	-0.09	0.01	-0.10	-0.07
Phosphorus	-2.82	1.51	-5.84	0.10
Surphur	49.06	6.66	36.19	62.11
1997/98=1	-310.48	17.40	-344.64	-275.69
MH17 variety	87.56	124.46	-150.00	337.79
MH18 variety	130.91	120.84	-106.32	368.63
Soil texture (medium=1)	162.16	15.85	131.53	193.41
District fixed effects	Yes	Yes	Yes	Yes

Appendix A: Multi-output crop response functions

Perhaps the most difficult problem in estimating non-experimental/observational agricultural production functions is that input data typically are not available by crop (Just, Zilberman and Horchman 1983). As such the separate ordinary least squares or seemingly unrelated regressions may be limited in capturing the crop specific response to fertilizer when the crops are intercropped. In the case of multi-crop farming systems, most researchers therefore use the dual profit or cost function approach. In the recent literature it has been observed that duality does not hold when there is measurement error or when there is less variability in the price data (Lusk et al. 2002) which is the case in the Malawian data. We therefore consider a new approach that was initially introduced by Löthgren (1997) and later expounded by Löthgren (2000), Barrett and Hogset (2003) and Henningsen et al. (2015). Other studies that have used the ray production function approach include Fousekis (2002) in fisheries production and (Löthgren 2000) in rice production.

First consider a single crop, say maize output production function denoted as:

$$\mathrm{f(x)} = \max\{\mathrm{y} \geq 0\colon \mathrm{y} \in \mathrm{Y(x)}\} \qquad \textbf{(A1)}$$

where Y(x) is the output set satisfying all the basic axioms of convexity, closedness and free disposal. This production function can be analyzed using a single regression equation specifying the relation between output and inputs. In the case of Malawi and other African countries, it is common for farmers to grow multiple crops on the same plot during the same growing period. Here, we follow Löthgren (1997) ray production function approach which is based on the polar-coordinate representation of the output vector;

$$\mathrm{y} = \tilde{\mathrm{y}}\mathrm{m}(\lambda) \qquad \textbf{(A2)}$$

where $\tilde{y} = \left[\sum_{j=1}^{p} y_j^2\right]^{0.5}$ denotes the Euclidean norm of the output vector y. The function $m\colon \left[0, \frac{\pi}{2}\right]^{p-1} \rightarrow [0,1]^p$ is defined by $m_i(\lambda) = \cos\lambda_j \prod_{m=0}^{i-1} \sin\lambda_m$ for all $j = 1, \ldots, p$ crops where $\lambda \in \left[0, \frac{\pi}{2}\right]^{p-1}$ and by definition $\sin\lambda_0 = \cos\lambda_p = 1$ represents a transformation of the polar coordinate angle vector to output mix $m(\lambda) = \frac{y}{\tilde{y}}$ with norm $m(\lambda) = 1$. The coordinate angles for the five crops considered in this study can then be recursively determined from the inverse transformation. The challenge though is that the sample of farmers who grow all the crops on the same plot is very small. The sample size may be appropriate for maize-single legume intercrops. As such, we calculated separate yield norms and coordinate angles for maize and each of the legumes as follows

$$\lambda_{\text{maize,legume}} = \arccos\frac{y_{\text{maize}}}{\tilde{y}_{maize,legume}} \text{ and} \tag{A3}$$

$$\lambda_{\text{legume}} = \arccos\frac{y_{\text{legume}}}{\tilde{y}_{maize,legume}\sin\lambda_{\text{maize,legume}}}. \tag{A4}$$

Henningsen et al. (2019) showed that to get robust results, the coordinate angles should not include the recursive sin components. Following this suggestion, we then estimate a single equation production function ($f(x,\lambda)$) with the Euclidean norm of the output vectors as the dependent variable and inputs and the polar output coordinates as independent variables. The advantage of this approach is that in cases where the farmer is growing a single crop, the production function is the same as the traditional single equation production function. In addition, the ray production function is equivalent to the output distance function which makes it consistent with multi-output production economics theory (Lothgren 1997). Figure A2 shows the polar representation of the production possibilities frontier in the maize-legume space.

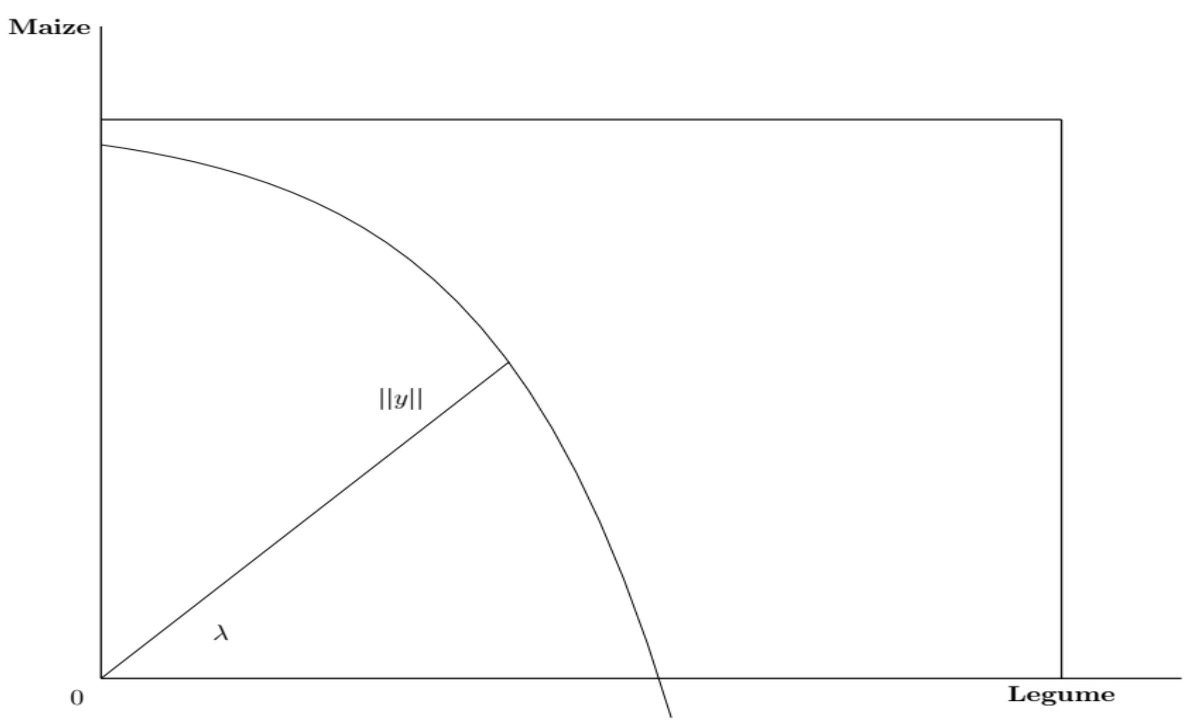


Figure A2: Polar representation of production possibilities frontier in maize-legume output space

Table A3 shows that the coordinate angles estimate are negative. This is consistent with agronomic evidence in Malawi. The fertilizer use estimate for the ray productions are within overlapping credible intervals with the maize -only production function. This is the case because maize is the main crop within each of the multi-crop farming systems analyzed. It is also apparent that standard deviations of the linear part of the crop responses are larger in inter-cropping farming systems than in sole cropping systems. In the sensitivity analyses we considered the model with maize only. Since my main concern is with crop response estimates, we do not discuss the other parameters but only point out that they are all in line with other studies on crop response function estimation in Africa.

Table A3: Bayesian linear model results for maize-legume intercrops

	Dependent variables							
	Maize and bean yield norm		Maize and Groundnuts yield norm		Maize and pigeon peas yield norm		Maize and soya beans	
	Median	Std.Dev	Median	Std.Dev	Median	Std.Dev	Median	Std.Dev
(Intercept)	3691.58	1732.20	-1764.44	1437.42	2918.53	871.50	14601.59	4428.11
Polar coordinate (legume)	-3438.23	994.79	-4.8E+9	4.2E+9	-2736.80	468.36	-625.32	2456.52
Total inorganic N fertilizer use	8.15	1.21	4.81	1.24	3.33	0.44	2.67	5.09
Squared N fertilizer use	-0.01	0.00	-0.00	0.00	-0.00	0.00	0.01	0.02
organic fertilizer use(No=1)	28.24	99.45	-139.32	99.87	20.87	43.63	-634.93	272.07
Total family labor per ha	0.11	0.05	0.17	0.05	0.21	0.01	0.14	0.20
Seed use	0.95	0.28	0.85	0.50	0.51	0.11	33.05	7.38
Soil type (Between sandy and clay=1)	201.19	116.80	42.69	91.65	-39.82	36.47	105.93	400.59
Soil type (Clay=1)	208.86	134.22	44.54	121.88	-33.30	51.05	168.97	489.20
Soil type (Other=1)	-102.15	182.33	83.38	156.28	53.25	81.27	67.58	748.93
Soil quality (Fair=1)	-196.42	77.77	139.27	83.63	-135.14	30.76	236.44	254.04
Soil quality (Poor=1)	6.95	126.85	-175.32	113.40	-211.17	51.14	-1023.78	351.55
Rainfall	-0.05	1.39	2.81	1.43	0.80	0.62	-14.64	3.85
Elevation	0.32	0.23	-0.10	0.27	0.54	0.12	-0.74	1.36
Slope	-14.68	7.07	11.96	10.36	-2.10	3.41	48.40	33.96
Area planted (GPS recorded)	-402.65	142.38	-181.49	110.92	-362.40	48.52	192.48	303.73
Household head gender (Female=1)	-153.59	85.76	-119.20	80.89	-40.69	32.83	592.78	313.52
Household head age	-6.10	2.24	-2.99	2.13	-0.23	0.91	5.93	6.80
Household size	15.25	19.49	2.73	18.85	2.47	7.71	127.68	61.59
Poverty(poor=1)	-55.25	83.75	-188.91	86.29	-33.49	32.48	16.26	231.78
Education (Primary=1)	-320.67	115.14	390.17	123.72	-14.51	50.52	487.35	434.74
Education (Secondary=1)	-41.77	106.41	73.95	120.66	61.69	47.89	770.02	347.30

Education (Tertiary=1)	-175.47	337.24	-181.49	419.33	310.06	252.95	-298.51	192.23
Distance from plot to house	-45.23	32.67	27.78	34.39	2.02	12.31	33.05	7.38
District fixed effect	Yes		Yes		Yes		Yes	

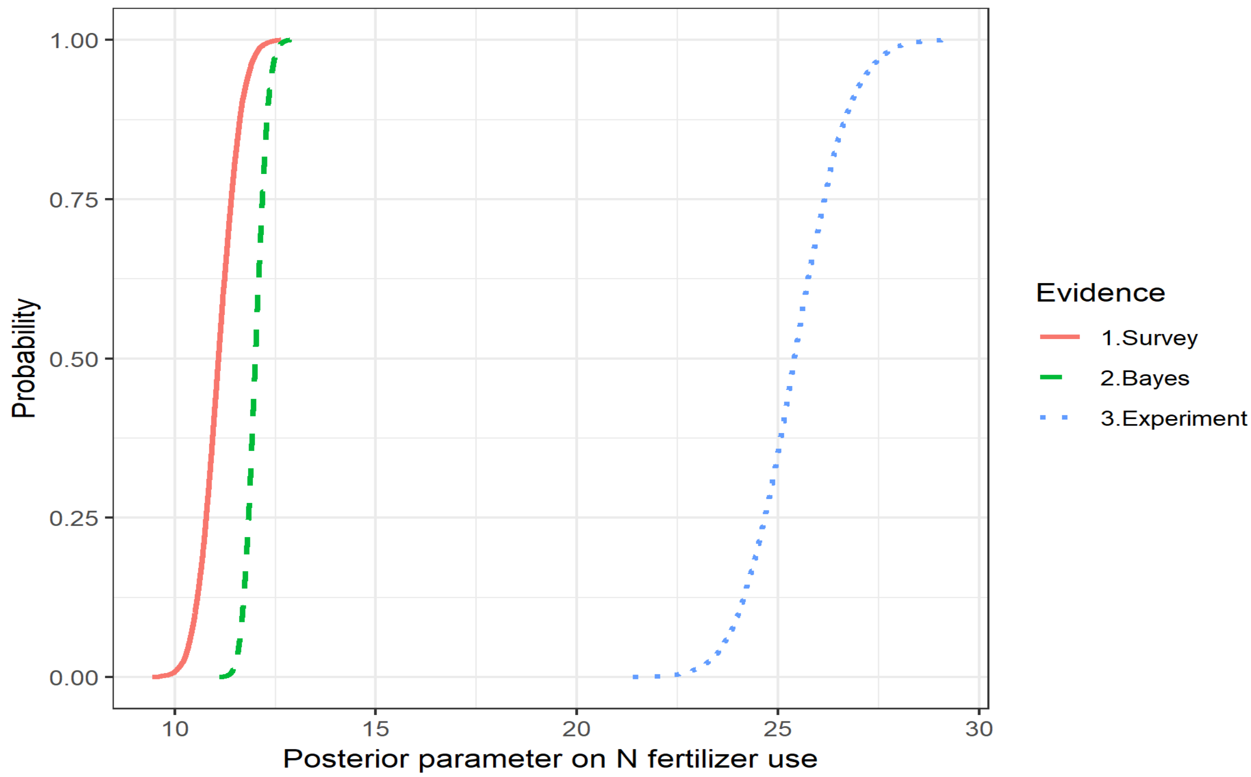


Figure A3: Crop responses to nitrogen application. Note: the Bayes curve uses means and precision parameters from the survey and experimental curves

Appendix B: Profitability analysis and cross-validation

To determine whether Bayesian recommendations are the most profitable to farmers, we follow a three-stage validation procedure.

Stage 1

In the first stage, we use 50% of the survey sample and 100% of the experimental survey data to calculate the optimal rates of nitrogen at a fixed price ratio. We also calculate value-cost ratios (VCR) at the economically optimal recommendations.

$$VCR - EOBR = \frac{p\left(\beta_1^{Bayes} + \beta_2^{Bayes} x^{*Bayes}\right)}{w}$$

$$VCR - EOER = \frac{p\left(\beta_1^{Experiment} + \beta_2^{Experiment} x^{*Experiment}\right)}{w}$$

$$VCR - EOSR = \frac{p\left(\beta_1^{Survey} + \beta_2^{Survey} x^{*Survey}\right)}{w}$$

Stage 2

In the second stage, we re-estimate with 50% of the data using a non-informative Bayesian model on the survey data. We then use the parameters from this equation and optimal rates from the first stage calculate VCR at economically optimal recommendations.

$$VCR - EOBR - Validation = \frac{p\left(\alpha_1^{Validation} + \alpha_2^{Validation} x^{*Bayes}\right)}{w}$$

$$VCR - EOER - Validation = \frac{p\left(\alpha_1^{Validation} + \alpha_2^{Validation} x^{*Experiment}\right)}{w}$$

$$VCR - EOSR - Validation = \frac{p\left(\alpha_1^{Validation} + \alpha_2^{Validation} x^{*Survey}\right)}{w}$$

Stage 3

In the third stage, we calculate the gains and losses in the crop responses from first stage as compared to the second stage. This is the value to using a particular recommendation.

The value-cost ratio (VCR) analysis shows that Bayesian recommendations are the most profitable in the out-of-sample validation tests followed by experimental recommendations.

Chapter 2 Appendices

Appendix A: Additional Tables and Graphs

TableA1: Prior literature using characteristics/trait crop varietal adoption models

Study	Country and Crop	Dependent variables	Independent variables	Research Methods	Econometric Models	Identification strategy
Adesina and Zinnah (1993)	Sierra Leone, Rice	Proportion of the total varietal portfolio (i.e., local and improved varieties) that is constituted by improved varieties	*Variety characteristics*-farmer's valuation of taste, yield, ease of cooking, ease of threshing and tillering capacity Controls: Age, farm size, access to extension services, participation in on-farm trials, rice farming experience	Observational	Tobit	Selection on observables
Lunduka, Fisher, Snapp(2012)	Malawi, Maize	Share plant of local, hybrid, OPV and Recycled hybrid seed varieties	*Variety characteristics*- high yielding, drought tolerance, early maturing, storability, poundability, flour-to- grain ratio, and taste *Controls:* knowledge about varieties, household head, household education, number of adult household members, access to credit (landholding and wealth), distance to agricultural inputs and outputs markets, subsidy seed voucher receipt	Observational (one district) -	Tobit	Selection on observables
Ward et al (2014)	India, Rice	Hypothetical four rice seeds to be chosen by farmers based on characteristics	*Variety characteristics*- Hypothetical researcher introduced characteristics; duration to maturity, yield, grain can be stored and used as seed next season, seed price, seed rate. *Controls:* Age, household size, land area owned, farming experience, access to irrigation water, number of different varieties cultivated,	Discrete choice experiment	Random parameter logit model	Experiment
Zeng (2014)	Maize, Ethiopia	Binary adoption decision	*Variety characteristics*- farmers valuation of grain yield, disease resistance, storability, grain price, taste. *Controls*: prices of seeds, maize output and fertilizer, household size, total land holding, marital status, gender, age, access to credit, distance to the nearest market, seed dealer and fertilizer dealers, regional dummies, adoption decision of current season.	Observational	Mixed logit	Control function approach
Girma et al (2017)	Zimbabwe, Maize	18 hypothetical varieties	*Variety characteristics*: Yield, cob size, grain size, drought tolerance, grain texture, tip (husk) cover and seed price.	Discrete choice experiment	Generalized multinomial logit model	Experiment
Smale et al. (2001)	Mexico, Maize	Portion of household maize area planted to each variety	*Variety characteristics*- Suitability for market sale, consumption of the staple food, preparation of food consumed on special occasions, avoiding disastrous harvests, quality of feed or forage, farmer assessment of seed cost. Controls: regional characteristics (productivity potential and infrastructural development), remittances, percentage of sales, total maize area, percentage of maize area, tractor ownership, access to	informal interviews and field research - Household survey	Tobit	

			irrigation, number of soil types on the farm.			
Edmeades and Smale (2006) and Edmeades et al. (2008)	Uganda, Bananas	Number of banana mats	*Variety characteristics*-cooking quality, yield, brewing quality, yield loss from diseases	Household Survey	Zero Inflated Poisson	Selection on observables
Edmeades (2007)	Uganda, Bananas	Hedonic price	*Variety characteristics:* Quality, bunch size and fruit size	Discrete choice experiment	Two-Stage Least Squares (2-SLS)	Two-Stage Least Squares (2-SLS)
Useche et al. (2009)	USA, GM Corn	Discrete choice of variety	*Variety Characteristics*- seed cost, yield, insecticide use, herbicide use, labor *Other*-total operated acres, labor, education, concern on environmental effects, region dummies	Observational (Administrative Farm survey)	Mixed multinomial logit	Selection on observables

Table A2: Multinomial logit model results without variety characteristics

		2.50%	**50%**	**97.50%**
Intercept	muDK8033Mkangala	-2.60	-1.74	-0.83
Intercept	muDK8053Mapasa	-3.00	-2.12	-1.24
Intercept	muDK9089Fumba	-4.19	-2.70	-1.42
Intercept	muMH18Chokonoka	-3.36	-2.44	-1.55
Intercept	muPAN53	-6.17	-4.66	-3.20
Intercept	muSC403Kanyani	-0.72	-0.14	0.49
Intercept	muSC627Mkango	-1.47	-0.79	-0.14
Intercept	muSC719Njovu	-4.05	-2.86	-1.68
Age	muDK8033Mkangala	-0.03	-0.02	-0.01
Age	muDK8053Mapasa	-0.03	-0.02	-0.01
Age	muDK9089Fumba	-0.05	-0.03	-0.01
Age	muMH18Chokonoka	-0.02	-0.01	0.01
Age	muPAN53	-0.01	0.01	0.02
Age	muSC403Kanyani	-0.02	-0.01	0.00
Age	muSC627Mkango	-0.02	-0.01	0.00
Age	muSC719Njovu	-0.03	-0.02	0.00
Household size	muDK8033Mkangala	-0.04	0.03	0.10
Household size	muDK8053Mapasa	0.00	0.06	0.11
Household size	muDK9089Fumba	-0.05	0.05	0.14
Household size	muMH18Chokonoka	0.04	0.11	0.17
Household size	muPAN53	-0.05	0.03	0.11
Household size	muSC403Kanyani	0.01	0.06	0.11
Household size	muSC627Mkango	-0.02	0.03	0.09
Household size	muSC719Njovu	-0.02	0.05	0.12
Education	muDK8033Mkangala	0.01	0.06	0.10
Education	muDK8053Mapasa	0.03	0.08	0.12
Education	muDK9089Fumba	-0.01	0.06	0.12
Education	muMH18Chokonoka	-0.01	0.04	0.09
Education	muPAN53	0.01	0.08	0.15

Education	muSC403Kanyani	-0.02	0.02	0.05
Education	muSC627Mkango	0.00	0.04	0.07
Education	muSC719Njovu	-0.03	0.03	0.09
Sex	muDK8033Mkangala	-0.60	0.18	0.93
Sex	muDK8053Mapasa	-0.83	-0.05	0.64
Sex	muDK9089Fumba	-1.10	-0.13	0.98
Sex	muMH18Chokonoka	-0.62	0.19	1.10
Sex	muPAN53	-0.28	0.97	2.52
Sex	muSC403Kanyani	-0.48	0.02	0.48
Sex	muSC627Mkango	-0.77	-0.19	0.33
Sex	muSC719Njovu	-1.24	-0.40	0.50
Plot area	muDK8033Mkangala	-1.20	-0.62	-0.17
Plot area	muDK8053Mapasa	-0.24	-0.04	0.07
Plot area	muDK9089Fumba	-0.94	-0.28	0.04
Plot area	muMH18Chokonoka	-1.07	-0.51	-0.07
Plot area	muPAN53	-0.61	-0.15	0.06
Plot area	muSC403Kanyani	-1.52	-1.10	-0.71
Plot area	muSC627Mkango	-1.25	-0.78	-0.39
Plot area	muSC719Njovu	-0.75	-0.25	0.02
Fertilizer subsidy	muDK8033Mkangala	-1.12	-0.32	0.44
Fertilizer subsidy	muDK8053Mapasa	-1.41	-0.57	0.16
Fertilizer subsidy	muDK9089Fumba	-1.95	-0.64	0.48
Fertilizer subsidy	muMH18Chokonoka	-1.34	-0.53	0.21
Fertilizer subsidy	muPAN53	0.23	0.97	1.68
Fertilizer subsidy	muSC403Kanyani	-0.94	-0.38	0.13
Fertilizer subsidy	muSC627Mkango	-0.34	0.14	0.63
Fertilizer subsidy	muSC719Njovu	-1.65	-0.59	0.29
Seed subsidy	muDK8033Mkangala	0.23	0.96	1.71
Seed subsidy	muDK8053Mapasa	0.34	1.05	1.83
Seed subsidy	muDK9089Fumba	0.37	1.46	2.82
Seed subsidy	muMH18Chokonoka	-0.40	0.32	1.13
Seed subsidy	muPAN53	-1.10	-0.44	0.30
Seed subsidy	muSC403Kanyani	0.46	0.97	1.51
Seed subsidy	muSC627Mkango	-0.28	0.19	0.71
Seed subsidy	muSC719Njovu	0.12	0.96	1.91
Marital status	muDK8033Mkangala	-0.38	0.43	1.26
Marital status	muDK8053Mapasa	0.05	0.79	1.63
Marital status	muDK9089Fumba	-0.20	0.85	2.04
Marital status	muMH18Chokonoka	-0.35	0.44	1.29
Marital status	muPAN53	-0.75	0.29	1.57
Marital status	muSC403Kanyani	-0.58	-0.12	0.35
Marital status	muSC627Mkango	-0.25	0.24	0.84
Marital status	muSC719Njovu	0.61	1.64	2.74

Appendix Table A3: Bayesian hierarchical multinomial logit with district heterogeneity

	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
Intercept	1.7700	0.1900	1.4100	2.1500
Price	-0.0100	0.0000	-0.0100	-0.0100
Variety age (years)	0.1500	0.0100	0.1400	0.1600
Days to maturity	-0.0600	0.0000	-0.0600	-0.0600
Flint	1.2400	0.0400	1.1600	1.3300
Yield potential	0.0000	0.0000	0.0000	0.0000
Drought tolerance	1.1500	0.0500	1.0600	1.2400
District(16) level Effects				
sd(Intercept)	0.4400	0.1700	0.1400	0.8300
sd(Price)	0.0000	0.0000	0.0000	0.0000
sd(Variety age)	0.0200	0.0000	0.0100	0.0300
sd(MaturityMidPoint)	0.0000	0.0000	0.0000	0.0000
sd(FlintDentFlint)	0.0300	0.0300	0.0000	0.0900
sd(PotYieldMidPoint)	0.0000	0.0000	0.0000	0.0000
sd(DroughtToleranceYes)	0.1400	0.0300	0.0900	0.2100
cor(Intercept,Price)	0.1000	0.4500	-0.6400	0.8000
cor(Intercept,Variety age)	-0.6100	0.3000	-0.9500	0.0400
cor(Price,Variety age)	-0.7200	0.2000	-0.9500	-0.2400
cor(Intercept,MaturityMidPoint)	-0.1600	0.3600	-0.7800	0.5600
cor(Price,MaturityMidPoint)	-0.3100	0.4000	-0.9000	0.5600
cor(Variety age,MaturityMidPoint)	0.2700	0.3800	-0.5400	0.8900
cor(Intercept,FlintDentFlint)	-0.0700	0.3200	-0.6600	0.5200
cor(Price,FlintDentFlint)	0.0300	0.3800	-0.6200	0.7000
cor(Variety age,FlintDentFlint)	-0.0400	0.3800	-0.8500	0.6300
cor(MaturityMidPoint,FlintDentFlint)	-0.0800	0.3300	-0.7200	0.5300
cor(Intercept,PotYieldMidPoint)	-0.6600	0.2100	-0.9000	-0.1100
cor(Price,PotYieldMidPoint)	-0.5100	0.3300	-0.9400	0.1800
cor(Variety age,PotYieldMidPoint)	0.7000	0.2100	0.1000	0.9300
cor(MaturityMidPoint,PotYieldMidPoint)	0.0300	0.3400	-0.6200	0.7200
cor(FlintDentFlint,Yield potential)	-0.0600	0.3900	-0.7200	0.7000
cor(Intercept,DroughtToleranceYes)	-0.6200	0.2700	-0.9500	0.0400
cor(Price,DroughtToleranceYes)	-0.6300	0.2400	-0.9500	0.0300
cor(Variety age,DroughtToleranceYes)	0.8200	0.1200	0.5100	0.9500
cor(MaturityMidPoint,DroughtToleranceYes)	0.2600	0.4100	-0.6000	0.8700
cor(FlintDentFlint,DroughtToleranceYes)	0.1200	0.3500	-0.6800	0.7400
cor(Yield potential,DroughtToleranceYes)	0.7000	0.1800	0.2700	0.9600

Appendix Table A4: Bayesian hierarchical multinomial logit with individual household heterogeneity

	Estimate	**Est.Error**	**l-95% CI**	**u-95% CI**
Intercept	1.4400	0.0700	1.2700	1.5500
Price	-0.0100	0.0000	-0.0100	-0.0100
Variety age (years)	0.1600	0.0000	0.1600	0.1700
Days to maturity	-0.0700	0.0000	-0.0800	-0.0600
Flint	1.4200	0.0400	1.3000	1.4800
Yield potential	0.0000	0.0000	0.0000	0.0000
Drought tolerance	1.2100	0.0200	1.1900	1.2600
Individual (1651) level effects				
sd(Intercept)	0.0200	0.0100	0.0000	0.0500
sd(Price)	0.0000	0.0000	0.0000	0.0000
sd(Variety age)	0.0100	0.0100	0.0000	0.0200
sd(Days to maturity)	0.0100	0.0000	0.0000	0.0100
sd(FlintDentFlint)	0.1800	0.0600	0.0200	0.2500
sd(Yield potential)	0.0000	0.0000	0.0000	0.0000
sd(DroughtToleranceYes)	0.1000	0.0800	0.0100	0.1900
cor(Intercept,Price)	-0.1900	0.6400	-0.8900	0.5900
cor(Intercept,Variety age)	0.1500	0.5600	-0.5400	0.7200
cor(Price,Variety age)	-0.8300	0.0500	-0.9200	-0.7500
cor(Intercept,MaturityMidPoint)	0.5100	0.3000	-0.0600	0.8900
cor(Price,MaturityMidPoint)	-0.1000	0.8900	-0.9900	0.8700
cor(Variety age,MaturityMidPoint)	-0.0500	0.7700	-0.9000	0.7400
cor(Intercept,FlintDentFlint)	-0.4500	0.2200	-0.7100	-0.1100
cor(Price,FlintDentFlint)	-0.0300	0.7800	-0.8700	0.7800
cor(Variety age,FlintDentFlint)	-0.0700	0.9000	-0.9800	0.8900
cor(MaturityMidPoint,FlintDentFlint)	-0.7900	0.2300	-1.0000	-0.1800
cor(Intercept,PotYieldMidPoint)	-0.1900	0.1500	-0.4400	0.0000
cor(Price,PotYieldMidPoint)	-0.3700	0.4600	-0.8600	0.1600
cor(Variety age,PotYieldMidPoint)	0.4700	0.3900	-0.1100	0.8900
cor(MaturityMidPoint,PotYieldMidPoint)	-0.5800	0.4100	-1.0000	0.0900
cor(FlintDentFlint,PotYieldMidPoint)	0.4900	0.4900	-0.1000	1.0000
cor(Intercept,DroughtToleranceYes)	0.2700	0.5800	-0.5500	0.8500
cor(Price,DroughtToleranceYes)	-0.5700	0.4600	-0.9900	0.3800
cor(Variety age,DroughtToleranceYes)	0.5300	0.4500	-0.4100	0.9200
cor(MaturityMidPoint,DroughtToleranceYes)	0.4400	0.5300	-0.3800	0.9400
cor(FlintDentFlint,DroughtToleranceYes)	-0.3800	0.4800	-0.8700	0.3700
cor(PotYieldMidPoint,DroughtToleranceYes)	-0.0100	0.2100	-0.4800	0.3400

Appendix B: Multidimensional representations of the characteristic's technology frontier

(a) Biplot analysis

We can use a biplot to show a two-dimensional approximation of the multi-dimensional representation of the characteristics data.

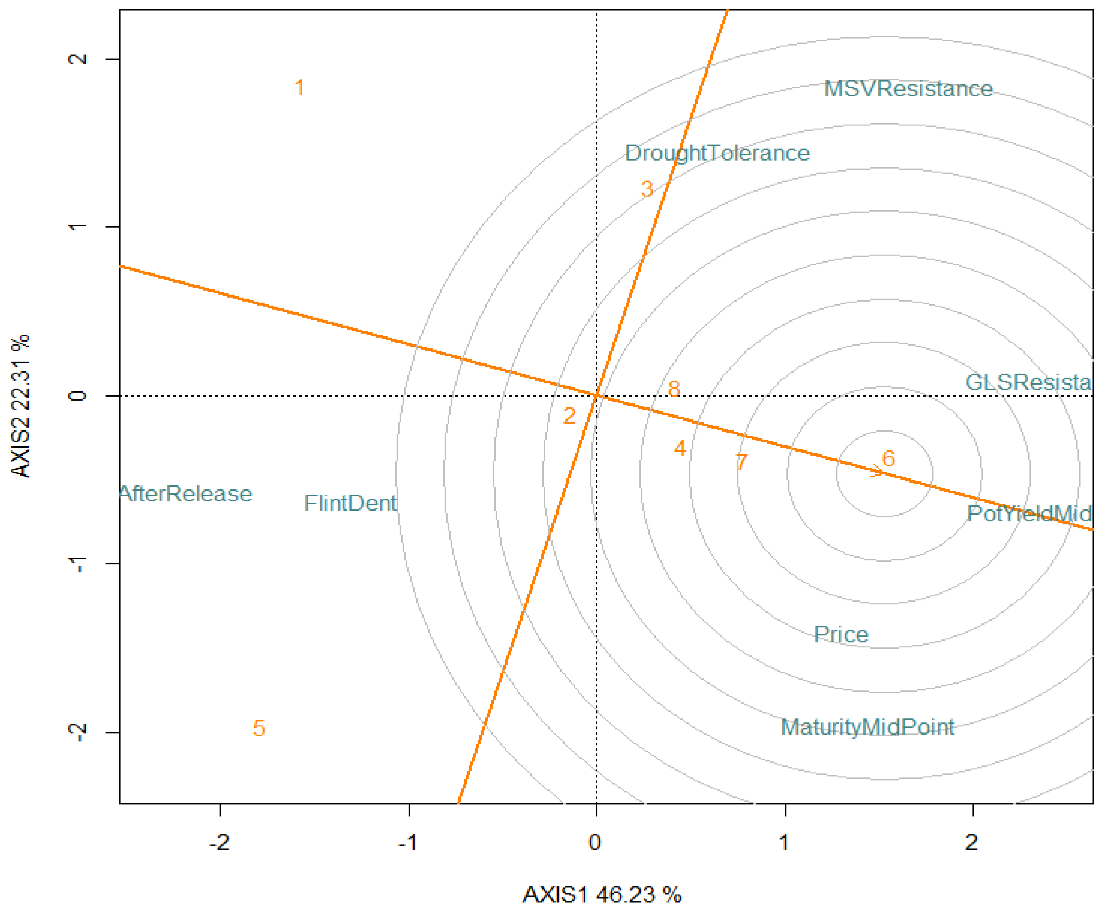


Figure A1: Biplot for ranking varieties

Notes: The numbers represent the varieties where 1=SC403, 2=SC 627, 3=DK8033, 4=DK8053, 5=MH 18, 6=SC 719, 7=PAN53, 8=DK9089.

The varieties closer to the origin have most of the characteristics while those far from the origin like 1 (SC 403) and 5 (MH 18) have specialized traits. Most importantly, these varieties are the oldest. The closeness of the variety to a trait reveals the trait profile of that variety. For instance, variety 6 (SC 4719) has the highest yield while varieties 1 (SC 403) and 5 (MH 18) have the highest yield potential.

(b) Data envelopment and revealed preference analysis

As we increase the number of characteristics, it is difficult to visualize the ranking of the varieties. One can therefore consider a mathematical representation of the characteristics technology that allows a multi-dimensional ranking of the varieties. Data Envelopment Analysis (DEA) as used in production efficiency analysis can be used for such purpose. Choi and Oh (2010) suggest that DEA can be used for a Lancaster characteristics space analysis by treating price as an input and the rest of the product characteristics as outputs. We leave this for future research.

Chapter 3 Appendices

Appendix A: Partial tableau

Table A1: A partial tableau (one-district) illustrating the structure of the Malawi Food Sector Programming Model (MAFOSP)

	MAIZE	RICE	CASSAVA	POTATOES	BEANS	G/NUTS	2000	2001	2002	2003	---	2009	INPUT SUPPLY	PRODUCT DEMAND	RHS	EQUATION NUMBER
MAX: NET BENEFIT													–	+		(1)
X-MAIZE	1						-a	-a	-a	-a	-a	-a			≤ 0	(2)
X-RICE		1					-a	-a	-a	-a	-a	-a			≤ 0	(2)
X-CASSAVA			1				-a	-a	-a	-a	-a	-a			≤ 0	(2)
X-POTATOES				1			-a	-a	-a	-a	-a	-a			≤ 0	(2)
X-BEANS					1		-a	-a	-a	-a	-a	-a			≤ 0	(2)
X-GROUNDNUTS						1	-a	-a	-a	-a	-a	-a			≤ 0	(2)
TOTAL LAND							-a	-a	-a	-a	-a	-a			≤ 0	(2)
CROP MIX CONVEXITY							1	1	1	1	1	1			= 1	(2)
PRODUCT CONSTRAINTS	–	–	–	–	–	–							+		≤ 0	(2)
INPUT CONSTRAINTS	+	+	+	+	+	+								–	≤ 0	(3)

Appendix B: Food crop flows in Malawi without Cropping Restrictions

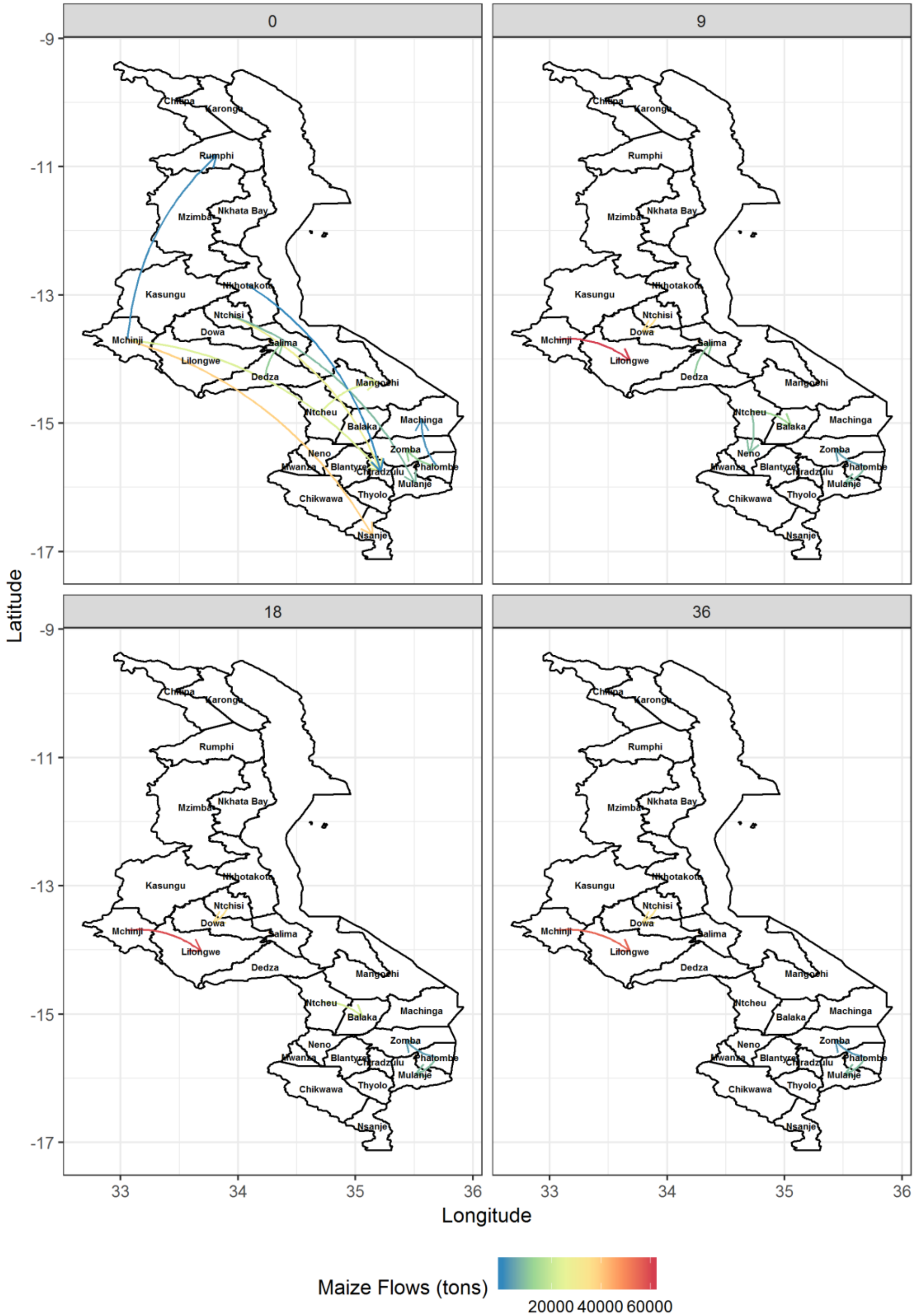


Figure B1: Maize flows

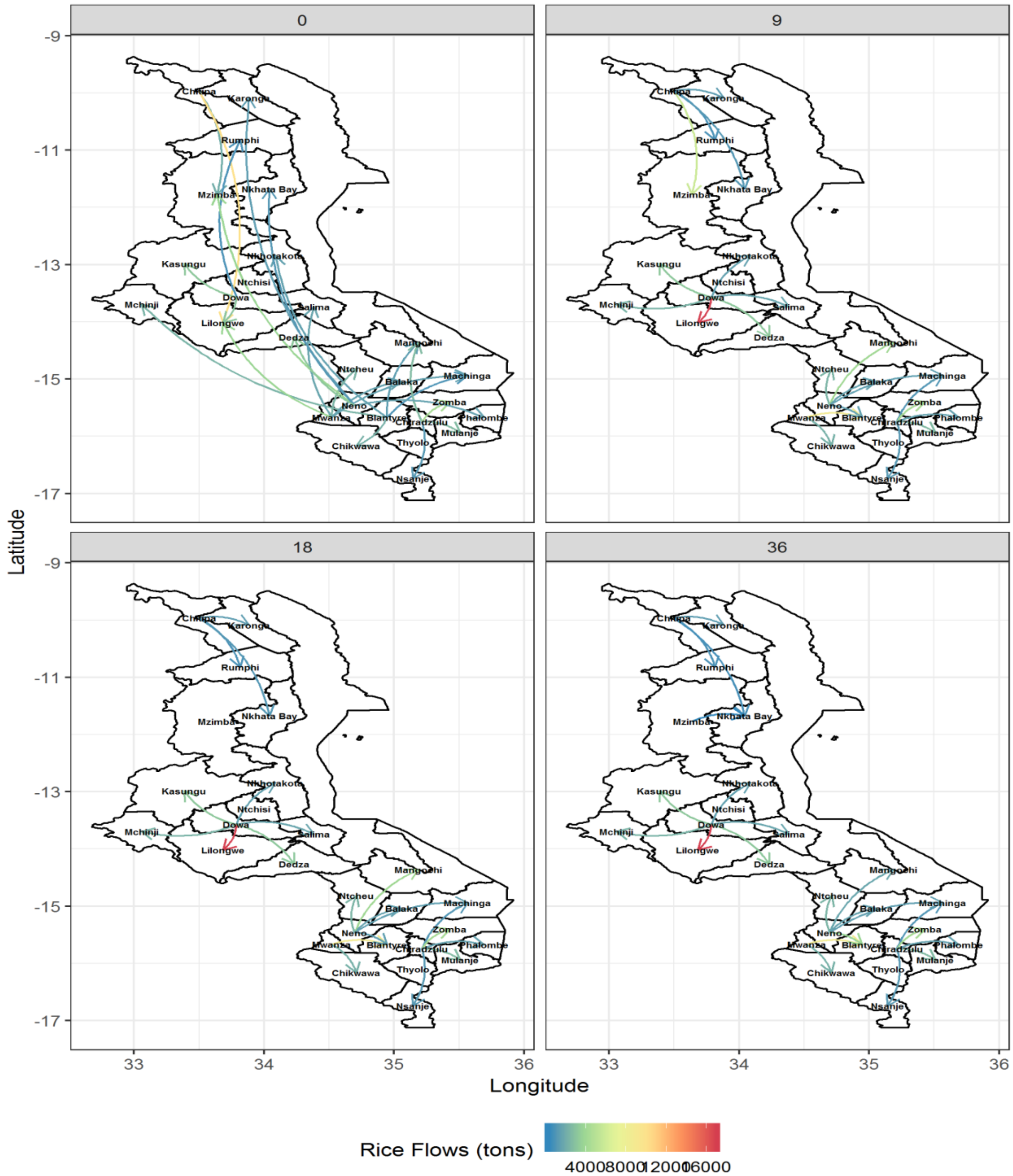


Figure B2: Rice flows

Notes: The rice flows are not consistent with the production patterns prevailing in the country. The major rice producing districts are Karonga, Nkhotakota, Zomba, Chikwawa and Nsanje. In the figure, each of these is buying rice. A water balance module will help correct this pattern.

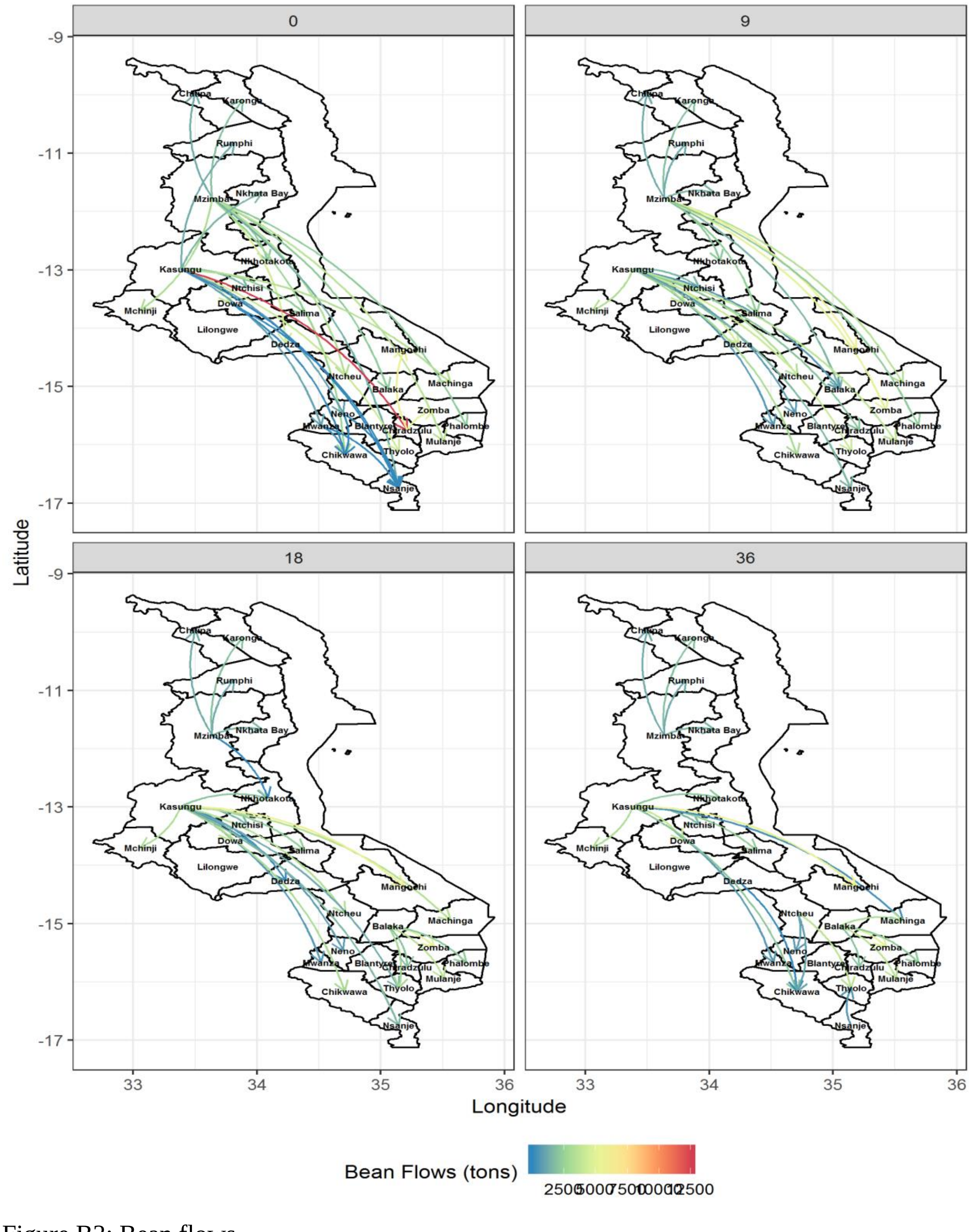


Figure B3: Bean flows

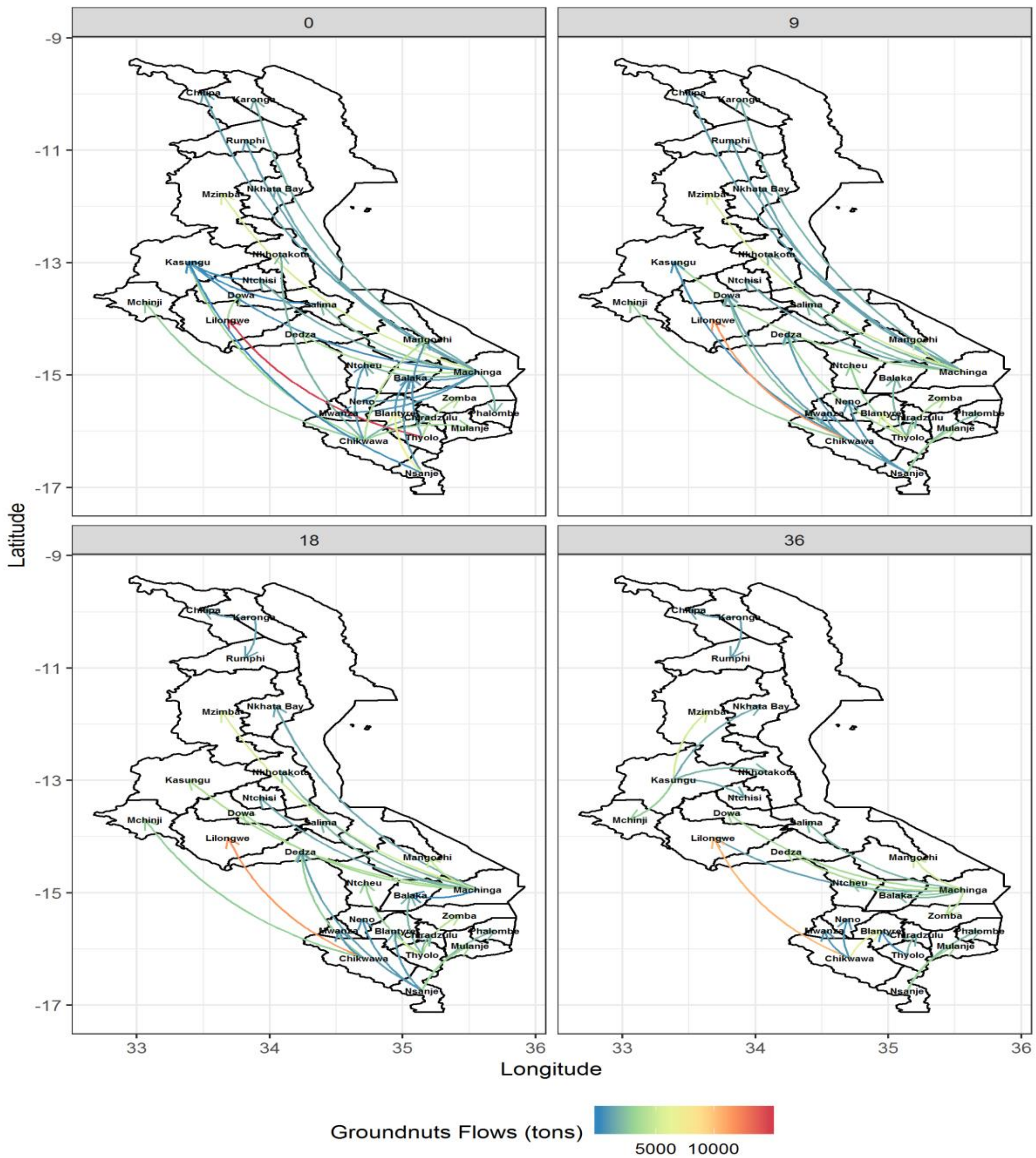


Figure B4: Groundnut flows

Appendic C: Food crop flows in Malawi with crop mix restrictions

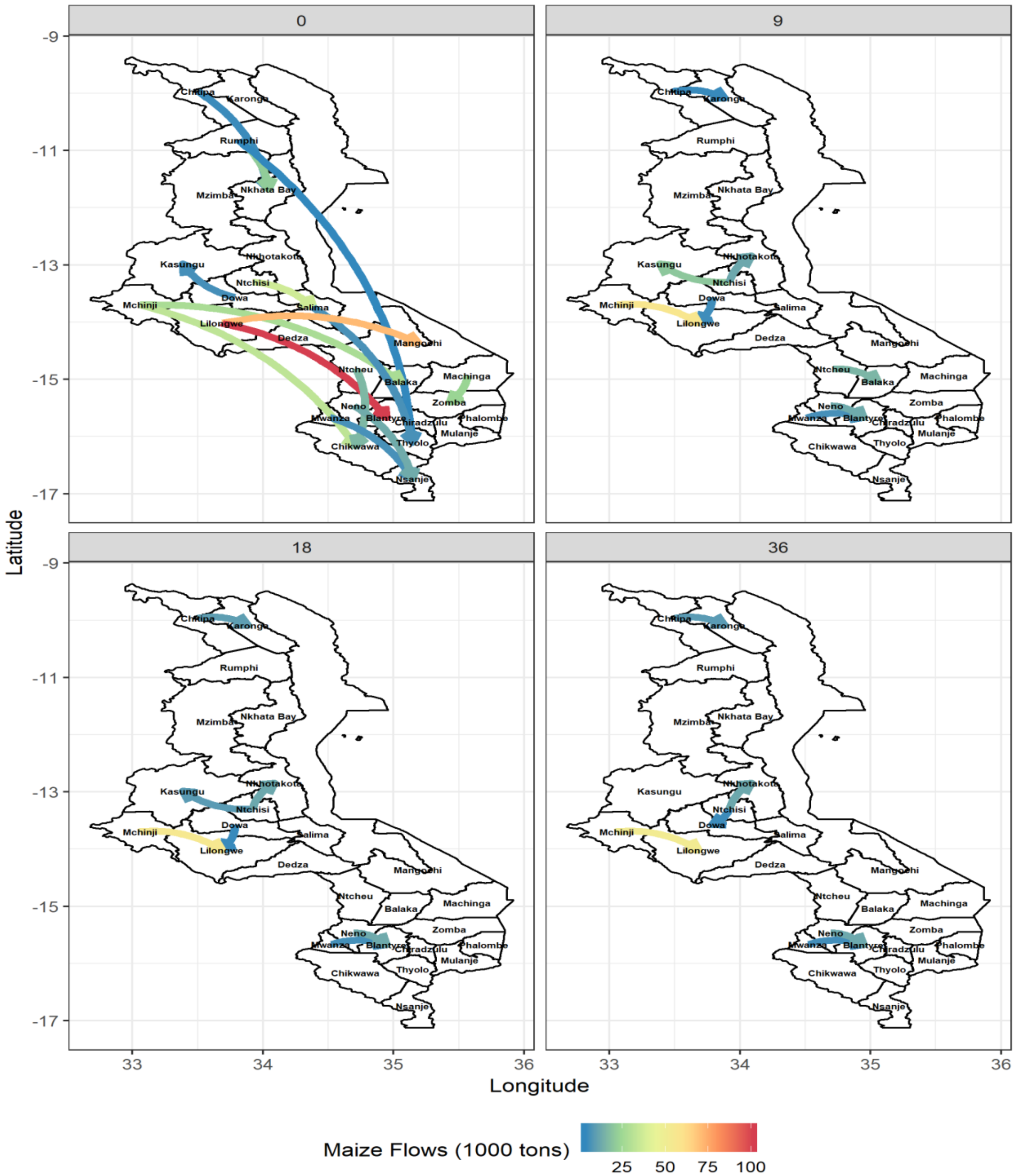


Figure C1: Maize flows

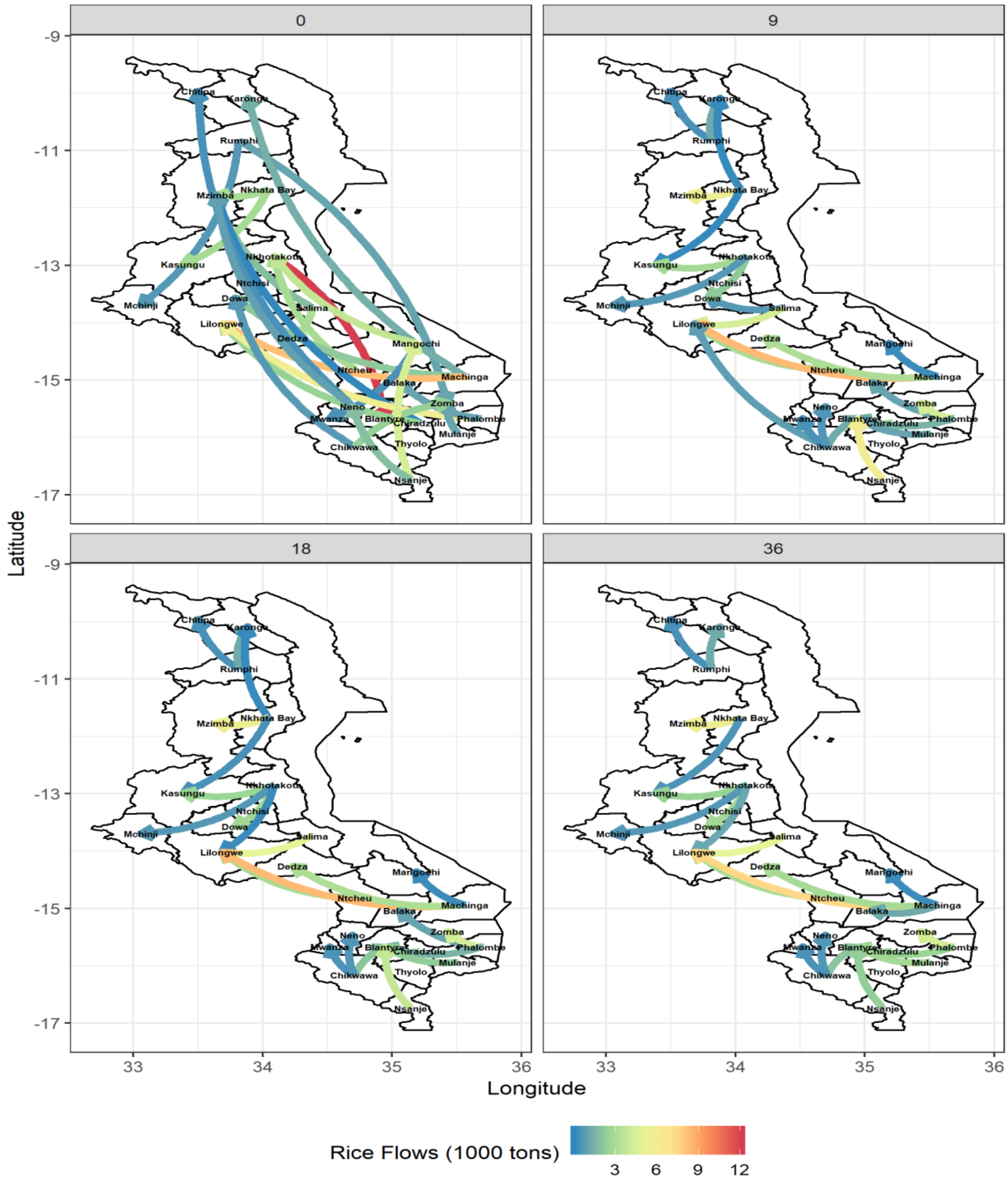


Figure C2: Rice flows

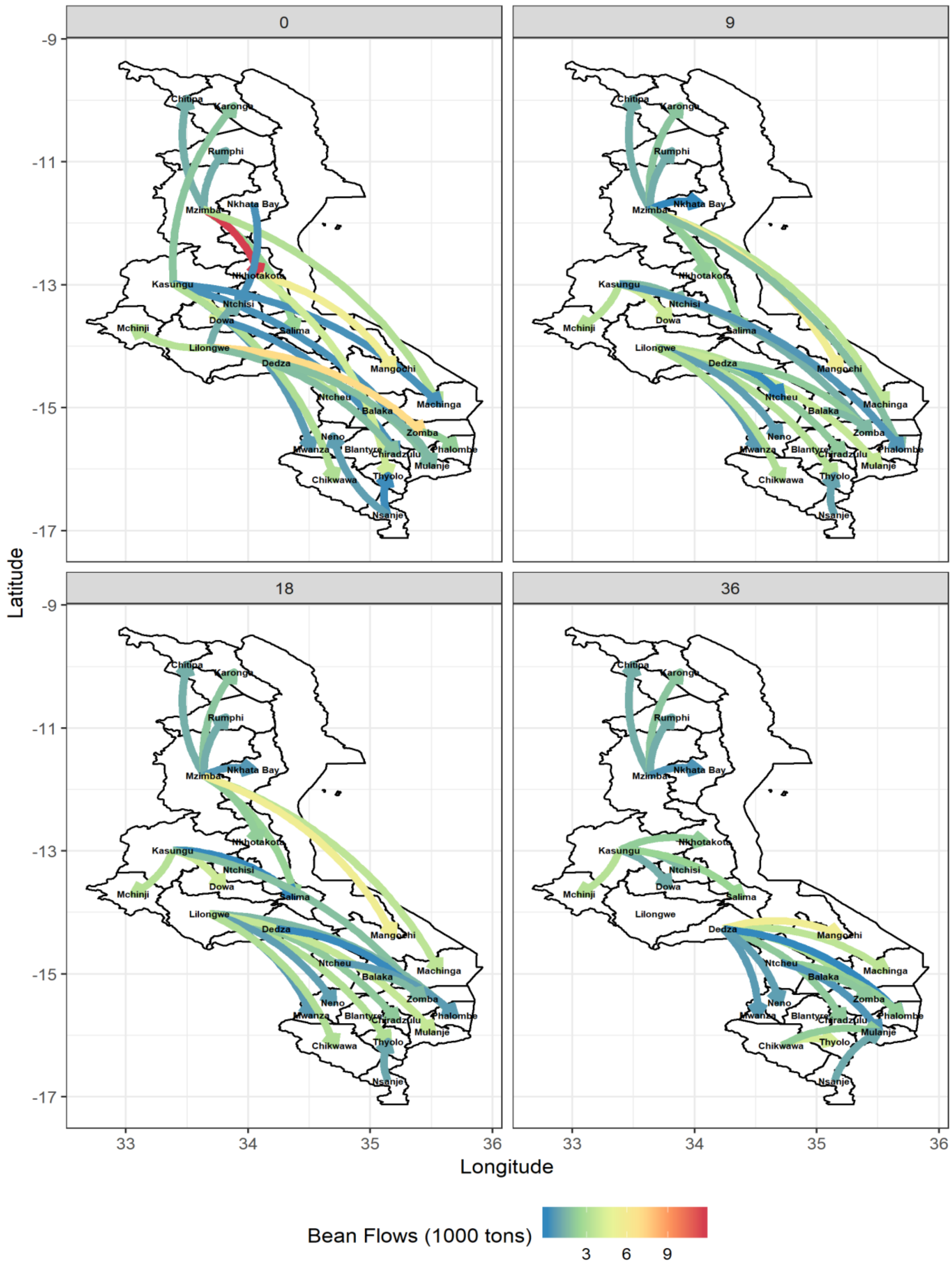


Figure C3: Bean flows

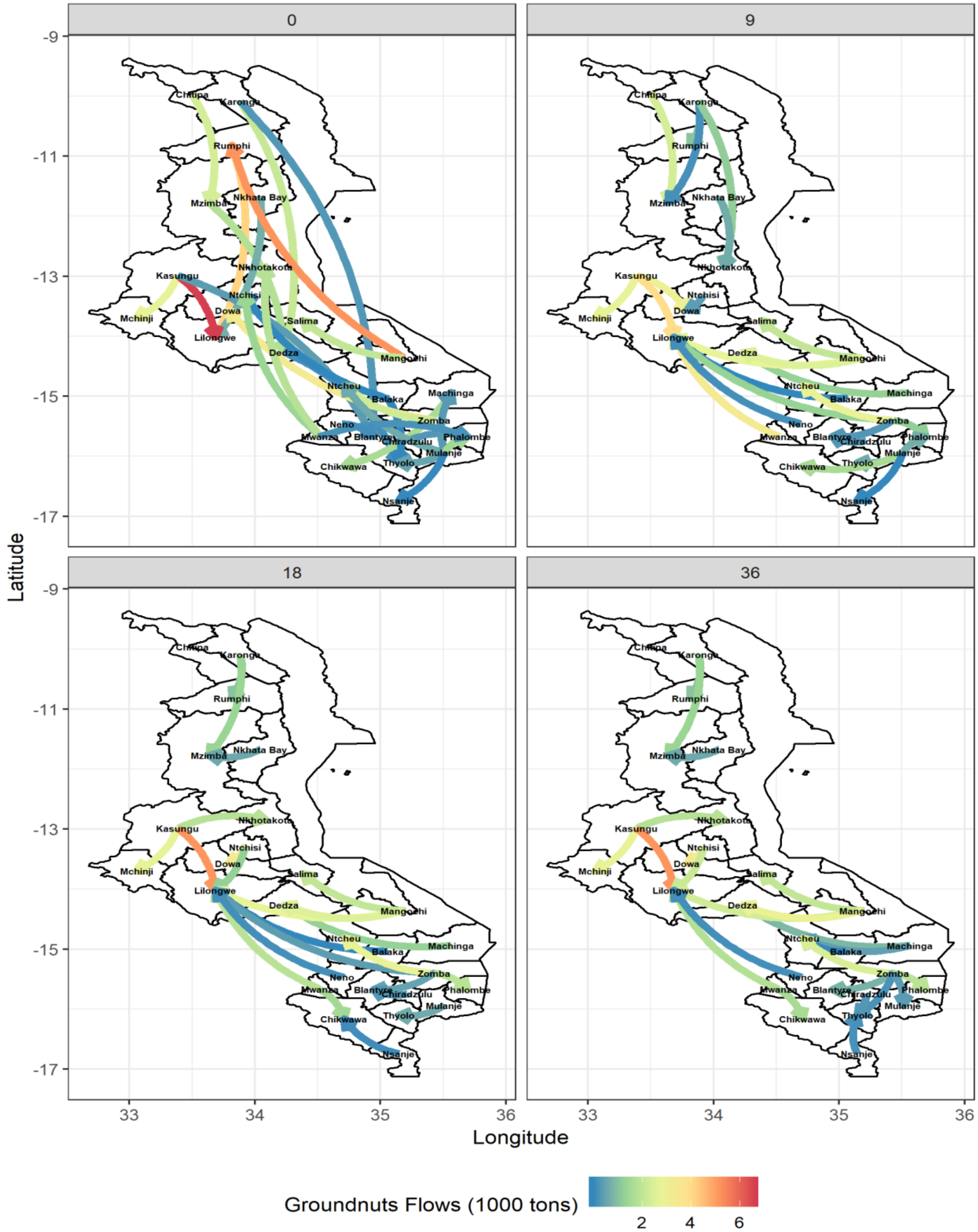


Figure C4: Groundnut flows

Appendix D: GAMS code for the agricultural sector model

```
$TITLE: A MALAWI AGR SECTOR PROTOTYPE MODEL FOR POLICY ANALYSIS; 2019-REF
$OFFSYMXREF
$OFFSYMLIST
*
* JEFFREY APLAND
* MAXWELL MKONDIWA
* DEPT OF APPLIED ECONOMICS
* UNIVERSITY OF MINNESOTA
*
OPTION LIMROW=100, LIMCOL=100, SOLPRINT=ON, ITERLIM=2000;

* SETS CONTROLLING SOLUTION REPORTS

* MAXWELL: 27TH FEB 2019
* Add Policy Scenarios.  Active only the line
*associated to the scenario of interest and run de model
SET SWITCH /BASELINE, ZEROTRANSPORTCOST,HALFTRANSPORTCOST,DOUBLETRANSPORTCOST/;

PARAMETERS SCENARIOS(SWITCH) BASELINESC /BASELINE 1,ZEROTRANSPORTCOST
0,HALFTRANSPORTCOST 0,DOUBLETRANSPORTCOST 0/;
*PARAMETERS SCENARIOS(SWITCH) ZEROTRANSPORTCOSTSC /BASELINE 0,ZEROTRANSPORTCOST
1,HALFTRANSPORTCOST 0,DOUBLETRANSPORTCOST 0/;
*PARAMETERS SCENARIOS(SWITCH) HALFTRANSPORTCOSTSC /BASELINE 0,ZEROTRANSPORTCOST
0,HALFTRANSPORTCOST 1,DOUBLETRANSPORTCOST 0/;
*PARAMETERS SCENARIOS(SWITCH) DOUBLETRANSPORTCOSTSC /BASELINE 0,ZEROTRANSPORTCOST
0,HALFTRANSPORTCOST 0,DOUBLETRANSPORTCOST 1/;


SET  KRT  RESULTS TABLES
   /01  OBJECTIVE FUNCTION VALUE,                02  REGIONAL PRODUCTION ACTIVITIES,
    03  CROP PRODUCTION ACTIVITIES,              04  REGIONAL INPUT SUPPLY ACTIVITIES,
    05  CROP LAND INPUT SUPPLY ACTIVITIES,       06  REGIONAL PRODUCT DEMAND
ACTIVITIES,
    07  INTER-REG TRANSPORTATION ACTIVITIES,     08  REGIONAL IMPORT ACTIVITIES,
    09  REGIONAL EXPORT ACTIVITIES,
    31  REGIONAL INPUT CONSTRAINTS,              32  CROP LAND INPUT CONSTRAINTS,
    33  REGIONAL PRODUCT CONSTRAINTS,
    41  CROP ACTIVITIES BY REG & LAND TYPE,      42  CROP ACTIVITIES SUMMED BY REGION,
    43  TOTAL CROP PROD ACTIVITY LEVELS,
    71  REG INPUT & PROD RESULTS BY REGION,      72  REG INPUT & PROD RESULTS BY ITEM,
    73  REG INPUT & PROD RESULTS SELECT ITEMS,
    81  INTER-REG INPUT & PRODUCT SHIPMENTS,
    91  CROP RESULTS BY REG TABLE 91,            92  CROP RESULTS TOTAL TABLE 92 /;

* PRINT TABLE IF PRNT = 1
PARAMETER  PRNT(KRT)
   /01*03 1, 04*09 1, 31*33 1, 41 1, 42 1, 43 1, 71 1, 72 1, 73 1, 81 1, 91 1, 92 1/;

$STITLE SET DECLARATIONS AND ASSIGNMENTS
* OVERVIEW OF SETS:
*-------------------------------------------------------------------------------------
------------------
* SET................
DESCRIPTION...............................................................
*-------------------------------------------------------------------------------------
------------------
* JR                    REGIONS
* JRS                   INTER-REGIONAL TRANSPORT SOURCE REGIONS, ALIAS JR
* JRD                   INTER-REGIONAL TRANSPORT DESTINATION REGIONS, ALIAS JR
*
* JIP                   INPUTS AND PRODUCTS
```

```
*
* JI(JIP)               INPUTS
* JIR(JI)               REGIONAL INPUTS
* JIT                   CROP LAND TYPES
* JIC(JI)               CROP LAND INPUTS
* JIN(JI)               NONREGIONAL INPUTS
* JRIR(JR,JIR)          REGIONAL INPUTS MAPPED TO REGIONS
* JRT(JR,JIT)           CROP LAND TYPES MAPPED TO REGIONS
* JRIC(JR,JIC)          CROP LAND INPUTS MAPPED TO REGIONS

* JP(JIP)               PRODUCTS
* JPR(JP)               REGIONAL PRODUCTS
* JPN(JP)               NONREGIONAL PRODUCTS
* JRPR(JR,JP)           REGIONAL PRODUCTS MAPPED TO REGIONS
*
* JRSD(JRS,JRD)         INTER-REGIONAL TRANSPORT DESTINATIONS MAPPED TO SOURCES
* JTIP(JIP)             TRANSPORTED REGIONAL INPUTS AND PRODUCTS
* JRSDIP(JRS,JRD,JIP)   INPUTS AND PRODUCTS MAPPED TO SOURCE AND DESTINATION REGIONS
*
* JIPX(JIP)             EXPORTED INPUTS AND PRODUCTS
* JIPM(JIP)             IMPORTED INPUTS AND PRODUCTS
* JIPRX(JR,JIP)         EXPORTED REGIONAL INPUTS AND PRODUCTS MAPPED TO REGIONS
* JIPRM(JR,JIP)         IMPORTED REGIONAL INPUTS AND PRODUCTS MAPPED TO REGIONS
* JIPNX(JIP)            EXPORTED NONREGIONAL INPUTS AND PRODUCTS
* JIPNM(JIP)            IMPORTED NONREGIONAL INPUTS AND PRODUCTS
*
* JX                    PRODUCTION ACTIVITIES
* JXR(JX)               REGIONAL ACTIVITIES
* JXC(JXR)              CROP PRODUCTION ACTIVITIES
* JXL(JXR)              LIVESTOCK PRODUCTION ACTIVITIES
* JXP(JXR)              PROCESSING ACTIVITIES
* JXO(JX)               OTHER ACTIVITIES
* JRX(JR,JX)            PRODUCTION ACTIVITIES MAPPED TO REGIONS
* JXCT(JXC,JIT)         CROP LAND TYPES MAPPED TO CROP PROD ACTIVITIES
* JRXT(JR,JXC,JIT)      CROP LAND TYPES AND CROP PROD ACTIVITIES MAPPED TO REGIONS
* JXN(JX)               NONREGIONAL PRODUCTION ACTIVITIES
*
* JB                    BUDGETS USED TO CONSTRUCT CROP AND LIVESTOCK ACTIVITIES
* JBR(JR,JB)            BUDGETS MAPPED TO REGIONS
* JBC(JB)               CROP BUDGETS USED TO CONSTRUCT CROP PRODUCTION ACTIVITIES
* JBCT(JBC,JIT)         CROP BUDGETS MAPPED TO CROP LAND TYPES
* JBCR(JR,JXC,JBC)      CROP BUDGETS MAPPED TO CROP PROD ACTIVITIES AND REGIONS
* JBL(JB)               LIVESTOCK BUDGETS USED TO CONSTRUCT LIVESTOCK PROD ACTIVITIES
* JBLR(JR,JXL,JBL)      LIVESTOCK BUDGETS MAPPED TO LIVESTOCK PROD ACTIVITIES AND
REGIONS
*
* JOC                   OTHER PRODUCTION ACTIVITY CONSTRAINTS
* JOCR(JOC)             OTHER REGIONAL PRODUCTION ACTIVITY CONSTRAINTS
* JROCR(JR,JOCR)        OTHER REGIONAL CONSTRAINTS MAPPED TO REGIONS
* JOCN(JOC)             OTHER NONREGIONAL PRODUCTION ACTIVITY CONSTRAINTS
*
* JU                    UNITS OF MEASURE
* JU2                   UNITS OF MEASURE, ALIAS JU
* JUX(JX,JU)            UNITS MAPPED TO PRODUCTION ACTIVITIES
* JUIPX(JIP,JU)         UNITS OF MEASURE MAPPED TO INPUTS AND PRODUCTS FOR PROD
ACTIVITIES
* JUB(JB,JU)            UNITS OF MEASURE MAPPED TO BUDGETS
* JUBL(JBC,JU)          UNITS OF MEASURE MAPPED TO LIVESTOCK BUDGETS
* JUIPB(JIP,JU)         UNITS OF MEASURE MAPPED TO INPUTS AND PRODUCTS FOR BUDGETS
*
* JSDP                  SUPPLY AND DEMAND PARAMETERS
* JXMP                  EXPORT AND IMPORT PARAMETERS
```

```
*--------------------------------------------------------------------------------------------------
*
* MOST SETS ARE DECLARED AND ELEMENTS ASSIGNED IN THE ORDER LISTED IN THE COMMENTS
ABOVE.  HOWEVER,
* SOME SETS ARE ASSIGNED AFTER NECESSARY DATA HAVE BEEN ENTERED.  A GAMS SET IS CALLED
DYNAMIC IF
* ITS ELEMENTS ARE ASSIGNED BASED ON THE ELEMENTS OF OTHER SETS OR THE VALUES OF
CERTAIN PARAMETERS.
* DYNAMIC SETS ARE MARKED ABOVE WITH **.  ONCE MEMBERSHIP HAS BEEN ASSIGNED, IT IS
USEFUL TO EXCLUDE
* ELEMENTS FROM CERTAIN SETS IN ORDER TO MANAGE MODEL SIZE.  FOR EXAMPLE, ALL REGIONAL
INPUTS ARE
* ASSIGNED TO ALL REGIONS.  HOWEVER, IF SOME INPUTS ARE NOT USED IN SOME REGIONS, THE
NUMBER OF
* CONSTRAINTS MAY BE REDUCED BY EXCLUDING THE INPUTS FROM THOSE REGIONS.


SET  JR  REGIONS
*--------------------------------------------------------------------------------------------------
* REGION......    DESCRIPTION.................      REGION......
DESCRIPTION.................
*--------------------------------------------------------------------------------------------------
 / CHI          CHITIPA,                           KAR           KARONGA,
   RUM          RUMPHI,                            MZI           MZIMBA,
   KHA          NKHATABAY,                         KAS           KASUNGU,
   DOW          DOWA,                              NTCH          NTCHISI,
   MCH          MCHINJI,                           SA            SALIMA,
   KK           NKHOTAKOTA,                        LL            LILONGWE,
   DED          DEDZA,                             NU            NTCHEU,
   MAC          MACHINGA,                          MAN           MANGOCHI,
   BLK          BALAKA,                            ZA            ZOMBA,
   BT           BLANTYRE,                          TO            THYOLO,
   CZ           CHIRADZULU,                        PHA           PHALOMBE,
   MU           MULANJE,                           MWA           MWANZA,
   NEN          NENO,                              CHK           CHIKWAWA,
   NSA          NSANJE /;
*--------------------------------------------------------------------------------------------------

ALIAS(JR,JRS);    ALIAS(JR,JRD);


SET  JIP  INPUTS AND PRODUCTS
*--------------------------------------------------------------------------------------------------
* INPUT/PROD..    DESCRIPTION.................      INPUT/PROD..
DESCRIPTION.................
*--------------------------------------------------------------------------------------------------
 /MAIZESEED     MAIZE SEED,                        RICESEED      RICE SEED,
  CASSAVASEED   CASSAVA SEED,                      POTATOSEED    POTATO SEED,
  BEANSSEED     BEANS SEED,                        GNUTSSEED     GROUNDNUTS SEED,
  FERT1         FERTLIZER BASAL,                   FERT2         FERTILIZER TOP
DRESSING,
  PESTICIDES    PESTICIDES,                        TRANSPORT     TRANSPORT,
  PACKAGINGM    PACKAGING MATERIALS,               LABOR         LABOR USE,
  MAIZE         MAIZE,                             RICE          RICE,
  CASSAVA       CASSAVA,                           POTATOES      POTATOES,
  BEANS         BEANS,                             GROUNDNUTS    GROUNDNUTS,
  CL-MAIZE      MAIZE CROPLAND,                    CL-RICE       RICE CROPLAND,
```

```
  CL-CASSAVA      CASSAVA CROPLAND,                  CL-POTATOES     POTATOES
CROPLAND,
  CL-BEANS        BEANS CROPLAND,                    CL-GROUNDNUTS   GROUNDNUTS
CROPLAND,
  CL-TOTAL        TOTAL CROPLAND,                    CONVEX          CONVEXITY
  BLANK-I,                                           BLANK-P/;
*-------------------------------------------------------------------------------------
------------------


SET  JI(JIP)  INPUTS  /MAIZESEED, RICESEED, CASSAVASEED, POTATOSEED, BEANSSEED,
GNUTSSEED, FERT1, FERT2,
                       PESTICIDES, TRANSPORT, PACKAGINGM, LABOR, CL-MAIZE, CL-RICE,
CL-CASSAVA,
                       CL-POTATOES, CL-BEANS, CL-GROUNDNUTS,CL-TOTAL, CONVEX, BLANK-
I/;

SET  JIR(JI)  REGIONAL INPUTS  /MAIZESEED, RICESEED, CASSAVASEED, POTATOSEED,
BEANSSEED, GNUTSSEED,
                                FERT1, FERT2, PESTICIDES, TRANSPORT, PACKAGINGM,
LABOR/;
$EJECT;
*    CROP LAND INPUTS PLAY A UNIQUE ROLE IN MAGS.  THEY DEFINE THE LEVEL OF
AGGREGATION FOR THE CROP
*    PRODUCTION ACTIVITIES.  RELEVANT CROP BUDGETS ARE CREATED FOR EACH TYPE OF CROP
LAND.  CROP
*    PRODUCTION ACTIVITIES ARE CONSTRUCTED USING THESE BUDGETS, AND CROP PRODUCTION
ACTIVITIES ARE
*    MAPPED TO EACH REGION IN WHICH THE CORRESPONDING LAND TYPE OCCURS.

SET  JIT  CROP LAND TYPES  /CLT-CHI, CLT-KAR, CLT-RUM, CLT-KHA, CLT-MZI, CLT-KAS, CLT-
MCH, CLT-DOW, CLT-NTCH, CLT-LL, CLT-DED, CLT-NU, CLT-KK, CLT-SA, CLT-MAN,
                            CLT-MAC, CLT-ZA, CLT-BLK, CLT-BT, CLT-TO, CLT-CZ, CLT-PHA,
CLT-MU, CLT-MWA, CLT-NEN, CLT-CHK, CLT-NSA/;

SET  JIC(JI)  CROP LAND INPUTS  /CL-MAIZE, CL-RICE, CL-CASSAVA, CL-POTATOES, CL-BEANS,
CL-GROUNDNUTS,
                                 CL-TOTAL, CONVEX/;

SET  JIN(JI)  NONREGIONAL INPUTS /BLANK-I/;


SET  JRIR(JR,JIR)  REGIONAL INPUTS MAPPED TO REGIONS;

     JRIR(JR,JIR) = YES;

SET  JRT(JR,JIT)  CROP LAND TYPES MAPPED TO REGIONS  /CHI.CLT-CHI, KAR.CLT-KAR,
RUM.CLT-RUM, KHA.CLT-KHA, MZI.CLT-MZI,
                  KAS.CLT-KAS, MCH.CLT-MCH, DOW.CLT-DOW, NTCH.CLT-NTCH, LL.CLT-LL,
DED.CLT-DED, NU.CLT-NU, KK.CLT-KK, SA.CLT-SA, MAN.CLT-MAN,
                  MAC.CLT-MAC, ZA.CLT-ZA, BLK.CLT-BLK, BT.CLT-BT, TO.CLT-TO, CZ.CLT-
CZ, PHA.CLT-PHA, MU.CLT-MU, MWA.CLT-MWA, NEN.CLT-NEN,
                  CHK.CLT-CHK, NSA.CLT-NSA /;

SET  JRIC(JR,JIC)  CROP LAND INPUTS MAPPED TO REGIONS;

     JRIC(JR,JIC) = YES;
$EJECT;
SET  JP(JIP)  PRODUCTS  /MAIZE, RICE, CASSAVA, POTATOES, BEANS, GROUNDNUTS, BLANK-P/;

SET  JPR(JP)  REGIONAL PRODUCTS  /MAIZE, RICE, CASSAVA, POTATOES, BEANS, GROUNDNUTS/;

SET  JPN(JP)  NONREGIONAL PRODUCTS  /BLANK-P/;
```

```
SET  JRPR(JR,JP)  REGIONAL PRODUCTS MAPPED TO REGIONS;

*    MAP ALL REGIONAL PRODUCTS TO ALL REGIONS, THEN IDENTIFY EXCEPTIONS

     JRPR(JR,JPR) = YES;

     JRPR("TO","RICE") = NO;
     JRPR("NTCH","RICE") = NO;

*    CANCEL MAPPINGS OF NONREGIONAL PRODUCTS TO REGIONS

     JRPR(JR,JP)$(NOT JPR(JP)) = NO;

SET  JRSD(JRS,JRD)  INTER-REGIONAL TRANSPORT ROUTES - DESTINATIONS MAPPED TO SOURCES;

*    ALLOW TRANSPORT BETWEEN ALL REGIONS BUT PRECLUDE ROUTES W THE SAME SOURCE AND
DESTINATION

     JRSD(JRS,JRD) = YES;   JRSD(JRS,JRS) = NO;

SET  JTIP(JIP)  TRANSPORTED REGIONAL INPUTS AND PRODUCTS  /MAIZESEED, RICESEED,
CASSAVASEED, POTATOSEED,
                BEANSSEED, GNUTSSEED, FERT1, FERT2, PESTICIDES, MAIZE, RICE, CASSAVA,
POTATOES, BEANS,
                GROUNDNUTS/;

SET  JRSDIP(JRS,JRD,JIP)  INPUTS AND PRODUCTS MAPPED TO SOURCES AND DESTINATIONS;
     JRSDIP(JRS,JRD,JIP) = YES;


*    THE FOLLOWING REGIONAL PRODUCTS MAY BE SHIPPED BETWEEN ANY TWO REGIONS IN WHICH
THEY OCCUR

*     JRSDIP(JRS,JRD,"MAIZE") = YES;            JRSDIP(JRS,JRD,"BEANS") = YES;
*     JRSDIP(JRS,JRD,"RICE") = YES;             JRSDIP(JRS,JRD,"GROUNDNUTS") = YES;
*     JRSDIP(JRS,JRD,"CASSAVA") = YES;
*     JRSDIP(JRS,JRD,"POTATOES") = YES;

*    INPUTS AND PRODUCTS THAT ARE NOT TRANSPORTED ARE BELOW EXCLUDED FROM MAPPINGS TO
ROUTES.  ALSO
*    EXCLUDED ARE SHIPMENTS WHERE THE ROUTE IS NOT AN IDENTIFIED SHIPPING ROUTE, OR IF
THE SOURCE OR
*    DESTINATION HAS NOT MARKET FOR THE ITEM

     JRSDIP(JRS,JRD,JIP)$(NOT JTIP(JIP)) = NO;
     JRSDIP(JRS,JRD,JTIP)$(NOT JRSD(JRS,JRD)) = NO;
     JRSDIP(JRS,JRD,JIR)$((NOT JRIR(JRS,JIR)) OR (NOT JRIR(JRD,JIR))) = NO;
     JRSDIP(JRS,JRD,JPR)$((NOT JRPR(JRS,JPR)) OR (NOT JRPR(JRD,JPR))) = NO;


$EJECT;
SET  JX  PRODUCTION ACTIVITIES
*----------------------------------------------------------------------------------------
------------------
* ACTIVITY....  DESCRIPTION......................     ACTIVITY....
DESCRIPTION......................
*----------------------------------------------------------------------------------------
------------------
 /X-MAIZE      MAIZE PRODUCTION,                        X-RICE        RICE PRODUCTION,
  X-CASSAVA    CASSAVA PRODUCTION,                      X-POTATOES    POTATOES
PRODUCTION,
```

```
  X-BEANS       BEANS PRODUCTION,                      X-GROUNDNUTS  GROUNDNUTS
PRODUCTION,
  X-MIX-00      CROP MIX IN 2000,                      X-MIX-01      CROP MIX IN 2001,
  X-MIX-02      CROP MIX IN 2002,                      X-MIX-03      CROP MIX IN 2003,
  X-MIX-04      CROP MIX IN 2004,                      X-MIX-05      CROP MIX IN 2005,
  X-MIX-06      CROP MIX IN 2006,                      X-MIX-07      CROP MIX IN 2007,
  X-MIX-08      CROP MIX IN 2008,                      X-MIX-09      CROP MIX IN
2009/;
*-------------------------------------------------------------------------------------
------------------


SET  JXR(JX)  REGIONAL PRODUCTION ACTIVITIES  /X-MAIZE, X-RICE, X-CASSAVA, X-POTATOES,
X-BEANS,
              X-GROUNDNUTS, X-MIX-00,X-MIX-01,X-MIX-02,X-MIX-03,X-MIX-04,X-MIX-05,X-
MIX-06,X-MIX-07,X-MIX-08,X-MIX-09/;

SET  JXC(JXR)  CROP PRODUCTION ACTIVITIES  /X-MAIZE,X-RICE,X-CASSAVA,X-POTATOES,X-
BEANS,X-GROUNDNUTS,
     X-MIX-00,X-MIX-01,X-MIX-02,X-MIX-03,X-MIX-04,X-MIX-05,X-MIX-06,X-MIX-07,X-MIX-
08,X-MIX-09/;

SET  JXL(JXR)  LIVESTOCK PRODUCTION ACTIVITIES;

SET  JXP(JXR)  PROCESSING ACTIVITIES;

SET  JXO(JX)  OTHER ACTIVITIES;

* Maxwell
*SET YR YEAR /00*09/;
SET  JRX(JR,JXR)  PRODUCTION ACTIVITIES MAPPED TO REGIONS;

*    MAP ALL REGIONAL PROD ACTIVITIES TO ALL REGIONS, THEN EXCLUDE EXCEPTION

     JRX(JR,JXR) = YES;

     JRX("NTCH","X-RICE") = NO;   JRX("TO","X-RICE") = NO;


SET  JXCT(JXC,JIT)  CROP LAND TYPES MAPPED TO CROP PROD ACTIVITIES;
*    MAP ALL LAND TYPES TO ALL CROP PRODUCTION ACTIVITIES , THEN EXCLUDE EXCEPTION
     JXCT(JXC,JIT) = YES;

SET  JRXT(JR,JXC,JIT)  CROP LAND TYPES MAPPED TO CROP PROD ACTIVITIES AND REGIONS;

*    EXCLUDE ALL COMBINATIONS, THEN MAP LAND TYPES AND CROP PROD ACT WHEN BOTH ARE
MAPPED TO THAT REGION
     JRXT(JR,JXC,JIT) = NO;
     JRXT(JR,JXC,JIT)$(JRT(JR,JIT) AND JRX(JR,JXC) AND JXCT(JXC,JIT)) = YES;

SET  JXN(JX)  NONREGIONAL PRODUCTION ACTIVITIES;
     JXN(JX) = NO;

$EJECT;
SET  JB  BUDGETS USED TO CONSTRUCT CROP AND LIVESTOCK PRODUCTION ACTIVITIES
*-------------------------------------------------------------------------------------
------------------
* BUDGET......  DESCRIPTION.......................     BUDGET......
DESCRIPTION.......................
*-------------------------------------------------------------------------------------
------------------
 /B-MAIZE       MAIZE PRODUCTION,                      B-MIX-04      CROP MIX IN 2004,
  B-RICE        RICE PRODUCTION,                       B-MIX-05      CROP MIX IN 2005,
```

```
  B-CASSAVA     CASSAVA PRODUCTION,                         B-MIX-06      CROP MIX IN 2006,
  B-POTATOES    POTATOES PRODUCTION,                        B-MIX-07      CROP MIX IN 2007,
  B-BEANS       BEANS PRODUCTION,                           B-MIX-08      CROP MIX IN 2008,
  B-GROUNDNUTS  GROUNDNUTS PRODUCTION,                      B-MIX-09      CROP MIX IN 2009,
  B-MIX-00      CROP MIX IN 2000,                           B-MIX-01      CROP MIX IN 2001,
  B-MIX-02      CROP MIX IN 2002,                           B-MIX-03      CROP MIX IN
2003/;
*-------------------------------------------------------------------------------------
------------------

SET  JBR(JR,JB)  BUDGETS MAPPED TO REGIONS;
     JBR(JR,JB) = YES;

SET  JBC(JB)  CROP BUDGETS USED TO CONSTRUCT CROP PRODUCTION ACTIVITIES  /B-MAIZE, B-
RICE, B-CASSAVA,
     B-POTATOES, B-BEANS, B-GROUNDNUTS, B-MIX-00*B-MIX-09/;

* Maxwell, 2nd July
*SET  YR2(JBC) CROP AREA BY YEAR /B-MIX-00*B-MIX-09/

SET YEAR HISTORICAL MIXES YEARS /2000*2009 /;


SET  JBL(JB)  LIVESTOCK BUDGETS USED TO CONSTRUCT LIVESTOCK PRODUCTION ACTIVITIES;

SET  JBCT(JBC,JIT)  CROP BUDGETS MAPPED TO CROP LAND TYPES;
*    MAP ALL CROP BUDGETS TO CROP LAND TYPES THEN EXCLUDE EXCEPTIONS
     JBCT(JBC,JIT)=YES;

SET  JBCR(JR,JXC,JBC)  CROP BUDGETS MAPPED TO CROP PROD ACTIVITIES AND REGIONS;
     JBCR(JR,JXC,JBC) = NO;

     JBCR(JR,JXC,JBC)
     $(JRX(JR,JXC) AND JBR(JR,JBC) AND
      (SUM(JIT$(JRT(JR,JIT) AND JRX(JR,JXC) AND JBCT(JBC,JIT)),1)>0)) = YES;

SET JOC  OTHER PRODUCTION ACTIVITY CONSTRAINTS;
SET JOCR(JOC) OTHER REGIONAL PRODUCTION ACTIVITY CONSTRAINTS;
SET JROCR(JR,JOCR)  OTHER REGIONAL CONSTRAINTS MAPPED TO REGIONS;
SET JOCN(JOC)  OTHER NONREGIONAL PRODUCTION ACTIVITY CONSTRAINTS;


$EJECT;
SET  JU  UNITS OF MEASURE
*-------------------------------------------------------------------------------------
------------------
* UNITS........  DESCRIPTION.....................    UNITS........
DESCRIPTION.....................
*-------------------------------------------------------------------------------------
------------------
 /AC            ACRES,                                     AC1000        THOUSAND ACRES,
  HA            HECTARES,                                  HA1000        THOUSAND
HECTARES,
  LB            POUNDS,                                    CWT           HUNDREDWEIGHT,
  T-US          US TONS,                                   T-US1000      THOUSAND US TONS,
  KG            KILOGRAMS,                                 MT            METRIC TONS,
  MT1000        THOUSAND METRIC TONS,                      BU            BUSHELS,
  BU1000        THOUSAND BUSHELS,                          BU-CORN       BUSHELS OF CORN
GRAIN,
  BU-SOY        BUSHELS OF SOYBEANS,                       BU-WHT        BUSHELS OF WHEAT,
  BG-CS         BAGS OF CORN SEED,                         GAL           GALLONS,
  GAL1000       THOUSAND GALLONS,                          MGAL          MILLION GALLONS,
  L             LITERS,                                    ML            MILLION LITERS,
```

```
  BRL           BARRELS,                          MBRL          MILLION BARRELS,
  HD            HEAD,                             HD1000        THOUSAND HEAD,
  LTR           LITTERS,                          LTR1000       THOUSAND LITTERS,
  HR            HOURS,                            HR1000        THOUSAND HOURS,
  KW            KILOWATTS - THOUSAND WATTS,       MW            MEGAWATTS -
MILLION WATTS,
  GW            GIGAWATTS - BILLION WATTS,        USD           DOLLARS,
  USD1000       THOUSAND DOLLARS,                 AN-USD        ANNUAL DOLLARS,
  AN-USD1000    THOUSAND ANNUAL DOLLARS,          OTHER         OTHER UNITS
CONVERT ONLY TO OTHER

  MANDAYS       PERSON DAYS OF LABOR
  KWACHA        MALAWI KWACHA CURRENCY
  50KG          50 KILOGRAMS /;
*-------------------------------------------------------------------------------------
------------------

ALIAS(JU,JU2);

SET  JUX(JX,JU)  UNITS MAPPED TO PRODUCTION ACTIVITIES
/ X-MAIZE.HA, X-RICE.HA, X-CASSAVA.HA, X-POTATOES.HA, X-BEANS.HA, X-GROUNDNUTS.HA, X-
MIX-00.OTHER,
  X-MIX-01.OTHER, X-MIX-02.OTHER, X-MIX-03.OTHER, X-MIX-04.OTHER, X-MIX-05.OTHER, X-
MIX-06.OTHER,
  X-MIX-07.OTHER, X-MIX-08.OTHER, X-MIX-09.OTHER /;

SET  JUIPX(JIP,JU)  UNITS OF MEASURE MAPPED TO INPUTS & PRODUCTS FOR CONST &
ACTIVITIES
/ MAIZESEED.KG, RICESEED.KG, CASSAVASEED.KG, POTATOSEED.KG, BEANSSEED.KG,
GNUTSSEED.KG, FERT1.50KG,
  FERT2.50KG, PESTICIDES.L,LABOR.MANDAYS, MAIZE.KG, RICE.KG, CASSAVA.KG, POTATOES.KG,
BEANS.KG,
  GROUNDNUTS.KG, TRANSPORT.KWACHA, PACKAGINGM.KWACHA/;

SET  JUB(JB,JU)  UNITS MAPPED TO BUDGETS

/ B-MAIZE.HA, B-RICE.HA, B-CASSAVA.HA, B-POTATOES.HA, B-BEANS.HA, B-GROUNDNUTS.HA, B-
MIX-00.OTHER,
  B-MIX-09.OTHER, B-MIX-01.OTHER, B-MIX-02.OTHER, B-MIX-03.OTHER, B-MIX-04.OTHER, B-
MIX-05.OTHER,
  B-MIX-06.OTHER, B-MIX-07.OTHER, B-MIX-08.OTHER/;

SET  JUIPB(JIP,JU)  UNITS OF MEASURE MAPPED TO INPUTS AND PRODUCTS FOR BUDGETS
/ MAIZESEED.KG,RICESEED.KG, CASSAVASEED.KG, POTATOSEED.KG, BEANSSEED.KG, GNUTSSEED.KG,
FERT1.50KG,
  FERT2.50KG, PESTICIDES.L,LABOR.MANDAYS, MAIZE.KG, RICE.KG, CASSAVA.KG, POTATOES.KG,
BEANS.KG,
  GROUNDNUTS.KG, TRANSPORT.KWACHA, PACKAGINGM.KWACHA/;


SET  JSDP  SUPPLY & DEMAND PARAMETERS  /INTERCEPT, SLOPE, PBAR, QBAR, ELAST, QMIN,
QMAX/;

SET  JXMP  EXPORT & IMPORT PARAMETERS  /X-PRICE, X-MIN, X-MAX, M-PRICE, M-MIN, M-MAX/;
$STITLE PARAMETER DECLARATIONS AND ASSIGNMENTS
* OVERVIEW OF MODEL PARAMETERS
*-------------------------------------------------------------------------------------
------------------
* PARAMETER............
DESCRIPTION..............................................................
*-------------------------------------------------------------------------------------
------------------
* CNV(JU,JU2)            CONVERSION FACTOR JU TO JU2
```

```
* CNV1(JBC,JXC)          CONVERSION MULTIPLIER BUDGET UNITS TO ACTIVITY UNITS
* CNV2(JIP)              CONVERSION MULTIPLIER FOR INPUT OR PRODUCT BUDGET TO ACTIVITY
*
* CSR(JR,JIR,JSDP)       REGIONAL INPUT SUPPLY PARAMETERS
* CSC(JR,JIT,JIC,JSDP)   CROP LAND INPUT SUPPLY PARAMETERS
* CSN(JIN,JSDP)          NONREGIONAL INPUT SUPPLY PARAMETERS
*
* CDR(JR,JPR,JSDP)       REGIONAL PRODUCT DEMAND PARAMETERS
* CDN(JPN,JSDP)          NONREGIONAL PRODUCT DEMAND PARAMETERS
*
* CXMR(JR,JIP,JXMP)      REGIONAL IMPORT AND EXPORT PARAMETERS
* CXMN(JIP,JXMP)         NONREGIONAL IMPORT AND EXPORT PARAMETERS
*
* DX(JR)                 X COORDINATE FOR REGION
* DY(JR)                 Y COORDINATE FOR REGION
* DT(JRS,JRD)            DISTANCE IN MILES BETWEEN REGIONS
*
* CTIP(JIP)              UNIT TRANSPORTATION COST PER MILE BY PRODUCT
* CTR(JRS,JRD,JIP)       UNIT TRANSPORTATION COST BY SOURCE, DESTINATION AND PRODUCT
*
* AC(JIT,JBC,JIP)        PRODUCT AND INPUT COEFFICIENTS FOR CROP BUDGETS
* AOC(JIT,JBC,JOC)       OTHER CONSTRAINT COEFFICIENTS FOR CROP BUDGETS
* ACM(JR,JBC,JXC)        IO COEF MULTIPLIER CROP BUDGET TO PRODUCTION ACTIVITY
*
* AL(JR,JBL,JIP)         PRODUCT AND INPUT COEFFICIENTS FOR LIVESTOCK BUDGETS
* AOL(JR,JBL,JOC)        OTHER CONSTRAINT COEFFICIENTS FOR LIVESTOCK BUDGETS
* ALM(JR,JBL,JXL)        IO COEF MULTIPLIER LIVESTOCK BUDGET TO PRODUCTION ACTIVITY
*
* ARP(JR,JXP,JIP)        PRODUCT AND INPUT COEF FOR REGIONAL PROCESSING ACTIVITIES
* AORP(JR,JXP,JOCR)      OTHER CONSTRAINT COEF FOR REGIONAL PROCESSING ACTIVITIES
*
* AR(JR,JXR,JIP)         PRODUCT AND INPUT COEF FOR REGIONAL PRODUCTION ACTIVITIES
* ARC(JR,JXC,JIT,JIP)    PRODUCT AND INPUT COEF FOR REGIONAL CROP PRODUCTION
ACTIVITIES
* AOR(JR,JXR,JOCR)       OTHER CONSTRAINT COEF FOR REGIONAL PRODUCTION ACTIVITIES
* AORC(JR,JXC,JIT,JOCR)  OTHER CONSTRAINT COEF FOR REGIONAL CROP PRODUCTION ACTIVITIES
* BOR(JR,JOCR)           RIGHTHAND SIDES FOR OTHER REGIONAL CONSTRAINTS
*
* AN(JXN,JIP)            PRODUCT AND INPUT COEF FOR NONREGIONAL PRODUCTION ACTIVITIES
* AON(JXN,JOCN)          OTHER CONSTRAINT COEF FOR NONREGIONAL PRODUCTION ACTIVITIES
* BON(JOCN)              RIGHTHAND SIDES FOR OTHER NONREGIONAL CONSTRAINTS
*-------------------------------------------------------------------------------------
------------------
$EJECT
SCALAR  ACHA  ACRES PER HECTARE  /2.469955/;       SCALAR  LBKG  POUNDS PER KILOGRAM
/2.20462/;
SCALAR  GL  GALLONS PER LITER  /0.264172/;         SCALAR  MKM  MILES PER KILOMETER
/1.609344/;

PARAMETER  CNV(JU,JU2)  MULTIPLIER FOR CONVERSION OF UNITS JU TO UNITS JU2;

CNV(JU,JU2) = 0;                              CNV(JU,JU) = 1;

CNV("AC","AC1000") = 1/1000;                  CNV("AC1000","AC") = 1000;
CNV("AC","HA") = 1/ACHA;                      CNV("HA","AC") = ACHA;
CNV("AC","HA1000") = 1/(ACHA*1000);           CNV("HA1000","AC") = ACHA*1000;
CNV("AC1000","HA") = 1000/ACHA;               CNV("HA","AC1000") = ACHA/1000;
CNV("AC1000","HA1000") = 1/ACHA;              CNV("HA1000","AC1000") = ACHA;
CNV("HA","HA1000") = 1/1000;                  CNV("HA1000","HA") = 1000;

CNV("LB","CWT") = 1/100;                      CNV("CWT","LB") = 100;
CNV("LB","T-US") = 1/2000;                    CNV("T-US","LB") = 2000;
CNV("LB","T-US1000") = 1/(1000*2000);         CNV("T-US1000","LB") = 1000*2000;
```

```
CNV("CWT","T-US") = 100/2000;                CNV("T-US","CWT") = 2000/100;
CNV("CWT","T-US1000") = 100/(1000*2000);     CNV("T-US1000","CWT") = (1000*2000)/100;
CNV("CWT","KG") = 100/LBKG;                  CNV("KG","CWT") = LBKG/100;
CNV("CWT","MT") = 100/(LBKG*1000);           CNV("MT","CWT") = (1000*LBKG)/100;
CNV("CWT","MT1000") = 100/(LBKG*1000*1000);  CNV("MT1000","CWT") =
(1000*1000*LBKG)/100;

CNV("T-US","T-US1000") = 1/1000;             CNV("T-US1000","T-US") = 1000;
CNV("LB","KG") = 1/LBKG;                     CNV("KG","LB") = LBKG;
CNV("LB","MT") = 1/(1000*LBKG);              CNV("MT","LB") = 1000*LBKG;
CNV("LB","MT1000") = 1/(1000*1000*LBKG);     CNV("MT1000","LB") = 1000*1000*LBKG;
CNV("T-US","KG") = 2000/LBKG;                CNV("KG","T-US") = LBKG/2000;
CNV("T-US","MT") = 2000/(1000*LBKG);         CNV("MT","T-US") = 1000*LBKG/2000;
CNV("T-US","MT1000") = 2000/(1000*1000*LBKG); CNV("MT1000","T-US") =
1000*1000*LBKG/2000;
CNV("T-US1000","KG") = 1000*2000/LBKG;       CNV("KG","T-US1000") = LBKG/(1000*2000);
CNV("T-US1000","MT") = 2000/LBKG;            CNV("MT","T-US1000") = LBKG/2000;
CNV("T-US1000","MT1000") = 2000/(1000*LBKG); CNV("MT1000","T-US1000") =
1000*LBKG/2000;
CNV("KG","MT") = 1/1000;                     CNV("MT","KG") = 1000;
CNV("KG","MT1000") = 1/(1000*1000);          CNV("MT1000","KG") = 1000*1000;
CNV("MT","MT1000") = 1/1000;                 CNV("MT1000","MT") = 1000;

CNV("BU","BU1000") = 1/1000;                 CNV("BU1000","BU") = 1000;

CNV("BU-CORN","LB") = 56;                    CNV("LB","BU-CORN") = 1/56;
CNV("BU-CORN","T-US") = 56/2000;             CNV("T-US","BU-CORN") = 2000/56;
CNV("BU-CORN","T-US1000") = 56/(2000*1000);  CNV("T-US1000","BU-CORN") =
(2000*1000)/56;
CNV("BU-CORN","KG") = 56/LBKG;               CNV("KG","BU-CORN") = LBKG/56;
CNV("BU-CORN","MT") = 56/(LBKG*1000);        CNV("MT","BU-CORN") = (LBKG*1000)/56;
CNV("BU-CORN","MT1000")=56/(1000*1000*LBKG); CNV("MT1000","BU-
CORN")=(1000*1000*LBKG)/56;

CNV("BU-SOY","LB") = 60;                     CNV("LB","BU-SOY") = 1/60;
CNV("BU-SOY","T-US") = 60/2000;              CNV("T-US","BU-SOY") = 2000/60;
CNV("BU-SOY","T-US1000") = 60/(2000*1000);   CNV("T-US1000","BU-SOY") =
(2000*1000)/60;
CNV("BU-SOY","KG") = 60/LBKG;                CNV("KG","BU-SOY") = LBKG/60;
CNV("BU-SOY","MT") = 60/(LBKG*1000);         CNV("MT","BU-SOY") = (LBKG*1000)/60;
CNV("BU-SOY","MT1000")=60/(1000*1000*LBKG);  CNV("MT1000","BU-
SOY")=(1000*1000*LBKG)/60;

CNV("BU-WHT","LB") = 60;                     CNV("LB","BU-WHT") = 1/60;
CNV("BU-WHT","T-US") = 60/2000;              CNV("T-US","BU-WHT") = 2000/60;
CNV("BU-WHT","T-US1000") = 60/(2000*1000);   CNV("T-US1000","BU-WHT") =
(2000*1000)/60;
CNV("BU-WHT","KG") = 60/LBKG;                CNV("KG","BU-WHT") = LBKG/60;
CNV("BU-WHT","MT") = 60/(LBKG*1000);         CNV("MT","BU-WHT") = (LBKG*1000)/60;
CNV("BU-WHT","MT1000")=60/(1000*1000*LBKG);  CNV("MT1000","BU-
WHT")=(1000*1000*LBKG)/60;

CNV("BG-CS","LB") = 50;                      CNV("LB","BG-CS") = 1/50;
CNV("BG-CS","T-US") = 50/2000;               CNV("T-US","BG-CS") = 2000/50;
CNV("BG-CS","T-US1000") = 50/(2000*1000);    CNV("T-US1000","BG-CS") =
(2000*1000)/50;
CNV("BG-CS","KG") = 50/LBKG;                 CNV("KG","BG-CS") = LBKG/50;
CNV("BG-CS","MT") = 50/(LBKG*1000);          CNV("MT","BG-CS") = (LBKG*1000)/50;
CNV("BG-CS","MT1000") = 50/(1000*1000*LBKG); CNV("MT1000","BG-CS") =
(1000*1000*LBKG)/50;

CNV("GAL","GAL1000") = 1/1000;               CNV("GAL1000","GAL") = 1000;
CNV("GAL","MGAL") = 1/1000000;               CNV("MGAL","GAL") = 1000000;
```

```
CNV("GAL","L") = 1/GL;                      CNV("L","GAL") = GL;
CNV("GAL","ML") = 1/(GL*1000000);           CNV("ML","GAL") = 1000000*GL;
CNV("GAL1000","MGAL") = 1/1000;             CNV("MGAL","GAL1000") = 1000;
CNV("GAL1000","L") = 1000/GL;               CNV("L","GAL1000") = GL/1000;
CNV("GAL1000","ML") = 1000/(GL*1000000);    CNV("ML","GAL1000") = (GL*1000000)/1000;
CNV("MGAL","L") = 1000000/GL;               CNV("L","MGAL") = GL/1000000;
CNV("MGAL","ML") = 1/GL;                    CNV("ML","MGAL") = GL;
CNV("L","ML") = 1/1000000;                  CNV("ML","L") = 1000000;
CNV("BRL","MBRL") = 1/1000000;              CNV("MBRL","BRL") = 1000000;

CNV("HD","HD1000") = 1/1000;                CNV("HD1000","HD") = 1000;
CNV("LTR","LTR1000") = 1/1000;              CNV("LTR1000","LTR") = 1000;

CNV("KW","MW") = 1/1000;                    CNV("MW","KW") = 1000;
CNV("KW","GW") = 1/1000000;                 CNV("GW","KW") = 1000000;
CNV("MW","GW") = 1/1000;                    CNV("GW","MW") = 1000;

CNV("USD","USD1000") = 1/1000;              CNV("USD1000","USD") = 1000;
CNV("AN-USD","AN-USD1000") = 1/1000;        CNV("AN-USD1000","AN-USD") = 1000;


PARAMETERS  CNV1(JB,JX)  CONVERSION MULTIPLIER BUDGET UNITS TO PROD ACT UNITS,
            CNV2(JIP)  CONVERSION MULTIPLIER BY ITEM BUDGET TO CONSTRANT UNITS;

            CNV1(JB,JX) = SUM(JU$JUB(JB,JU),SUM(JU2$JUX(JX,JU2),CNV(JU,JU2)));
            CNV2(JIP) = SUM(JU$JUIPB(JIP,JU),SUM(JU2$JUIPX(JIP,JU2),CNV(JU,JU2)));

$STITLE SUPPLY AND DEMAND PARAMETER DECLARATIONS AND ASSIGNMENTS
* REGIONAL AND NONREGIONAL INPUTS, AND CROP LAND INPUTS, HAVE SUPPLY ACTIVITIES.
SINCE THE OBJECTIVE
* FUNCTION OF THE MODEL IS CONSUMER PLUS PRODUCER SURPLUS, THE OBJECTIVE FUNCTION
ENTRY FOR AN INPUT
* SUPPLY ACTIVITY IS MINUS THE SUPPLY FUNCTION INTEGRAL.  SUPPLY FUNCTIONS ARE ASSUMED
TO BE LINEAR IN
* THIS VERSION OF MAGS, HOWEVER RELAXATION OF THAT ASSUMPTION IS STRAIGHT FORWARD AND
CAN BE ACCOMODATED
* WITH GAMS.  PARAMETERS ARE ASSIGNED OR COMPUTED IN PRICE DEPENDENT FORM.  AN
INFINITELY ELASTIC SUPPLY
* IS INDICATED BY A ZERO SLOPE WITH THE INTERCEPT THEN BECOMING THE CONSTANT MARKET
PRICE.  IF IN THE
* FOLLOWING TABLES BOTH THE INTERCEPT AND SLOPE ARE ZERO, OR EQUIVALENTLY BLANK, THE
SLOPE AND INTERCEPT
* WILL BE COMPUTED WITH THE USER ASSIGNED PRICE, QUANTITY AND ELASTICITY (PBAR, QBAR
AND ELAST,
* RESPECTIVELY) IF ALL THREE OF THESE PARAMETERS ARE STRICTLY POSITIVE.
*
* FIXED SUPPLIES CAN BE MODELED BY SETTING APPROPRIATE VALUES FOR QMIN AND QMAX WHICH
ARE USED AS LOWER
* AND UPPER BOUNDS, RESPECTIVELY, ON THE SUPPLY ACTIVITIES.  IF EXOGENOUS SUPPLY
FUNCTIONS ARE TO BE
* USED WITHOUT OTHER LIMITS TO SUPPLY, QMIN SHOULD BE SET TO ZERO AND QMAX SHOULD BE
SET TO INFINITY
* (INF).  A POSITIVE QMIN AND OR A FINITE QMAX MAY BE SET, IF DESIRED, WITH AN
EXOGENOUS SUPPLY FUNCTION
* THAT IS PRICE RESPONSIVE OR INFINITELY ELASTIC FROM QMIN TO QMAX.
*
* PRICES ARE IN DOLLARS OR OTHER BASE CURRENCY USED IN THE MODEL.  QUANTITIES ARE IN
THE UNITS DEFINED
* BY THE USER FOR THE INPUT'S SUPPLY ACTIVITY AND CONSTRAINT.
*


TABLE  CSR(JR,JIR,JSDP)  REGIONAL INPUT SUPPLY PARAMETERS
```

```
$OFFLISTING;
*$INCLUDE C:\0 Teaching\ApEc 8202\Exams\CSR.INC
$INCLUDE G:\My Drive\MalawiAgSectorModel\GAMS Model\CSR8thNov.INC
;
$ONLISTING;

* CHECK TO SEE IF ELAST PBAR AND QBAR ARE TO BE USED TO CALCULATE THE INTERCEPT &
SLOPE FOR THE INVERSE
* SUPPLY FUNCTION FOR EACH REGIONAL INPUT IN EACH REGION. IF INDICATED, CALCULATE THE
SLOPE & INTERCEPT.

  SET CHECKCSR(JR,JIR);
  CHECKCSR(JR,JIR)=NO;
  CHECKCSR(JR,JIR)$((CSR(JR,JIR,"SLOPE")=0) AND (CSR(JR,JIR,"INTERCEPT")=0)
           AND (CSR(JR,JIR,"PBAR")>0) AND (CSR(JR,JIR,"QBAR")>0) AND
(CSR(JR,JIR,"ELAST")>0)) = YES;
  CSR(JR,JIR,"SLOPE")$CHECKCSR(JR,JIR) =
CSR(JR,JIR,"PBAR")/(CSR(JR,JIR,"QBAR")*CSR(JR,JIR,"ELAST"));
  CSR(JR,JIR,"INTERCEPT")$CHECKCSR(JR,JIR)
      = CSR(JR,JIR,"PBAR") - (CSR(JR,JIR,"PBAR")/CSR(JR,JIR,"ELAST"));




PARAMETER  CSC(JR,JIT,JIC,JSDP)  CROP LAND INPUT SUPPLY PARAMETERS;
           CSC(JR,JIT,JIC,JSDP) = 0.0;
           CSC(JR,JIT,JIC,"QMIN") = -INF;
*           CSC(JR,JIT,"CL-RICE","QMAX") = 20000;
           CSC(JR,JIT,"CONVEX","QMIN") = 1.0;
           CSC(JR,JIT,"CONVEX","QMAX") = 1.0;


* CHECK TO SEE IF ELAST PBAR AND QBAR ARE TO BE USED TO CALCULATE THE INTERCEPT &
SLOPE FOR THE INVERSE
* SUPPLY FUNCTION FOR EACH CROP LAND INPUT FOR EACH CROP LAND TYPE IN EACH REGION. IF
INDICATED,
* CALCULATE THE SLOPE & INTERCEPT.

$OFFTEXT
  SET CHECKCSC(JR,JIT,JIC);
  CHECKCSC(JR,JIT,JIC)=NO;
  CHECKCSC(JR,JIT,JIC)$((CSC(JR,JIT,JIC,"SLOPE")=0) AND
(CSC(JR,JIT,JIC,"INTERCEPT")=0)
           AND (CSC(JR,JIT,JIC,"PBAR")>0) AND (CSC(JR,JIT,JIC,"QBAR")>0)
           AND (CSC(JR,JIT,JIC,"ELAST")>0)) = YES;
  CSC(JR,JIT,JIC,"SLOPE")$CHECKCSC(JR,JIT,JIC)
                        =
CSC(JR,JIT,JIC,"PBAR")/(CSC(JR,JIT,JIC,"QBAR")*CSC(JR,JIT,JIC,"ELAST"));
  CSC(JR,JIT,JIC,"INTERCEPT")$CHECKCSC(JR,JIT,JIC)
      = CSC(JR,JIT,JIC,"PBAR") - (CSC(JR,JIT,JIC,"PBAR")/CSC(JR,JIT,JIC,"ELAST"));
$EJECT;
* REGIONAL AND NONREGIONAL PRODUCTS HAVE DEMAND ACTIVITIES.  SINCE THE OBJECTIVE
FUNCTION OF THE MODEL
* IS CONSUMER PLUS PRODUCER SURPLUS, THE OBJECTIVE FUNCTION ENTRY FOR A PRODUCT DEMAND
ACTIVITY IS THE
* DEMAND FUNCTION INTEGRAL.  DEMAND FUNCTIONS ARE ASSUMED TO BE LINEAR, HOWEVER
RELAXATION OF THAT
* ASSUMPTION IS STRAIGHT FORWARD AND CAN BE ACCOMODATED WITH GAMS.  PARAMETERS ARE
ASSIGNED OR COMPUTED
* IN PRICE DEPENDENT FORM.  AN INFINITELY ELASTIC DEMAND IS INDICATED BY A ZERO SLOPE
WITH THE INTERCEPT
* THEN BECOMING THE CONSTANT MARKET PRICE.  IF IN THE FOLLOWING TABLES BOTH THE
INTERCEPT AND SLOPE ARE
```

```
* ZERO, OR EQUIVALENTLY BLANK, THE SLOPE AND INTERCEPT WILL BE COMPUTED WITH THE USER
ASSIGNED PRICE,
* QUANTITY AND ELASTICITY (PBAR, QBAR AND ELAST, RESPECTIVELY) IF THE PBAR AND QBAR
PARAMETERS ARE
* POSITIVE AND THE ELAST PARAMETER IS NEGATIVE.
*
* FIXED DEMANDS CAN BE MODELED BY SETTING APPROPRIATE VALUES FOR QMIN AND QMAX WHICH
ARE USED AS LOWER
* AND UPPER LIMITS, RESPECTIVELY, ON THE DEMAND ACTIVITIES.  IF EXOGENOUS DEMAND
FUNCTIONS ARE TO BE
* USED WITHOUT OTHER LIMITS TO DEMAND, QMIN SHOULD BE SET TO ZERO AND QMAX SHOULD BE
SET TO INFINITY
* (INF).  A POSITIVE QMIN AND/OR A FINITE QMAX MAY BE SET, IF DESIRED, WITH AN
EXOGENOUS DEMAND FUNCTION
* THAT IS PRICE RESPONSIVE OR INFINITELY ELASTIC FROM QMIN TO QMAX.
*
* PRICES ARE IN DOLLARS OR OTHER BASE CURRENCY USED IN THE MODEL.  QUANTITIES ARE IN
THE UNITS DEFINED
* BY THE USER FOR THE PRODUCT'S DEMAND ACTIVITY AND CONSTRAINT.


TABLE  CDR(JR,JPR,JSDP)  REGIONAL PRODUCT DEMAND PARAMETERS
$OFFLISTING;
*$INCLUDE C:\0 Teaching\ApEc 8202\Exams\CDR.INC
$INCLUDE G:\My Drive\MalawiAgSectorModel\GAMS Model\CDR.INC
;
$ONLISTING;

* CHECK TO SEE IF ELAST PBAR AND QBAR ARE TO BE USED TO CALCULATE THE INTERCEPT &
SLOPE FOR THE INVERSE
* DEMAND FUNCTION FOR EACH REGIONAL PRODUCT IN EACH REGION. IF INDICATED, CALCULATE
THE SLOPE &
* INTERCEPT.


  SET CHECKCDR(JR,JPR);
  CHECKCDR(JR,JPR)=NO;
  CHECKCDR(JR,JPR)$((CDR(JR,JPR,"SLOPE")=0) AND (CDR(JR,JPR,"INTERCEPT")=0)
           AND (CDR(JR,JPR,"PBAR")>0) AND (CDR(JR,JPR,"QBAR")>0) AND
(CDR(JR,JPR,"ELAST")>0)) = YES;
  CDR(JR,JPR,"SLOPE")$CHECKCDR(JR,JPR) =
CDR(JR,JPR,"PBAR")/(CDR(JR,JPR,"QBAR")*CDR(JR,JPR,"ELAST"));
  CDR(JR,JPR,"INTERCEPT")$CHECKCDR(JR,JPR)
      = CDR(JR,JPR,"PBAR") - (CDR(JR,JPR,"PBAR")/CDR(JR,JPR,"ELAST"));

*PARAMETER  CDN(JPN,JSDP)  NONREGIONAL PRODUCT DEMAND PARAMETERS;
*           CDN(JPN,JSDP) = 0.0;
*           CDN(JPN,"QMAX") = INF;


* CHECK TO SEE IF ELAST PBAR AND QBAR ARE TO BE USED TO CALCULATE THE INTERCEPT &
SLOPE FOR THE INVERSE
* DEMAND FUNCTION FOR EACH NON-REGIONAL PRODUCT. IF INDICATED, CALCULATE THE SLOPE &
INTERCEPT.

*  SET CHECKCDN(JPN);
*  CHECKCDN(JPN)=NO;
*  CHECKCDN(JPN)$((CDN(JPN,"SLOPE")=0) AND (CDN(JPN,"INTERCEPT")=0)
*           AND (CDN(JPN,"PBAR")>0) AND (CDN(JPN,"QBAR")>0) AND (CDN(JPN,"ELAST")>0))
= YES;
*  CDN(JPN,"SLOPE")$CHECKCDN(JPN) =
CDN(JPN,"PBAR")/(CDN(JPN,"QBAR")*CDN(JPN,"ELAST"));
```

```
*  CDN(JPN,"INTERCEPT")$CHECKCDN(JPN) = CDN(JPN,"PBAR") -
(CDN(JPN,"PBAR")/CDN(JPN,"ELAST"));


$STITLE TRADE PARAMETER DECLARATIONS AND ASSIGNMENTS
* FOR TRADED PRODUCTS AND INPUTS, IMPORT AND EXPORT PRICES IN THE MODEL ARE EXOGENOUS.
PRICES AND
* MINIMUM AND MAXIMUM TRADE LEVELS MAY BE SET IN THE FOLLOWING TABLES FOR REGIONAL AND
NONREGIONAL
* INPUTS AND PRODUCTS.

PARAMETER  CXMR(JR,JIP,JXMP)  REGIONAL IMPORT AND EXPORT PARAMETERS;

PARAMETER  CXMN(JIP,JXMP)  NONREGIONAL IMPORT AND EXPORT PARAMETERS;

$STITLE INTER-REGIONAL TRANSPORTATION PARAMETER DECLARATIONS AND ASSIGNMENTS
TABLE  DT(JRS,JRD)  DISTANCE IN MILES BETWEEN REGIONS
$OFFLISTING;
*$INCLUDE C:\0 Teaching\ApEc 8202\Exams\DT.INC
$INCLUDE G:\My Drive\MalawiAgSectorModel\GAMS Model\DT.INC
;
$ONLISTING;


DT(JRS,JRS) = 9999999.0;


PARAMETER CTIP(JIP) UNIT TRAN COST PER KM PER PRODUCT;
        CTIP(JIP)=0.018*SCENARIOS("BASELINE");
*        CTIP(JIP)=0.0000000001*SCENARIOS("ZEROTRANSPORTCOST");
*        CTIP(JIP)=0.009*SCENARIOS("HALFTRANSPORTCOST");
*        CTIP(JIP)=0.036*SCENARIOS("DOUBLETRANSPORTCOST");

PARAMETER  CTR(JRS,JRD,JIP)  UNIT INTER-REGIONAL TRANS COST BY SOURCE DESTINATION &
ITEM;

           CTR(JRS,JRD,JIP) = 999999.9;

           CTR(JRS,JRD,JTIP)$JRSDIP(JRS,JRD,JTIP) = DT(JRS,JRD)*CTIP(JTIP);

$STITLE PRODUCTION ACTIVITY TECHNICAL COEFFICIENTS AND SUPPORTING BUDGET DATA
* INPUT/OUTPUT COEFFICIENTS AND OTHER CONSTRAINT COEFFICIENTS FOR PRODUCTION
ACTIVITIES MAY BE ENTERED
* DIRECTLY, OR THEY MAY BE COMPUTED USING COEFFICIENTS FROM UNIT BUDGETS.  FOR CROP
AND LIVESTOCK
* PRODUCTION ACTIVITIES, THE CONSTRAINT COEFFICIENTS MAY BE COMPUTED WITH COEFFICIENTS
AND MULTIPLIERS
* FOR RELATED CROP AND LIVESTOCK BUDGETS.
*
* THE INPUT/OUTPUT COEFFICIENTS SHOULD BE INTERPRETED AS THE NET USE OF THE ITEM PER
UNIT OF THE
* ACTIVITY.  THUS AN INPUT REQUIREMENT WOULD BE INDICATED BY A POSITIVE COEFFICIENT.
A NEGATIVE INPUT
* COEFFICIENT IMPLIES THAT THE ACTIVITY IS A NET SUPPLIER OF THE INPUT.  A PRODUCT
WOULD TYPICALLY HAVE
* A NEGATIVE COEFFICIENT INDICATING NET SUPPLY.  ENDOGENOUS DEMAND FOR A PRODUCT WOULD
BE IMPLIED BY A
* PRODUCTION ACTIVITY WITH A POSITIVE INPUT COEFFICIENT.

TABLE  AC(JIT,JBC,JIP)  PRODUCT AND INPUT COEFFICIENTS FOR CROP BUDGETS
$OFFLISTING;
*$INCLUDE C:\0 Teaching\ApEc 8202\Exams\AC.INC
$INCLUDE G:\My Drive\MalawiAgSectorModel\GAMS Model\AC.INC
```

```
;
$ONLISTING;

* ASSIGN INPUT REQUIREMENTS FOR CROP LAND INPUTS

AC(JIT,"B-MAIZE","CL-MAIZE")$JBCT("B-MAIZE",JIT) = 1.0;
AC(JIT,"B-MAIZE","CL-TOTAL")$JBCT("B-MAIZE",JIT) = 1.0;

AC(JIT,"B-RICE","CL-RICE")$JBCT("B-RICE",JIT) = 1.0;
AC(JIT,"B-RICE","CL-TOTAL")$JBCT("B-RICE",JIT) = 1.0;

AC(JIT,"B-CASSAVA","CL-CASSAVA")$JBCT("B-CASSAVA",JIT) = 1.0;
AC(JIT,"B-CASSAVA","CL-TOTAL")$JBCT("B-CASSAVA",JIT) = 1.0;

AC(JIT,"B-POTATOES","CL-POTATOES")$JBCT("B-POTATOES",JIT) = 1.0;
AC(JIT,"B-POTATOES","CL-TOTAL")$JBCT("B-POTATOES",JIT) = 1.0;

AC(JIT,"B-BEANS","CL-BEANS")$JBCT("B-BEANS",JIT) = 1.0;
AC(JIT,"B-BEANS","CL-TOTAL")$JBCT("B-BEANS",JIT) = 1.0;

AC(JIT,"B-GROUNDNUTS","CL-GROUNDNUTS")$JBCT("B-GROUNDNUTS",JIT) = 1.0;
AC(JIT,"B-GROUNDNUTS","CL-TOTAL")$JBCT("B-GROUNDNUTS",JIT) = 1.0;

AC(JIT,"B-MIX-00","CONVEX")$JBCT("B-MIX-00",JIT) = 1.0;
AC(JIT,"B-MIX-01","CONVEX")$JBCT("B-MIX-01",JIT) = 1.0;
AC(JIT,"B-MIX-02","CONVEX")$JBCT("B-MIX-02",JIT) = 1.0;
AC(JIT,"B-MIX-03","CONVEX")$JBCT("B-MIX-03",JIT) = 1.0;
AC(JIT,"B-MIX-04","CONVEX")$JBCT("B-MIX-04",JIT) = 1.0;
AC(JIT,"B-MIX-05","CONVEX")$JBCT("B-MIX-05",JIT) = 1.0;
AC(JIT,"B-MIX-06","CONVEX")$JBCT("B-MIX-06",JIT) = 1.0;
AC(JIT,"B-MIX-07","CONVEX")$JBCT("B-MIX-07",JIT) = 1.0;
AC(JIT,"B-MIX-08","CONVEX")$JBCT("B-MIX-08",JIT) = 1.0;
AC(JIT,"B-MIX-09","CONVEX")$JBCT("B-MIX-09",JIT) = 1.0;

PARAMETER  AOC(JIT,JBC,JOC)  OTHER CONSTRAINT COEFFICIENTS FOR CROP BUDGETS;


PARAMETER  ACM(JR,JBC,JXC)  IO COEF MULTIPLIER CROP BUDGET TO PRODUCTION ACTIVITY;

           ACM(JR,"B-MAIZE","X-MAIZE")$(JBR(JR,"B-MAIZE") AND JRX(JR,"X-MAIZE")) = 1;
           ACM(JR,"B-RICE","X-RICE")$(JBR(JR,"B-RICE") AND JRX(JR,"X-RICE")) = 1;
           ACM(JR,"B-CASSAVA","X-CASSAVA")$(JBR(JR,"B-CASSAVA") AND JRX(JR,"X-
CASSAVA")) = 1;
           ACM(JR,"B-POTATOES","X-POTATOES")$(JBR(JR,"B-POTATOES") AND JRX(JR,"X-
POTATOES")) = 1;
           ACM(JR,"B-BEANS","X-BEANS")$(JBR(JR,"B-BEANS") AND JRX(JR,"X-BEANS")) = 1;
           ACM(JR,"B-GROUNDNUTS","X-GROUNDNUTS")$(JBR(JR,"B-GROUNDNUTS") AND
JRX(JR,"X-GROUNDNUTS"))= 1;

           ACM(JR,"B-MIX-00","X-MIX-00")$JRX(JR,"X-MIX-00") = 1;
           ACM(JR,"B-MIX-01","X-MIX-01")$JRX(JR,"X-MIX-01") = 1;
           ACM(JR,"B-MIX-02","X-MIX-02")$JRX(JR,"X-MIX-02") = 1;
           ACM(JR,"B-MIX-03","X-MIX-03")$JRX(JR,"X-MIX-03") = 1;
           ACM(JR,"B-MIX-04","X-MIX-04")$JRX(JR,"X-MIX-04") = 1;
           ACM(JR,"B-MIX-05","X-MIX-05")$JRX(JR,"X-MIX-05") = 1;
           ACM(JR,"B-MIX-06","X-MIX-06")$JRX(JR,"X-MIX-06") = 1;
           ACM(JR,"B-MIX-07","X-MIX-07")$JRX(JR,"X-MIX-07") = 1;
           ACM(JR,"B-MIX-08","X-MIX-08")$JRX(JR,"X-MIX-08") = 1;
           ACM(JR,"B-MIX-09","X-MIX-09")$JRX(JR,"X-MIX-09") = 1;

JBCR(JR,JXC,JBC)$(JXR(JXC) AND (ACM(JR,JBC,JXC) NE 0)) = YES;

JRXT(JR,JXC,JIT)$(JXR(JXC) AND JRT(JR,JIT)) = YES;
```

```
$EJECT;

PARAMETER  AL(JR,JBL,JIP)  PRODUCT AND INPUT COEF FOR LIVESTOCK BUDGETS;

PARAMETER  AOL(JR,JB,JOC)  OTHER CONSTRAINT COEF FOR LIVESTOCK BUDGETS;

PARAMETER  ALM(JR,JBL,JXL)  IO COEF MULTIPLIER LIVESTOCK BUDGET TO PROD ACTIVITY;

PARAMETER  ARP(JR,JXP,JIP)  PRODUCT AND INPUT COEF FOR REGIONAL PROCESSING ACTIVITIES;

PARAMETER  AORP(JR,JXP,JOC)  OTHER CONSTRAINT COEF FOR REGIONAL PROCESSING ACTIVITIES;

PARAMETER  AR(JR,JXR,JIP)  PRODUCT AND INPUT COEF FOR REGIONAL PROD ACTIVITIES;

* CALCULATE INPUT/OUTPUT COEFFICIENTS FOR LIVESTOCK PRODUCTION ACTIVITIES USING BUDGET
COEFFICIENTS AND
* CORRESPONDING BUDGET MULTIPLIERS, ADJUSTING FOR UNIT CONVERSIONS

*          AR(JR,JXL,JIP)$JRX(JR,JXL) = SUM(JBL$(JBR(JR,JBL) AND (CNV1(JBL,JXL)>0)),
*
CNV2(JIP)*ALM(JR,JBL,JXR)*AL(JR,JBL,JIP)/CNV1(JBL,JXR));

*         AR(JR,JXP,JIP)$JRX(JR,JXP) = ARP(JR,JXP,JIP);

PARAMETER  ARC(JR,JXC,JIT,JIP)  PRODUCT AND INPUT COEF FOR REG CROP PROD ACTIVITIES;

PARAMETER  AOR(JR,JXR,JOC)  OTHER CONSTRAINT COEF FOR REGIONAL PROD ACTIVITIES;

PARAMETER  AORC(JR,JXC,JIT,JOC)  OTHER CONSTRAINT COEF FOR REG CROP PROD ACTIVITIES;

PARAMETER  BOR(JR,JOC)  RHS FOR OTHER REGIONAL CONSTRAINTS;

$EJECT


* FOR PRODUCTION ACTIVITIES MAPPED TO CROP AND LIVESTOCK BUDGETS, CALCULATE THE INPUT
AND OUTPUT
* COEFFICIENTS.  IF A COEFFICIENT VALUE HAS BEEN ASSIGNED TO THE ACTIVITY AND THAT
ACTIVITY IS MAPPED TO
* A CROP OR LIVESTOCK BUDGET, THE COEFFICIENT IS REASSIGNED TO ITS CURRENT VALUE PLUS
THE BUDGET VALUE
* TIMES THE CORRESPONDING MULTIPLIER.  COEFFICIENTS FOR OTHER CONSTRAINTS ARE COMPUTED
THE SAME WAY.


PARAMETER  ARC(JR,JXC,JIT,JIP) PRODUCT AND INPUT COEF FOR REG CROP PROD ACTIVITIES;

ARC(JR,JXC,JIT,JIP)$JRXT(JR,JXC,JIT)
              = SUM(JBC$(JBCT(JBC,JIT)),
                    ACM(JR,JBC,JXC)*AC(JIT,JBC,JIP));

* THESE WERE ADDED ON 3RD APRIL 2019 BECAUSE THE MODEL WAS NOT PRODUCING
*  WITHIN THE CROP MIX CONVEX COMBINATION

ARC(JR,"X-MAIZE",JIT,JI)$JRXT(JR,"X-MAIZE",JIT) = ARC(JR,"X-MAIZE",JIT,JI);
ARC(JR,"X-MAIZE",JIT,JIC) = 0;
ARC(JR,"X-MAIZE",JIT,"CL-MAIZE")$JRXT(JR,"X-MAIZE",JIT) = ARC(JR,"X-MAIZE",JIT,"CL-
MAIZE");
ARC(JR,"X-MAIZE",JIT,"CL-TOTAL")$JRXT(JR,"X-MAIZE",JIT) = ARC(JR,"X-MAIZE",JIT,"CL-
TOTAL");
*ARC(JR,"X-CORN2",JIT,JP)$JRXT(JR,"X-CORN2",JIT) = 0.95*ARC(JR,"X-CORN",JIT,JP);
```

```
ARC(JR,"X-RICE",JIT,JI)$JRXT(JR,"X-RICE",JIT) = ARC(JR,"X-RICE",JIT,JI);
ARC(JR,"X-RICE",JIT,JIC) = 0;
ARC(JR,"X-RICE",JIT,"CL-RICE")$JRXT(JR,"X-RICE",JIT) = ARC(JR,"X-RICE",JIT,"CL-RICE");
ARC(JR,"X-RICE",JIT,"CL-TOTAL")$JRXT(JR,"X-RICE",JIT) = ARC(JR,"X-RICE",JIT,"CL-
TOTAL");

ARC(JR,"X-CASSAVA",JIT,JI)$JRXT(JR,"X-CASSAVA",JIT) = ARC(JR,"X-CASSAVA",JIT,JI);
ARC(JR,"X-CASSAVA",JIT,JIC) = 0;
ARC(JR,"X-CASSAVA",JIT,"CL-CASSAVA")$JRXT(JR,"X-CASSAVA",JIT) = ARC(JR,"X-
CASSAVA",JIT,"CL-CASSAVA");
ARC(JR,"X-CASSAVA",JIT,"CL-TOTAL")$JRXT(JR,"X-CASSAVA",JIT) = ARC(JR,"X-
CASSAVA",JIT,"CL-TOTAL");

ARC(JR,"X-POTATOES",JIT,JI)$JRXT(JR,"X-POTATOES",JIT) = ARC(JR,"X-POTATOES",JIT,JI);
ARC(JR,"X-POTATOES",JIT,JIC) = 0;
ARC(JR,"X-POTATOES",JIT,"CL-POTATOES")$JRXT(JR,"X-POTATOES",JIT) = ARC(JR,"X-
POTATOES",JIT,"CL-POTATOES");
ARC(JR,"X-POTATOES",JIT,"CL-TOTAL")$JRXT(JR,"X-POTATOES",JIT) = ARC(JR,"X-
POTATOES",JIT,"CL-TOTAL");

ARC(JR,"X-BEANS",JIT,JI)$JRXT(JR,"X-BEANS",JIT) = ARC(JR,"X-BEANS",JIT,JI);
ARC(JR,"X-BEANS",JIT,JIC) = 0;
ARC(JR,"X-BEANS",JIT,"CL-BEANS")$JRXT(JR,"X-BEANS",JIT) = ARC(JR,"X-BEANS",JIT,"CL-
BEANS");
ARC(JR,"X-BEANS",JIT,"CL-TOTAL")$JRXT(JR,"X-BEANS",JIT) = ARC(JR,"X-BEANS",JIT,"CL-
TOTAL");

ARC(JR,"X-GROUNDNUTS",JIT,JI)$JRXT(JR,"X-GROUNDNUTS",JIT) = ARC(JR,"X-
GROUNDNUTS",JIT,JI);
ARC(JR,"X-GROUNDNUTS",JIT,JIC) = 0;
ARC(JR,"X-GROUNDNUTS",JIT,"CL-GROUNDNUTS")$JRXT(JR,"X-GROUNDNUTS",JIT) = ARC(JR,"X-
GROUNDNUTS",JIT,"CL-GROUNDNUTS");
ARC(JR,"X-GROUNDNUTS",JIT,"CL-TOTAL")$JRXT(JR,"X-GROUNDNUTS",JIT) = ARC(JR,"X-
GROUNDNUTS",JIT,"CL-TOTAL");

PARAMETER  AN(JX,JIP)  PRODUCT AND INPUT COEFFICIENTS FOR NONREGIONAL PROD ACTIVITIES;

PARAMETER  AON(JX,JOC)  OTHER CONSTRAINT COEFFICIENTS FOR NONREGIONAL PROD ACTIVITIES;

PARAMETER  BON(JOC)  RHS FOR OTHER NONREGIONAL CONSTRAINTS;

*AreaData (JR,JXC,JBC)

$STITLE  DATA CHECKS AND DISPLAY
* THE FOLLOWING SETS ARE CONSTRUCTED TO SEE IF INPUTS, PRODUCTS, BUDGETS, OR
CONSTRAINTS AND ACTIVITIES
* LACK INDICATED UNITS OR IF DUPLICATE UNITS ARE SPECIFIED.  INPUT AND PRODUCT ITEMS
DO NOT NEED BUDGET
* UNITS IF THEY DO NO APPEAR IN A CROP OR LIVESTOCK BUDGET

SET  ICK01(JBC)  BUDGETS WITH NO UNITS SPECIFIED;
SET  ICK02(JBC)  BUDGETS WITH DUPLICATE UNITS SPECIFIED;
SET  ICK03(JXC)  ACTIVITIES WITH NO UNITS SPECIFIED;
SET  ICK04(JXC)  ACTIVITIES WITH DUPLICATE UNITS SPECIFIED;
SET  ICK05(JBC,JXC)  BUDGET - ACTIVITY COMBOS W ZERO CONVERSION MULT;

SET  ICK06(JIP)  INPUTS AND PRODUCTS USED IN CROP OR LIVESTOCK BUDGETS;
SET  ICK07(JIP)  INPUTS AND PRODUCTS W NO BUDGET UNITS SPECIFIED;
SET  ICK08(JIP)  INPUTS AND PRODUCTS W DUPLICATE BUDGET UNITS SPECIFIED;
SET  ICK09(JIP)  INPUTS AND PRODUCTS W NO MODEL UNITS SPECIFIED;
SET  ICK10(JIP)  INPUTS AND PRODUCTS W DUPLICATE MODEL UNITS SPECIFIED;
SET  ICK11(JIP)  INPUTS AND PRODUCTS W ZERO CONVERSION MULT;
```

```
PARAMETERS  KNT1(JBC,JXC)  UNIT CONVERSION COMB0S FOR BUDGET AND ACT - MUST = 1,
            KNT2(JIP)  UNIT CONVERSION COMBOS FOR INPUT OR PRODUCT - MUST = 1;

OPTION  CNV1:6:1:1;     DISPLAY  CNV1;
OPTION  CNV2:6:0:1;     DISPLAY  CNV2;

ICK01(JBC) = NO;      ICK01(JBC)$(SUM(JU$JUB(JBC,JU),1)<1) = YES;     DISPLAY ICK01;
ICK02(JBC) = NO;      ICK02(JBC)$(SUM(JU$JUB(JBC,JU),1)>1) = YES;     DISPLAY ICK02;
ICK03(JXC) = NO;      ICK03(JXC)$(SUM(JU$JUX(JXC,JU),1)<1) = YES;     DISPLAY ICK03;
ICK04(JXC) = NO;      ICK04(JXC)$(SUM(JU$JUX(JXC,JU),1)>1) = YES;     DISPLAY ICK04;

ICK05(JBC,JXC) = NO;
ICK05(JBC,JXC)
  $(CNV1(JBC,JXC)<=0 AND (SUM(JR$(JBCR(JR,JXC,JBC) AND ACM(JR,JBC,JXC)>0),1)>0)) =
YES;
DISPLAY ICK05;

ICK06(JIP) = NO;
ICK06(JIP)$(SUM(JIT,SUM(JBC,ABS(AC(JIT,JBC,JIP)))) > 0) = YES;
*ICK06(JIP)$(SUM(JR,SUM(JBL,ABS(AL(JR,JBL,JIP)))) > 0) = YES;

ICK07(JIP) = NO;
ICK07(JIP)$((SUM(JU$JUIPB(JIP,JU),1)<1) AND ICK06(JIP)) = YES;  DISPLAY ICK07;

ICK08(JIP) = NO;      ICK08(JIP)$(SUM(JU$JUIPB(JIP,JU),1)>1) = YES;   DISPLAY ICK08;
ICK09(JIP) = NO;      ICK09(JIP)$(SUM(JU$JUIPX(JIP,JU),1)<1) = YES;   DISPLAY ICK09;
ICK10(JIP) = NO;      ICK10(JIP)$(SUM(JU$JUIPX(JIP,JU),1)>1) = YES;   DISPLAY ICK10;
ICK11(JIP) = NO;      ICK11(JIP)$(CNV2(JIP)<=0) = YES;                DISPLAY ICK11;

*PARAMETER ARCB(JR,JXC,JIT,JIP);

DISPLAY ACM;
DISPLAY AC;
DISPLAY ARC;

SET  IDTR(JR)  DISPLAY DATA FOR THESE REGIONS  /CHI,KAR/;

SET  IDTIC(JIT)  DISPLAY DATA FOR THESE CROP LAND TYPES  /CLT-CHI,CLT-KAR/;


PARAMETER  DT1(JR,JX,JB)  BUDGET MULT BY PROD ACT & REG;
           DT1(JR,JXC,JBC)$IDTR(JR) = ACM(JR,JBC,JXC);
           OPTION  DT1:4:1:1;      DISPLAY  DT1;

PARAMETER  DT2(JIT,JIP,JBC)  IP COEF FOR CROP BUDGETS;
           DT2(JIT,JIP,JBC)$IDTIC(JIT) = AC(JIT,JBC,JIP);   OPTION  DT2:5:1:1;
DISPLAY  DT2;

PARAMETER  DT3(JR,JIT,JIP,JXC)  PROD & INPUT COEF TABLE;
           DT3(JR,JIT,JIP,JXC)$(IDTR(JR) AND IDTIC(JIT)) = ARC(JR,JXC,JIT,JIP);
* OPTION  DT3:5:1:1;
           DISPLAY  DT3;

IDTR(JR) = NO;  IDTR("CHI") = YES;   IDTR("KAR") = YES;


$STITLE  DECLARE AND DEFINE VARIABLES, EQUATIONS AND THE SECTOR MODEL
VARIABLES
*--------------------------------------------------------------------------------
------------------
* VARIABLE........
DESCRIPTION...................................................................
```

```
*-------------------------------------------------------------------------------------
------------------
  ZOBJ              CONSUMER PLUS PRODUCER SURPLUS
  XR(JR,JXR)        REGIONAL PRODUCTION ACTIVITIES
  XRC(JR,JXC,JIT)   REGIONAL CROP PRODUCTION ACTIVITIES
  ZR(JR,JIR)        REGIONAL INPUT SUPPLY
  ZC(JR,JIT,JIC)    CROP LAND INPUT SUPPLY
*  LAMBDA2(JR,YEAR)    WEIGHT VARIABLES FOR HISTORICAL CROP MIXES
* CROPLAND2(JR,JIC) CROP LAND USED BY CROP
  YR(JR,JPR)        REGIONAL PRODUCT DEMAND
  TR(JRS,JRD,JIP)   INTER-REGIONAL TRANSPORTATION ACTIVITIES
  TMR(JR,JIP)       REGIONAL IMPORT ACTIVITIES
  TXR(JR,JIP)       REGIONAL EXPORT ACTIVITIES
  XN(JX)            NON-REGIONAL PRODUCTION ACTIVITIES
  ZN(JIN)           NON-REGIONAL INPUT SUPPLY
  YN(JPN)           NON-REGIONAL PRODUCT DEMAND
  TMN(JIP)          NON-REGIONAL IMPORT ACTIVITIES
  TXN(JIP)          NON-REGIONAL EXPORT ACTIVITIES;
*-------------------------------------------------------------------------------------
------------------

POSITIVE VARIABLES  XR, XRC, ZR, ZC, YR, TR, TMR, TXR, XN, ZN, YN, TMN, TXN;
*LAMBDA2, CROPLAND2;

*
* SET THE LOWER AND UPPER BOUNDS ON INPUT SUPPLY AND PRODUCT DEMAND VARIABLES TO THE
ASSOCIATED QMIN
* AND QMAX VALUES, RESPECTIVELY


ZR.LO(JR,JIR) $JRIR(JR,JIR) = CSR(JR,JIR,"QMIN");
ZR.UP(JR,JIR) $JRIR(JR,JIR) = CSR(JR,JIR,"QMAX");

ZC.LO(JR,JIT,JIC) $(JRT(JR,JIT) AND JRIC(JR,JIC)) = CSC(JR,JIT,JIC,"QMIN");
ZC.UP(JR,JIT,JIC) $(JRT(JR,JIT) AND JRIC(JR,JIC)) = CSC(JR,JIT,JIC,"QMAX");

YR.LO(JR,JPR) $JRPR(JR,JPR) = CDR(JR,JPR,"QMIN");
YR.UP(JR,JPR) $JRPR(JR,JPR) = CDR(JR,JPR,"QBAR");

$EJECT
EQUATIONS
*-------------------------------------------------------------------------------------
------------------
* EQUATION..............
DESCRIPTION..........................................................
*-------------------------------------------------------------------------------------
------------------
 OBJ                       OBJECTIVE FUNCTION SECTOR WELFARE 1000'S
 INPUTREG(JR,JIR)          REGIONAL INPUT CONSTRAINTS
 INPUTREGCL(JR,JIT,JIC)    CROP LAND INPUT CONSTRAINTS
 PRODREG(JR,JPR)           REGIONAL PRODUCT CONSTRAINTS;
* CONVEX2(JR)               CONVEXITY REQUIREMENT
* HISTAREA2(JR,JIC)         TOTAL LAND CONSTRAINT CROPS RESTRICTED TO HISTORICAL MIXES;
*INPUTNON(JIN)             NONREGIONAL INPUT CONSTRAINTS
* PRODNON(JPN)              NONREGIONAL PRODUCT CONSTRAINTS;
*-------------------------------------------------------------------------------------
------------------


OBJ..

 ZOBJ =E=
```

```
 .001*( SUM(JR, - SUM(JIR $JRIR(JR,JIR), (CSR(JR,JIR,"INTERCEPT")*ZR(JR,JIR)
                                     + 0.5*CSR(JR,JIR,"SLOPE")*(ZR(JR,JIR)**2)))

                - SUM(JIT $JRT(JR,JIT), SUM(JIC $JRIC(JR,JIC),

(CSC(JR,JIT,JIC,"INTERCEPT")*ZC(JR,JIT,JIC)
                                                +
0.5*CSC(JR,JIT,JIC,"SLOPE")*(ZC(JR,JIT,JIC)**2))))

                + SUM(JPR $JRPR(JR,JPR), (CDR(JR,JPR,"INTERCEPT")*YR(JR,JPR)
                                 + 0.5*CDR(JR,JPR,"SLOPE")*(YR(JR,JPR)**2)))

                - SUM(JRD, SUM(JIP $JRSDIP(JR,JRD,JIP),
CTR(JR,JRD,JIP)*TR(JR,JRD,JIP)))));

$ONTEXT
                - SUM(JIP $JIPRM(JR,JIP), CXMR(JR,JIP,"M-PRICE")*TMR(JR,JIP))

                + SUM(JIP $JIPRX(JR,JIP), CXMR(JR,JIP,"X-PRICE")*TXR(JR,JIP)))

        - SUM(JIN, (CSN(JIN,"INTERCEPT")*ZN(JIN) + 0.5*CSN(JIN,"SLOPE")*(ZN(JIN)**2)))

        + SUM(JPN, (CDN(JPN,"INTERCEPT")*YN(JPN) + 0.5*CDN(JPN,"SLOPE")*(YN(JPN)**2)))

        - SUM(JIP $JIPNM(JIP), CXMN(JIP,"M-PRICE")*TMN(JIP))

        + SUM(JIP $JIPNX(JIP), CXMN(JIP,"X-PRICE")*TXN(JIP)));

$OFFTEXT

INPUTREG(JR,JIR) $JRIR(JR,JIR)..

*SUM(JXR $(JRX(JR,JXR) AND (NOT JXC(JXR))), AR(JR,JXR,JIR)*XR(JR,JXR))

 SUM(JXC, SUM(JIT $JRXT(JR,JXC,JIT), ARC(JR,JXC,JIT,JIR)*XRC(JR,JXC,JIT)))

 - ZR(JR,JIR)

 + SUM(JRD $JRSDIP(JR,JRD,JIR), TR(JR,JRD,JIR))

 - SUM(JRS $JRSDIP(JRS,JR,JIR), TR(JRS,JR,JIR))=E= 0;

* - TMR(JR,JIR) $(JR,JIR) + TXR(JR,JIR) $JIPRX(JR,JIR) =E= 0;

*HISTAREA2(JR,JIC)..
*   SUM(YR2, LAMBDA2(JR,YEAR)*AREADATA(JR,JIC,YEAR))-CROPLAND2(JR,JIC)=L=0;

INPUTREGCL(JR,JIT,JIC) $(JRT(JR,JIT) AND JRIC(JR,JIC))..

  SUM(JXC $JRXT(JR,JXC,JIT), ARC(JR,JXC,JIT,JIC)*XRC(JR,JXC,JIT)) - ZC(JR,JIT,JIC) =L=
0;


PRODREG(JR,JPR) $JRPR(JR,JPR)..

  SUM(JXC, SUM(JIT $JRXT(JR,JXC,JIT), ARC(JR,JXC,JIT,JPR)*XRC(JR,JXC,JIT)))

 + YR(JR,JPR)

 + SUM(JRD $JRSDIP(JR,JRD,JPR), TR(JR,JRD,JPR))

 - SUM(JRS $JRSDIP(JRS,JR,JPR), TR(JRS,JR,JPR))=E= 0;
```

```
MODEL MAGS /ALL/;

*OPTION NLP = CONOPT
*OPTION NLP = CONOPT2
*OPTION NLP=MINOS5

SOLVE MAGS USING NLP MAXIMIZING ZOBJ;

$STITLE  GENERATE AND DISPLAY SOLUTION REPORTS
SET  JTH1  RESULTS TABLE HEADERS 1  /LOWER, LEVEL, UPPER, MARGINAL/;

SET  JTH2  RESULTS TABLE HEADERS 2  /AREA, PROD, YIELD, USE, QUANTITY, SHIP-IN, SHIP-
OUT, DEMAND,
     SUPPLY, EXPORT, IMPORT, LAMBDA, PRICE, VALUE/;

PARAMETER  RT01  CONSUMER PLUS PRODUCER SURPLUS $1000'S;
 RT01 = ZOBJ.L;
 OPTION  RT01:0:0:1;    DISPLAY$PRNT("01") RT01;

*PARAMETER  RT02(JR,JXR,JU,JTH1)  REGIONAL PRODUCTION ACTIVITIES;
* RT02(JR,JXR,JU,"LOWER")$(JUX(JXR,JU) AND JRX(JR,JXR)AND(NOT JXC(JXR))) =
XR.LO(JR,JXR);
* RT02(JR,JXR,JU,"LEVEL")$(JUX(JXR,JU) AND JRX(JR,JXR)AND(NOT JXC(JXR))) =
XR.L(JR,JXR);
* RT02(JR,JXR,JU,"UPPER")$(JUX(JXR,JU) AND JRX(JR,JXR)AND(NOT JXC(JXR))) =
XR.UP(JR,JXR);
* RT02(JR,JXR,JU,"MARGINAL")$(JUX(JXR,JU) AND JRX(JR,JXR)AND(NOT JXC(JXR))) =
XR.M(JR,JXR);
* OPTION  RT02:4:2:1;    DISPLAY$PRNT("02") RT02;

PARAMETER  RT03(JR,JXC,JIT,JU,JTH1)  REGIONAL CROP PRODUCTION ACTIVITIES;
 RT03(JR,JXC,JIT,JU,"LOWER")$(JRXT(JR,JXC,JIT) AND JUX(JXC,JU)) = XRC.LO(JR,JXC,JIT);
 RT03(JR,JXC,JIT,JU,"LEVEL")$(JRXT(JR,JXC,JIT) AND JUX(JXC,JU)) = XRC.L(JR,JXC,JIT);
 RT03(JR,JXC,JIT,JU,"UPPER")$(JRXT(JR,JXC,JIT) AND JUX(JXC,JU)) = XRC.UP(JR,JXC,JIT);
 RT03(JR,JXC,JIT,JU,"MARGINAL")$(JRXT(JR,JXC,JIT) AND JUX(JXC,JU)) =
XRC.M(JR,JXC,JIT);
 OPTION  RT03:4:3:1;    DISPLAY$PRNT("03") RT03;

PARAMETER  RT04(JR,JIR,JU,JTH1)  REGIONAL INPUT SUPPLY ACTIVITIES;
 RT04(JR,JIR,JU,"LOWER")$(JRIR(JR,JIR) AND JUIPX(JIR,JU)) = ZR.LO(JR,JIR);
 RT04(JR,JIR,JU,"LEVEL")$(JRIR(JR,JIR) AND JUIPX(JIR,JU)) = ZR.L(JR,JIR);
 RT04(JR,JIR,JU,"UPPER")$(JRIR(JR,JIR) AND JUIPX(JIR,JU)) = ZR.UP(JR,JIR);
 RT04(JR,JIR,JU,"MARGINAL")$(JRIR(JR,JIR) AND JUIPX(JIR,JU)) = ZR.M(JR,JIR);
 OPTION  RT04:4:3:1;    DISPLAY$PRNT("04") RT04;

PARAMETER  RT05(JR,JIT,JIC,JU,JTH1)  CROP LAND INPUT SUPPLY ACTIVITIES;
 RT05(JR,JIT,JIC,JU,"LOWER")$(JRT(JR,JIT) AND JRIC(JR,JIC) AND JUIPX(JIC,JU)) =
ZC.LO(JR,JIT,JIC);
 RT05(JR,JIT,JIC,JU,"LEVEL")$(JRT(JR,JIT) AND JRIC(JR,JIC) AND JUIPX(JIC,JU)) =
ZC.L(JR,JIT,JIC);
 RT05(JR,JIT,JIC,JU,"UPPER")$(JRT(JR,JIT) AND JRIC(JR,JIC) AND JUIPX(JIC,JU)) =
ZC.UP(JR,JIT,JIC);
 RT05(JR,JIT,JIC,JU,"MARGINAL")$(JRT(JR,JIT) AND JRIC(JR,JIC) AND JUIPX(JIC,JU)) =
ZC.M(JR,JIT,JIC);
 OPTION  RT05:4:3:1;    DISPLAY$PRNT("05") RT05;

PARAMETER  RT06(JR,JPR,JU,JTH1)  REGIONAL PRODUCT DEMAND ACTIVITIES;
 RT06(JR,JPR,JU,"LOWER")$(JRPR(JR,JPR) AND JUIPX(JPR,JU)) = YR.LO(JR,JPR);
 RT06(JR,JPR,JU,"LEVEL")$(JRPR(JR,JPR) AND JUIPX(JPR,JU)) = YR.L(JR,JPR);
 RT06(JR,JPR,JU,"UPPER")$(JRPR(JR,JPR) AND JUIPX(JPR,JU)) = YR.UP(JR,JPR);
 RT06(JR,JPR,JU,"MARGINAL")$(JRPR(JR,JPR) AND JUIPX(JPR,JU)) = YR.M(JR,JPR);
 OPTION  RT06:4:3:1;    DISPLAY$PRNT("06") RT06;
```

```
PARAMETER  RT07(JRS,JRD,JIP,JU,JTH1)  INTER-REGIONAL TRANSPORTATION ACTIVITIES;
 RT07(JRS,JRD,JIP,JU,"LOWER")$(JRSDIP(JRS,JRD,JIP) AND JUIPX(JIP,JU)) =
TR.LO(JRS,JRD,JIP);
 RT07(JRS,JRD,JIP,JU,"LEVEL")$(JRSDIP(JRS,JRD,JIP) AND JUIPX(JIP,JU)) =
TR.L(JRS,JRD,JIP);
 RT07(JRS,JRD,JIP,JU,"UPPER")$(JRSDIP(JRS,JRD,JIP) AND JUIPX(JIP,JU)) =
TR.UP(JRS,JRD,JIP);
 RT07(JRS,JRD,JIP,JU,"MARGINAL")$(JRSDIP(JRS,JRD,JIP) AND JUIPX(JIP,JU)) =
TR.M(JRS,JRD,JIP);
 OPTION  RT07:4:3:1;    DISPLAY$PRNT("07") RT07;

PARAMETER  RT31(JR,JIR,JTH1)  REGIONAL INPUT CONSTRAINTS;
 RT31(JR,JIR,"LOWER")$JRIR(JR,JIR) = INPUTREG.LO(JR,JIR);
 RT31(JR,JIR,"LEVEL")$JRIR(JR,JIR) = INPUTREG.L(JR,JIR);
 RT31(JR,JIR,"UPPER")$JRIR(JR,JIR) = INPUTREG.UP(JR,JIR);
 RT31(JR,JIR,"MARGINAL")$JRIR(JR,JIR) = INPUTREG.M(JR,JIR);
 OPTION  RT31:4:2:1;    DISPLAY$PRNT("31") RT31;

PARAMETER  RT32(JR,JIT,JIC,JTH1)  CROP LAND INPUT CONSTRAINTS;
 RT32(JR,JIT,JIC,"LOWER")$(JRT(JR,JIT) AND JRIC(JR,JIC)) = INPUTREGCL.LO(JR,JIT,JIC);
 RT32(JR,JIT,JIC,"LEVEL")$(JRT(JR,JIT) AND JRIC(JR,JIC)) = INPUTREGCL.L(JR,JIT,JIC);
 RT32(JR,JIT,JIC,"UPPER")$(JRT(JR,JIT) AND JRIC(JR,JIC)) = INPUTREGCL.UP(JR,JIT,JIC);
 RT32(JR,JIT,JIC,"MARGINAL")$(JRT(JR,JIT) AND JRIC(JR,JIC)) =
INPUTREGCL.M(JR,JIT,JIC);
 OPTION  RT32:4:2:1;    DISPLAY$PRNT("32") RT32;

PARAMETER  RT33(JR,JPR,JTH1)  REGIONAL PRODUCT CONSTRAINTS;
 RT33(JR,JPR,"LOWER")$JRPR(JR,JPR) = PRODREG.LO(JR,JPR);
 RT33(JR,JPR,"LEVEL")$JRPR(JR,JPR) = PRODREG.L(JR,JPR);
 RT33(JR,JPR,"UPPER")$JRPR(JR,JPR) = PRODREG.UP(JR,JPR);
 RT33(JR,JPR,"MARGINAL")$JRPR(JR,JPR) = PRODREG.M(JR,JPR);
 OPTION  RT33:4:2:1;    DISPLAY$PRNT("33") RT33;

PARAMETER  RT41(JR,JIT,JXC)  CROP PROD ACTIVITY LEVELS BY REGION & LAND TYPE;
 RT41(JR,JIT,JXC) $JRXT(JR,JXC,JIT) = XRC.L(JR,JXC,JIT);
 OPTION  RT41:3:1:1;     DISPLAY$PRNT("41")  RT41;

PARAMETER  RT42(JXC,JR)  CROP PROD ACTIVITY LEVELS SUMMED BY REGION;
 RT42(JXC,JR) = SUM(JIT $JRXT(JR,JXC,JIT), XRC.L(JR,JXC,JIT));
 OPTION  RT42:3:1:1;     DISPLAY$PRNT("42")  RT42;

PARAMETER  RT43(JXC)  TOTAL CROP PROD ACTIVITY LEVELS;
 RT43(JXC) = SUM(JR, SUM(JIT $JRXT(JR,JXC,JIT), XRC.L(JR,JXC,JIT)));
 OPTION  RT43:3:0:1;     DISPLAY$PRNT("43")  RT43;

PARAMETER  RT71(JR,JIP,JU,JTH2)  REGIONAL INPUT & PRODUCT RESULTS BY REGION;
 RT71(JR,JIP,JU,JTH2) = 0.0;
 RT71(JR,JIR,JU,"PROD")$JUIPX(JIR,JU)
   = SUM(JXC, SUM(JIT $(JRXT(JR,JXC,JIT) AND (ARC(JR,JXC,JIT,JIR)<0)),
         - ARC(JR,JXC,JIT,JIR)*XRC.L(JR,JXC,JIT)));
 RT71(JR,JPR,JU,"PROD")$JUIPX(JPR,JU)
   = SUM(JXC, SUM(JIT $(JRXT(JR,JXC,JIT) AND (ARC(JR,JXC,JIT,JPR)<0)),
         - ARC(JR,JXC,JIT,JPR)*XRC.L(JR,JXC,JIT)));
 RT71(JR,JIR,JU,"USE")$JUIPX(JIR,JU)
   = SUM(JXC, SUM(JIT $(JRXT(JR,JXC,JIT) AND (ARC(JR,JXC,JIT,JIR)>0)),
         ARC(JR,JXC,JIT,JIR)*XRC.L(JR,JXC,JIT)));
 RT71(JR,JPR,JU,"USE")$JUIPX(JPR,JU)
   = SUM(JXC, SUM(JIT $(JRXT(JR,JXC,JIT) AND (ARC(JR,JXC,JIT,JPR)>0)),
         ARC(JR,JXC,JIT,JPR)*XRC.L(JR,JXC,JIT)));
 RT71(JR,JIP,JU,"SHIP-IN")$JUIPX(JIP,JU) = SUM(JRS$JRSDIP(JRS,JR,JIP),
TR.L(JRS,JR,JIP));
 RT71(JR,JIP,JU,"SHIP-OUT")$JUIPX(JIP,JU) = SUM(JRD$JRSDIP(JR,JRD,JIP),
TR.L(JR,JRD,JIP));
```

```
 RT71(JR,JPR,JU,"DEMAND")$(JRPR(JR,JPR) AND JUIPX(JPR,JU)) = YR.L(JR,JPR);
 RT71(JR,JIR,JU,"SUPPLY")$(JRIR(JR,JIR) AND JUIPX(JIR,JU)) = ZR.L(JR,JIR);

 RT71(JR,JPR,JU,"PRICE")$(JRPR(JR,JPR) AND JUIPX(JPR,JU)) = PRODREG.M(JR,JPR);
 RT71(JR,JIR,JU,"PRICE")$(JRIR(JR,JIR) AND JUIPX(JIR,JU)) = INPUTREG.M(JR,JIR);
 RT71(JR,JPR,JU,"LAMBDA")$(JRPR(JR,JPR) AND JUIPX(JPR,JU)) = YR.M(JR,JPR);
 RT71(JR,JIR,JU,"LAMBDA")$(JRIR(JR,JIR) AND JUIPX(JIR,JU)) = ZR.M(JR,JIR);
 OPTION  RT71:3:2:1;     DISPLAY$PRNT("71")  RT71;

PARAMETER  RT72(JIP,JR,JU,JTH2)  REGIONAL INPUT & PRODUCT RESULTS BY ITEM;
 RT72(JIP,JR,JU,JTH2) = RT71(JR,JIP,JU,JTH2);
 OPTION  RT72:3:2:1;     DISPLAY$PRNT("72")  RT72;

SET  JIP73(JIP)
/MAIZESEED,RICESEED,CASSAVASEED,POTATOSEED,BEANSSEED,GNUTSSEED,FERT1,FERT2,
                       PESTICIDES,TRANSPORT,PACKAGINGM,LABOR/;

PARAMETER  RT73(JIP,JR,JU,JTH2)  REGIONAL INPUT & PRODUCT RESULTS FOR SELECT ITEMS;
 RT73(JIP,JR,JU,JTH2)$JIP73(JIP) = RT71(JR,JIP,JU,JTH2);
 OPTION  RT73:3:2:1;     DISPLAY$PRNT("73")  RT73;

PARAMETER  RT81(JIP,JU,JRS,JRD)  INTER-REG INPUT & PRODUCT SHIPMENTS;
 RT81(JIP,JU,JRS,JRD) = RT07(JRS,JRD,JIP,JU,"LEVEL");
 OPTION  RT81:3:1:1;     DISPLAY$PRNT("81")  RT81;

SET  JPR91(JPR)  CROP PRODUCTS FOR RESULTS TABLE 91
/MAIZE,RICE,CASSAVA,POTATOES,BEANS,GROUNDNUTS/;

PARAMETER  RT91(JPR91,JR,JTH2)  CROP PRODUCT RESULTS BY REGION TABLE 91;
 RT91(JPR91,JR,"PROD") = SUM(JU$JUIPX(JPR91,JU),RT71(JR,JPR91,JU,"PROD"));
 RT91(JPR91,JR,"USE") = SUM(JU$JUIPX(JPR91,JU),RT71(JR,JPR91,JU,"USE"));
 RT91(JPR91,JR,"SHIP-IN") = SUM(JU$JUIPX(JPR91,JU),RT71(JR,JPR91,JU,"SHIP-IN"));
 RT91(JPR91,JR,"DEMAND") = SUM(JU$JUIPX(JPR91,JU),RT71(JR,JPR91,JU,"DEMAND"));
 RT91(JPR91,JR,"EXPORT") = SUM(JU$JUIPX(JPR91,JU),RT71(JR,JPR91,JU,"EXPORT"));
 RT91(JPR91,JR,"IMPORT") = SUM(JU$JUIPX(JPR91,JU),RT71(JR,JPR91,JU,"IMPORT"));
 RT91(JPR91,JR,"PRICE") = SUM(JU$JUIPX(JPR91,JU),RT71(JR,JPR91,JU,"PRICE"));
 RT91(JPR91,JR,"AREA") = SUM(JU$JUIPX(JPR91,JU),
      SUM(JXC, SUM(JIT $(JRXT(JR,JXC,JIT) AND
(ARC(JR,JXC,JIT,JPR91)<0)),XRC.L(JR,JXC,JIT))));
 RT91(JPR91,JR,"YIELD")$(RT91(JPR91,JR,"AREA")>0) =
RT91(JPR91,JR,"PROD")/RT91(JPR91,JR,"AREA");
 OPTION  RT91:3:1:1;     DISPLAY$PRNT("91")  RT91;

PARAMETER  RT92(JPR91,JTH2)  CROP RESULTS TOTAL TABLE 92;
 RT92(JPR91,"AREA") = SUM(JR,RT91(JPR91,JR,"AREA"));
 RT92(JPR91,"PROD") = SUM(JR,RT91(JPR91,JR,"PROD"));
 RT92(JPR91,"USE") = SUM(JR,RT91(JPR91,JR,"USE"));
 RT92(JPR91,"DEMAND") = SUM(JR,RT91(JPR91,JR,"DEMAND"));
 RT92(JPR91,"EXPORT") = SUM(JR,RT91(JPR91,JR,"EXPORT"));
 RT92(JPR91,"IMPORT") = SUM(JR,RT91(JPR91,JR,"IMPORT"));
  RT92(JPR91,"SHIP-IN") = SUM(JR,RT91(JPR91,JR,"SHIP-IN"));
 RT92(JPR91,"YIELD")$(RT92(JPR91,"AREA")>0) = RT92(JPR91,"PROD")/RT92(JPR91,"AREA");


 OPTION  RT92:3:1:1;     DISPLAY$PRNT("92")  RT92;

*=== Export to Excel using GDX utilities

*=== First unload to GDX file (occurs during execution phase)
*execute_unload "G:\My Drive\MalawiAgSectorModel\GAMS Model\results.gdx"
execute_unload "G:\My Drive\MalawiAgSectorModel\GAMS Model\resultsBaseline.gdx"
*execute_unload "G:\My Drive\MalawiAgSectorModel\GAMS Model\resultsZeroTC.gdx"
*execute_unload "G:\My Drive\MalawiAgSectorModel\GAMS Model\resultsHalfTC.gdx"
```

```
*execute_unload "G:\My Drive\MalawiAgSectorModel\GAMS Model\resultsDoubleTC.gdx"
*=== Now write to variable levels to Excel file from GDX
*=== Since we do not specify a sheet, data is placed in first sheet
execute 'gdxxrw.exe resultsBaseline.gdx'
*execute 'gdxxrw.exe resultsZeroTC.gdx'
*execute 'gdxxrw.exe resultsHalfTC.gdx'
*execute 'gdxxrw.exe resultsDoubleTC.gdx'
* execute 'gdxxrw.exe results.gdx'
*=== Write marginals to a different sheet with a specific range
```