PROFITABILITY OF FERTILIZER USE ON MAIZE BY SMALL-SCALE FARMING HOUSEHOLDS IN ZAMBIA

Zhiying Xu, Thom S. Jayne, J. Roy Black, and Jones Govereh*

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island, July24-27, 2005

*Xu is a Ph.D. Candidate (xuzhiyin@msu.edu), Jayne and Black are Professors, Department of Agricultural Economics, Michigan State University, and Govereh is a research fellow, Food Security Research Project-Zambia, Lusaka, Zambia.

Profitability of Fertilizer Use on Maize by Small-scale Farming Households in Zambia

Abstract

Multi-year nationwide survey data is used to estimate maize yield response functions and determine profitability of fertilizer use by small-scale farmers in Zambia. Most previous research on economics of fertilization used estimates of yield response to nutrients based on experimental or simulation data and seldom investigated region-specific and management-specific effects. In this paper we address the main issues arising from using large survey data and estimate maize yield response functions for different groups of households that have various management practices and soil conditions in two major agro-climatic zones. Profitability of fertilizer use is determined for each group in each zone and the results provide the following messages. First, households that obtained fertilizer on time and used animal draught power or mechanical power for land preparation are more likely to find fertilizer use profitable than other groups of households located in the same district. Second, farmers' proximity to the provincial centers has a significant impact on the profitability of fertilizer use. Greater distances and transport costs from provincial centers erode the profitability of fertilizer use. Third, high time preferences for money also reduce the profitability of fertilizer use. Thus, despite achieving relatively high physical crop response rates to fertilizer use in some areas, small farmers may find fertilizer use unprofitable until efforts are made to reduce transportation costs and implicit interest rates as well as to ensure more timely delivery of fertilizer.

Introduction

Farm productivity growth in widely understood to be a precondition for broad based economic development in most of the developing world (Johnston and Mellor, 1961; Tiffen 2003). Achieving this productivity growth is likely to involve, among many other things, substantially increased use of inorganic fertilizer. Currently, fertilizer use in Sub-Saharan Africa averages 9 kilograms (kgs) per hectare, the lowest of any developing region by far (FAO, 2004). While African policy makers and international donors recognize the urgency of raising fertilizer use by small farmers, for achieving both poverty alleviation and agricultural growth objectives, there is little consensus on the most appropriate policy and programmatic course of action.

Most efforts to raise fertilizer use in Sub-Saharan Africa over the past decade have focused on fertilizer subsidies and targeted credit programs, with the hopes that these programs could later be withdrawn once the profitability of fertilizer use has been made clear to newly adopting farmers, and once they have become sufficiently capitalized to be able to afford fertilizer with their own working capital. Relatively little emphasis has been given to improving the profitability of fertilizer use through understanding the most productive levels and combinations of nutrient input for various agro-ecological areas, management practices, and market conditions. For example, in Zambia government extension messages and distribution programs are based on one nationally-recommended application rate of 200kgs of Compound D and 200kgs of urea on each hectare of maize despite the substantial heterogeneity in farmer tillage techniques, seed types, agro-ecological and market conditions.

The objective of this paper is to estimate maize yield response to fertilizers under a range of small farm conditions, management practices, and market conditions in Zambia, and to identify the potential to increase fertilizer use through more profitable site-specific levels and combinations of nutrient application. Nationwide household surveys containing detailed production information is used to estimate maize yield response and the economics of fertilization for various soils, climates, management practices, and relative prices.

However, there are several important challenges to getting good estimates of the parameters of response functions using survey data. These challenges are identified in following section. We then provide a theoretical framework for our analysis, and a review of relevant literature to facilitate model specification, a section describing the survey data and methods, followed by the main findings and conclusions.

Challenges

First, yields and inputs are often measured with significant error. A related problem arises when households farm more than one field/plot but only household level information is available. Similarly, a household may use local seed on part of a plot but hybrid seed on another part, a logical extension of the challenge of multiple plots. Second, measurement of the underlying soils and climate is imperfect and, in any event, requires aggregation of categories to meaningful sized groups. In the Zambian case, approximate estimates of soils, soil pH, and climate (rainfall) are available. Third, measures of timing of activities, particularly given weather events, are important. Because fertilizer delivered under government programs in Zambia are often reputed to arrive after the optimal planting time, the surveys asked respondents whether urea and basal fertilizer were available in a timely manner. Unfortunately, about 15 percent of the households using fertilizer do not report a response to the timeliness question. Fourth, collinearity between nitrogen and phosphorus use is a challenge in terms of separating out individual effects; some households use only urea while others use only basal but over 50 percent of households use N and P_2O_5 in the ratio, if not the level, recommended by the national extension service.

Phosphorus provides a particular challenge under both experimental and survey conditions since much of plant uptake in the current year is the result of previous applications and inherent soil fertility. Only a modest proportion of the phosphorous applied in the current year is absorbed by the crop in the current year. Second, the fact that the rate of the phosphorus is not observed creates an additional measurement error challenge from an estimation perspective. Under experimental conditions, an estimate of the available phosphorus is often taken from each replication of a treatment which is typically not feasible in a survey situation.

The approach taken in dealing with these challenges was to attempt to determine bias direction of parameter estimates and to use robust estimation techniques. In addition, supporting Monte Carlo simulation was done to provide better insight into the properties of the estimates of marginal products.

Theoretical Framework

Crop yields can be seen as a function of decision or management variables (fertilizer, for example) that are under the farmer's control and environmental variables such as weather and soil type that are beyond the farmer's control. The yield response model that maps decision and environmental variables to output can be written as

$$y = f(x_i, Env),\ i=1,\ldots,n. \qquad (1)$$

where y is the stochastic crop yield, x_i is the ith management input, and Env is a vector of environmental variables that are comprised of both stochastic and nonstochastic factors.

The farmer's expected profit maximizing decision is[1]

$$\underset{x_i}{Max}\ pE(y) - \sum_{i=1}^{n} w_i x_i \ \text{ subject to } x_i \geq 0 \qquad (2)$$

where p is the output price[2], w_i is the ith input price, and E is the expectation operator. If the yield response function is strictly concave which exhibits diminishing marginal product of input i, the first order condition for the optimal level of input i satisfies

$$p\frac{\partial E(y)}{\partial x_i} - w_i = 0 \qquad (3)$$

Simplifying (3) gives

$$\frac{\partial E(y)}{\partial x_i} = \frac{w_i}{p} \qquad (4)$$

Equation (4) suggests that the expected profit maximizing level of input i is the level at which the expected marginal product (MP) of input i is equal to the input-output price ratio. Because the optimal input level is directly affected by this price ratio, a change in the ratio leads to the corresponding alteration in the optimal solution. In addition, optimal input levels are expected to vary across agro-climatic regions since yield response functions are not likely to be the same across regions.

---

[1] Single enterprise is assumed for this model.
[2] All prices are assumed exogenous to the model.

If the yield response function is linear in input i, it exhibits constant marginal product and the optimal decision is either not to apply at all or apply as much as possible depending on whether the slope of the yield function is less or greater than the price ratio w_i/p. Similarly, if the yield response function is strictly convex in input i and at some input level x_0 the slope of the yield response function is equal to w_i/p, the optimal decision is to apply as much input i that is greater than x_0 as possible.

Above conventional input allocation rules are optimal for nitrogen but are not always optimal for phosphorous which has substantial carryover (storage) in the soil. Yield is affected by the total amount of available phosphorous which is determined by the amount added at the current period (x_t) and the stock of phosphorous carried over into the current period (s_t). Yield y is a function of the total available amount (x_t+s_t). If yield response to phosphorous is linear, the optimal input allocation rule discussed earlier still holds, i.e., the optimal allocation amount depends on the slope of the linear function and the input-output price ratio. However, if yield response is nonlinear in phosphorous, the conventional allocation rules are sub-optimal because the marginal product of phosphorous is now affected by both x_t and s_t instead of just x_t. For example, if yield response function is approximated by a quadratic function with decreasing MP, the approach that takes into account carryover has the following first order condition:

$$\beta_1 - \beta_2(x_t + s_t) = \frac{w_t}{p_t};\ \beta_1, \beta_2 > 0 \tag{5}$$

and the optimal amount is:

$$x_t = \frac{\beta_1}{\beta_2} - \frac{w_t}{\beta_2 p_t} - s_t \tag{6}$$

which is less than the amount without considering carryover:

$$x_t = \frac{\beta_1}{\beta_2} - \frac{w_t}{\beta_2 p_t} \tag{7}$$

Literature Review

Various functional forms have been used to model crop yield response to nutrients. Examples include polynomial, spline (bent-stick) and Mitscherlich-Baule. Also, there are well-defined statistical procedures to choose among non-nested alternatives.

Polynomial functions are commonly used to specify crop yield response to fertilizer applications as are linear to a plateau and quadratic to a plateau. These functional forms are relatively easy to estimate. The quadratic form of maize yield response to nitrogen and phosphorous can be expressed as:

$$Y = \beta_0 + \beta_1(N) + \beta_2(P) + \frac{1}{2}\beta_3(N)^2 + \frac{1}{2}\beta_4(P)^2 + \beta_5(NP) + \varepsilon \tag{8}$$

where Y is maize yield, N and P are nitrogen and phosphorous application rates, and ε is random error with mean 0 and variance σ^2. A linear functional form is nested when $\beta_3, \beta_4, \beta_5 = 0$. The function imposes nonzero elasticity of substitution ($\sigma \neq 0$) and exhibits diminishing marginal productivity for each factor if $\beta_1, \beta_2 > 0$, $\beta_3, \beta_4 < 0$. However, above conditions also imply that excessive nutrient application will cause yield decrease, which is questionable due to little empirical support. Polynomial functions do not allow for sharp bends on the response surface.

The spline function can be written as:

$$Y = \beta_0 + \beta_1(N) + \beta_2(P) + \beta_3 d_1(N - N^*) + \beta_4 d_2(P - P^*) + \varepsilon \tag{9}$$

where d_1=1 if $N \geq N^*$ and 0 if $N<N^*$, d_2=1 if $P \geq P^*$ and 0 if $P<P^*$, N^* and P^* are function parameters representing knots. Spline function imposes sharp bends on the response

surface and accommodates plateaus if we impose $\beta_3 = -\beta_1$ and $\beta_4 = -\beta_2$, which implies that maize will no longer respond to the applied nutrients after N^* and P^* levels of application rate. At this point, maize reaches maximum yield or the growth plateau. The limitation of spline function is that it imposes infinitely large elasticity of substitution between nutrients.

The Mitschelich-Baule function is written as follows:

$$Y = \beta_0[1-\exp(-\beta_1(\beta_2+N))][1-\exp(-\beta_3(\beta_4+P))]+\varepsilon \qquad (10)$$

where β_0 is the asymptotic yield plateau. It contrasts to the von Liebig in that it does not impose zero elasticity of factor substitution. This function accommodates near perfect or near zero factor substitution, and imposes a growth plateau.

Data and Methods

We use maize production data for 1996/97, 1997/98, 1998/99 and 1999/00 production seasons that are obtained from Central Statistical Office in Zambia to estimate maize yield response functions. The source of the maize production data is the Post Harvest Survey, a nationally representative annual survey covering roughly 7,500 rural households each year.

About 80% of commercial fertilizer is applied in two of Zambia's agro-climatic zones that are relatively well suited to maize production: IIa and III. Soil and rainfall conditions in the other two zones (I and IIb) have low cropping potential under rainfed conditions especially with fertilizer. Acrisols[3] is the dominant soil group in Zones IIa and III and the cases with Acrisols are evenly spread between these two zones. In addition,

---

[3] The soil survey unit at Mt Makulu has rated Acrisols as marginal to moderately suitable for maize production under low input management.

the majority of these cases fall in the pH levels in the range 4.1 and 4.2, principally 4.2. This opportunity allows comparison of response rates between IIa and III. The principal focus of this paper is in Zone IIa and III with the soil type of Acrisols and the predominant pH levels; this is where the much of the maize is produced and fertilizer is applied. More comprehensive analysis of maize yield response to fertilizers under a range of small farm conditions, management practices, and market conditions is contained in Govereh *et al* (forthcoming).

Variables collected in each of the years are the ones considered for pooling. Data in some years are more comprehensive than in other years but the additional data are excluded because they are not collected consistently throughout the four years (labor, for example). Households that have less than 0.15 hectare of maize planted area, or maize yields less than 300kg/ha or greater than 6000kg/ha, or nitrogen index[4] over 150kg/ha are excluded because of implausibility.

Farmers frequently follow the fertilizer application recommendation[5] in terms of ratios of N and P_2O_5, although not rates. This results in significant collinearity which is shown in Figure 1. In addition, there is measurement error in the amount of phosphorous available from carry-over and mineralization as well as the amount added from current application in the year in which it is applied.[6] Thus, nitrogen was used in the production function to capture the effect of the "package" of N and P_2O_5; some regressions were restricted to the predominant P_2O_5 to N ratios.

---

[4] Nitrogen index is calculated as 0.1*basal fertilizer (kg/ha)+0.46*top dressing fertilizer(kg/ha), and phosphorous index is 0.2*basal fertilizer(kg/ha).

[5] Four bags of basal and four bags of top dressing per hectare is recommended by the extension service.

[6] See Appendix.

Maize yield is affected by both controllable and uncontrollable factors such as (but not limited to) soil type and quality, rainfall amount and distribution, grower's managerial skills, amount of fertilizer applied, use of hybrid seed and other inputs, and timeliness of planting and applying the inputs. Having stratified the sample by agro-climatic zones, soil types and pH levels, and based on the production and management data available, this paper investigates a set of yield response functions with the following model specifications:

Maize yield=f(N, femaleHH, age, usepower, fertontime, usehybrid, seedontime, land)

where *N*=nitrogen index in kg/ha

femaleHH = 1 if the head of the household is female, 0 otherwise
age = age of the household head
usepower = 1 if the household used animal draft power or mechanical power for land preparation, 0 otherwise
fertontime = 1 if basal fertilizer was available on time, 0 otherwise
usehybrid = 1 if the household used hybrid seed, 0 otherwise
seedontime = 1 if seed was available on time, 0 otherwise
land = hectares of maize cultivated

The yield model is estimated for Zones IIa and III respectively and the estimates of the marginal products of *N*, which are exactly the coefficient estimates on *N,* are then used along with the price ratios to determine profitability of fertilizer use on maize for each group of households.

Maize price data used in this study are the farm-level median prices in each district obtained from the PHS surveys and the fertilizer prices are the farm-level median prices[7] for each district. These districts are segmented into groups: those that are at the provincial centers where the purchasing points are located and those considered remote, i.e., at least 200 kilometers away from the provincial centers.

Results

The effects of some independent variables[8] on maize yield are inconclusive due to measurement error/insufficient variation in these variables. Coefficient estimate on *FemaleHH*, for example, are not stable across specifications which is not surprising given that they are less than 15% of the total cases. Households are then divided into four groups according to whether they used power and whether fertilizer was available on time as follows: (i) *usepower*=0 and *fertontime*=0; (ii) *usepower*=1 and *fertontime*=0; (iii) *usepower*=0 and *fertontime*=1; and (iv) *usepower*=1 and *fertontime*=1.

Profitability of fertilizer use is expected to differ among these groups because of different nitrogen response functions. Plots of maize yield versus nitrogen index and the corresponding Lowess smoothing[9] curves for Zones IIa and III are presented in Figures 2-8 (Lowess smoothing cannot be carried out for group (ii) in Zone III due to insufficient observations). These graphs suggest a clear linear response up to the level of approximately 110kg/ha of nitrogen index.

We estimated both linear and quadratic functions and the quadratic terms were found insignificant for all these groups, so a linear function was used to model yield

---

[7] The median fertilizer prices represent the price in kwacha per kilogram of fertilizer assuming the 50th percentile charges of transport from purchasing point to the farm respectively in each district.

[8] These variables are *femaleHH, age, usehybrid, seedontime,* and *land.*

[9] Locally weighted regression of maize yield on nitrogen index.

response to nitrogen. Tables 1 and 2 report the regression results for each group in Zones IIa and III respectively.

Table 1: Regression Results for Zone IIa

	Constant	*N*	Number of Observations
Group (i): *usepower*=0 & *fertontime*=0	1191	1.40	12
Group (ii): *usepower*=1 & *fertontime*=0	768^{**}	15.23^{*}	21
Group (iii) *usepower*=0 & *fertontime*=1	829^{***}	17.44^{***}	99
Group (iv) *usepower*=1 & *fertontime*=1	824^{***}	21.24^{***}	202

Note: ***, **, ***** denote statistical significance at (5%,10%], (1%-5%], (0, 1%] levels respectively. *N*=nitrogen index in kg/ha.

Table 2: Regression Results for Zone III

	Constant	*N*	Number of Observations
Group (i): *usepower*=0 & *fertontime*=0	567	14.45^{**}	29
Group (ii): *usepower*=1 & *fertontime*=0	n/a	n/a	1
Group (iii): *usepower*=0 & *fertontime*=1	877^{***}	17.91^{***}	127
Group (iv): *usepower*=1 & *fertontime*=1	823^{**}	25.26^{***}	29

Note: ***, **, ***** denote statistical significance at (5%,10%], (1%-5%], (0, 1%] levels respectively. *N*=nitrogen index in kg/ha; n/a denotes not available due to insufficient data.

The coefficient estimates on nitrogen index are all significantly different from zero for all cases except for Group (i) in Zone IIa and Group (ii) in Zone III due to insufficient observations. In addition, the estimates of marginal products of *N* are lowest for Group (i) and highest for Group (iv) in both Zones, which is consistent with our expectation. Based on the results in Zone IIa, Group (iii) has a higher response than Group (ii), which might suggest that the availability of fertilizer on time has a greater impact on fertilizer efficiency than whether animal draft or mechanical power is used for land preparation.

Chow tests are conducted to compare the nitrogen response rates between Zones IIa and III for Groups (i), (iii) and (iv) respectively. We cannot reject the null hypothesis that response rates are the same for the same group of households in these two zones that have the same soil group and pH levels. This suggests that the difference in recipitation in Zones IIa and III with the same soil group of Acrisols and the pH levels between 4.05 and 4.55 has little impact on the nitrogen response rates.

Table 3 and Table 4 show the farm-level price ratios of median nitrogen index prices to median maize prices and the corresponding value-cost ratios (VCR) for the period 1999-2000 for each group of households in Zones IIa or III that are located at the provincial centers and in remote areas (200kms from the provincial center) respectively. Transport costs typically increase the price of fertilizer by 13 to 24 percent above prices in provincial centers.

Table 3: Farm-level Price Ratios and VCRs at Provincial Centers

			VCR			
District	Zone	Price Ratio	Group(i)	Group(ii)	Group(iii)	Group(iv)
Kabwe	IIa	13.09	0.11	1.16	1.33	1.62
Ndola	III	10.93	1.32	n/a	1.64	2.31
Chipata	IIa	18.81	0.07	0.81	0.93	1.13
Mansa	III	9.78	1.48	n/a	1.83	2.58
Lusaka	IIa	10.90	0.13	1.40	1.60	1.95
Kasama	III	12.33	1.17	n/a	1.45	2.05
Solwezi	III	11.31	1.28	n/a	1.58	2.23
Choma	IIa	16.30	0.09	0.93	1.07	1.30

Note: n/a denotes not available due to the absence of coefficient estimate in Table 2.

Table 4: Farm-level Price Ratios and VCRs in Remote Areas

			VCR			
District	Zone	Price Ratio	Group(i)	Group(ii)	Group(iii)	Group(iv)
Mumbwa	IIa	16.36	0.09	0.93	1.07	1.30
Lufwanyama	III	13.56	1.07	n/a	1.32	1.86
Lundazi	IIa	23.16	0.06	0.66	0.75	0.92
Kawambwa	III	12.05	1.20	n/a	1.49	2.10
Chongwe	IIa	13.74	0.10	1.11	1.27	1.55
Mporokoso	III	15.07	0.96	n/a	1.19	1.68
Kasempa	III	13.93	1.04	n/a	1.29	1.81
Namwala	IIa	20.26	0.07	0.75	0.86	1.05

Note: n/a denotes not available due to the absence of coefficient estimate in Table 2.

There are three clear patterns from Tables 3 and 4: (1) the *N*-maize price ratios are apparently higher in the remote districts than in their corresponding provincial centers; (2) for each district, VCR is the highest for Group (iv) and the lowest for Group (i); (3) for each group of households, VCR is lower in the remote districts than in their corresponding provincial centers.

If VCR greater than 2 is used as the criterion for determining fertilizer profitability, households that belong to Groups (i), (ii) or (iii) hardly benefited from fertilizer use, no matter whether they are located at the provincial centers or in remote areas. Fertilizer use was more likely to be profitable for Group (iv) households that are near/at provincial centers, obtained fertilizer on time and used animal draft power or mechanical power for land preparation. Households in remote areas face adverse conditions because of the high fertilizer-maize price ratios which result in low VCRs and consequently fertilizer use is unlikely to be profitable unless the households obtained fertilizer on time, used animal draft or mechanical power, and are located in districts where the price ratios were not so high.

Many smallholder farmers do not have enough liquid assets and have to apply for loans to purchase fertilizers. If interest rates on the loans are high, fertilizer use will not be profitable. For example, Group (iv) households that are located in Ndola and borrowed money to purchase fertilizers in the period 1999-2000 are not likely to financially benefit from fertilizer use if the short-term interest rate is 50% in which case the VCR becomes 1.54 instead of 2.31. The break-even interest rate for this case is 15.57% for VCR equal to 2. That is, if interest rates are higher than 15.57%, fertilizer use on maize for these groups of households is not likely to be profitable.

Conclusions

Post Harvest Survey data for the period 1996/1997 to 1999/2000 in Zambia are used to estimate maize yield responses to nitrogen index in Zones IIa and III with the soil group of Acrisols and the pH levels between 4.05 and 4.55. Statistical analyses of the estimation results suggest that the marginal product of nitrogen index is the highest for the group of households that obtained fertilizer on time and used animal draft or mechanical power for land preparation in each zone. The null hypothesis that Zones IIa and III have the same yield response function for the same group of households cannot be rejected.

Results from the economic analyses of fertilization suggest the following key messages. First, households that obtained fertilizer on time and used animal draft power or mechanical power in land preparation are more likely to find fertilizer use profitable than other groups of households located in the same district. Second, farmers’ proximity to the provincial centers has a significant impact on the profitability of fertilizer use.

Greater distances and transport costs from provincial centers erode the profitability of fertilizer use. Applying fertilizer is most likely to be profitable near provincial centers where the price ratio of fertilizers to maize is the lowest. Third, high time preferences for money also reduce the profitability of fertilizer use. Thus, despite achieving relatively high physical crop response rates to fertilizer use in some areas, small farmers may find fertilizer use unprofitable until efforts are made to reduce transportation costs and implicit interest rates as well as to ensure more timely delivery of fertilizer.

Future research is underway to examine the profitability of fertilizer use under a range of alternative levels and combinations for other soil groups and pH levels in Zambia.

Figure 1: Relationship between P_2O_5 and N Applied

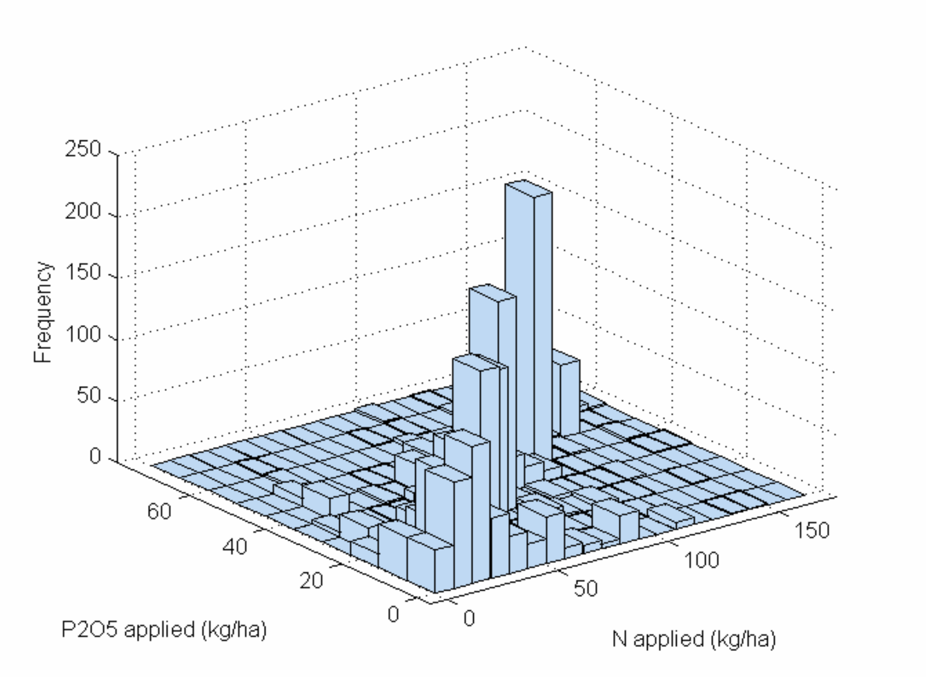


Figure 2: Yield versus Nitrogen Index for Group (i), Zone IIa

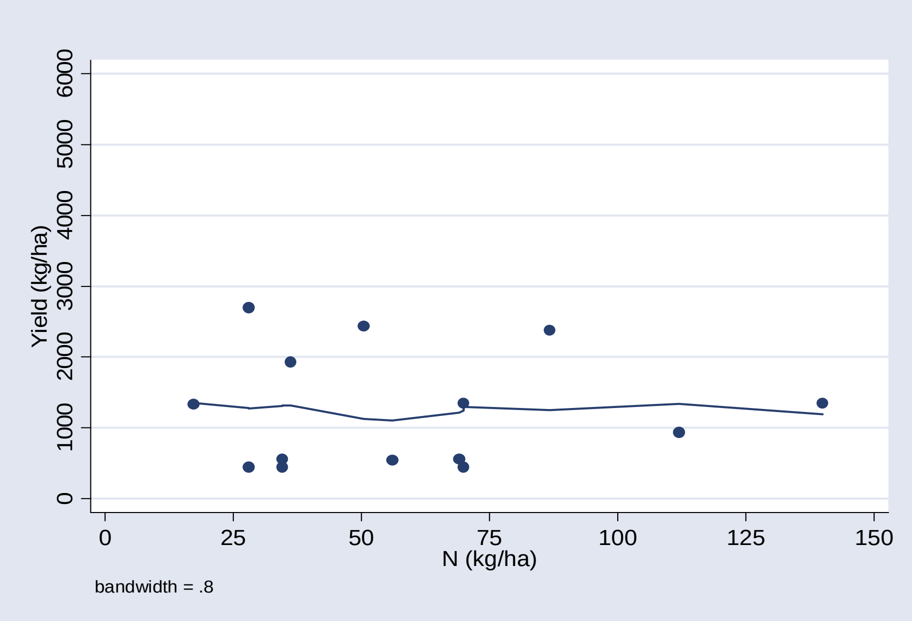


Figure 3: Yield versus Nitrogen Index for Group (ii), Zone IIa

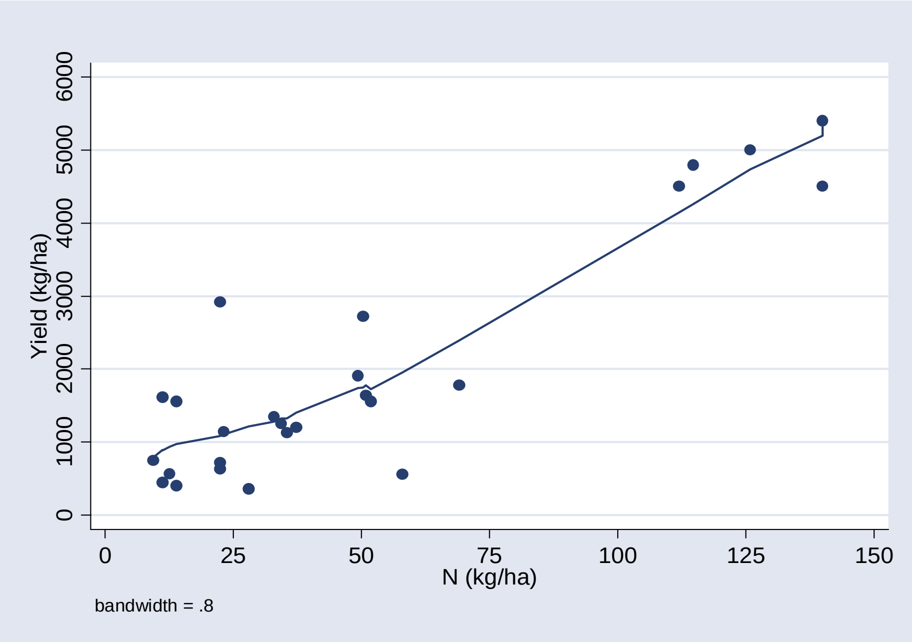


Figure 4: Yield versus Nitrogen Index for Group (iii), Zone IIa

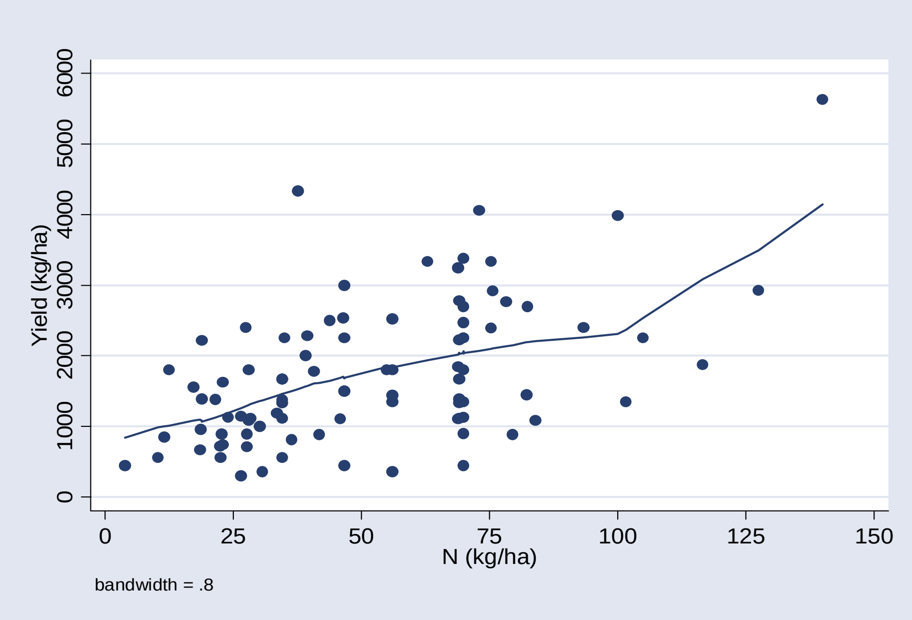


Figure 5: Yield versus Nitrogen Index for Group (iv), Zone IIa

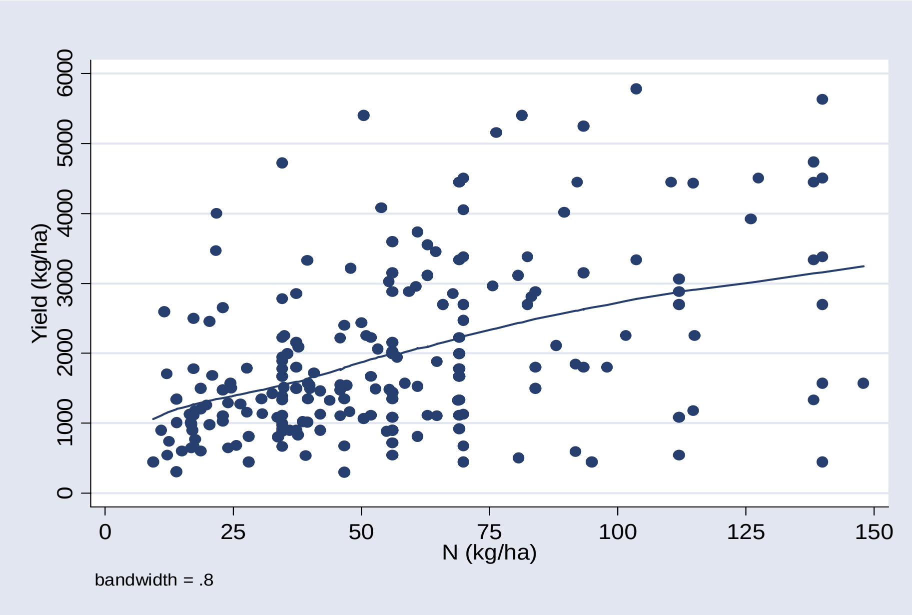


Figure 6: Yield versus Nitrogen Index for Group (i), Zone III

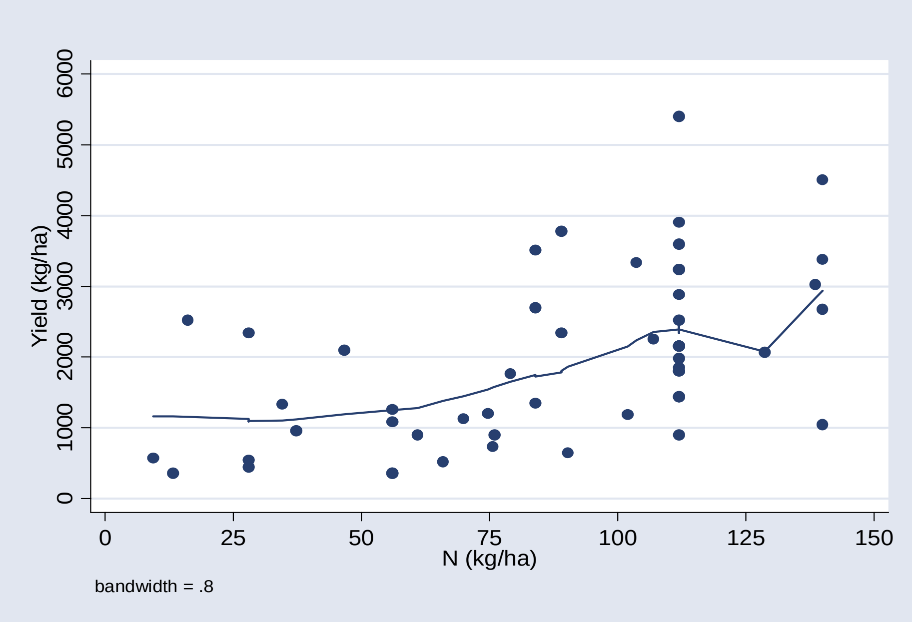


Figure 7: Yield versus Nitrogen Index for Group (iii), Zone III

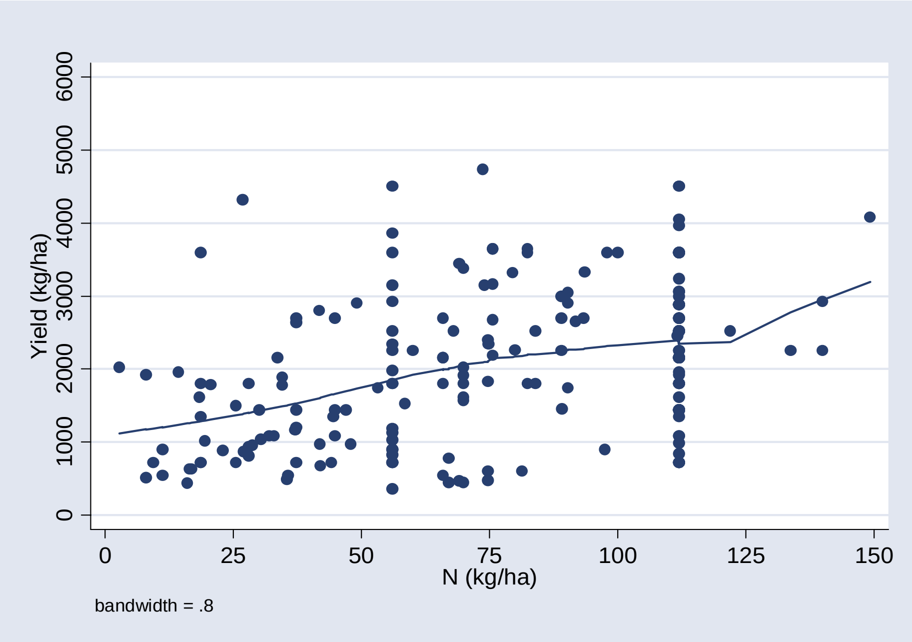
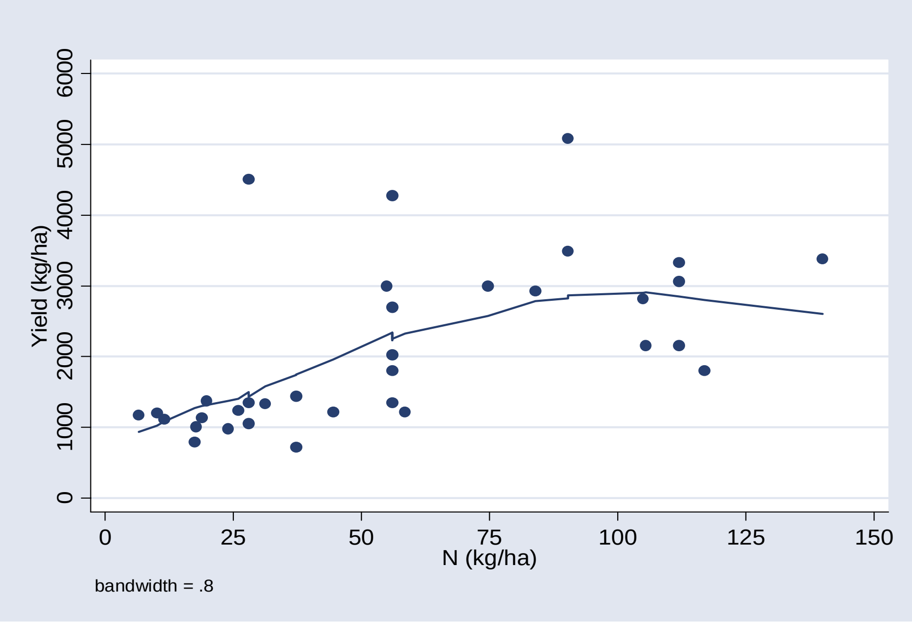


Figure 8: Yield versus Nitrogen Index for Group (iv), Zone III



References

Ackello-Ogutu, C., Paris, Q. and Williams, W.A. (1985) Testing a von Liebig crop response function against polynomial specifications. *Am. J. Agricultural Econ*. 67, 873-880.

Beattie, B.R. and Taylor, C.R. (1985) *The Economics of Production*. New York, John Wiley and Sons.

Berck, P. and Helfand, G. (1990) Reconciling the von Liebig and differentiable crop production functions. *Am.J. Agricultural Econ*. 72, 985-996.

Donovan, C., Damaseke, M., Govereh, J., and Simumba, D. (2002) Framework and intial analyses of fertilizer profitability in maize and cotton in Zambia. *Working Paper No.5, Food Security Research Project*, Lusaka, Zambia

Frank, M.D., Beattie, B.R. and Embleton, M.E. (1990) A comparison of alternative crop response models. *Am. J. Agricultural Econ*. 72, 597-603

Johnston, Bruce F., and John Mellor (1961). The Role of Agriculture in Economic Development. *American Economic Review* 51.4: 566-593.

Llewelyn, R. V. and Featherstone, A. M. (1997) A comparison of crop production functions using simulated data for irrigated corn in western Kansas. *Agricultural Systems*. 54, 521-538

Mjelde, J.W. et al (1991) Integrating data from various field experiments: the case of corn in Texas. *J. Production Agriculture* 4, 139-147

Paris Q. (1992) The von Liebig hypothesis. *Am. J. Agricultural Econ*. 74, 1019-1028
Schlegel, A.J. and Havlin, J.L. (1995) Crop response to long-term nitrogen and phosphorous fertilization. *J. Production Agriculture* 8, 181-185

Tiffen, Mary. 2003. Transition in Sub-Saharan Africa: Agriculture, Urbanization and Income Growth. *World Development,* 31(8): 1343-1366.

Vanotti, M.B. and Bundy, L.G. (1994) An alternative rationale for corn nitrogen fertilizer recommendations. *J. Production Agriculture* 7, 243-249

Appendix I

Measurement errors cause bias in parameter estimates. The essential concepts can be illustrated by analyzing a simple regression model without intercept term. Suppose that

$$y=\beta x^{*}+\varepsilon \qquad (1)$$

satisfies all the assumptions of the classical regression model. We assume that data on x^{*}, say, phosphorous, is not available. Our observed data x contain measurement error:

$$x=x^{*}+u,\ u\sim Normal(0,\ \sigma^{2}_{u})\ ,\ u \text{ is independent of } x^{*} \text{ and } y \qquad (2)$$

Substituting (2) in (1), we obtain

$$y=\beta(x-u)+\varepsilon=\beta x+(-\beta u+\varepsilon)=\beta x+v \qquad (3)$$

where x and v are correlated:

$$Cov(x,\ v)=Cov(x^{*}+u,\ -\beta u+\varepsilon)=-\beta\sigma^{2}_{u}$$

The assumption of no correlation between explanatory variable and the disturbance term is violated. Therefore least squares estimator for β is biased and inconsistent with a persistent bias toward zero. The larger the variability in the measurement error, the greater the bias would be.

We can also view (3) as having the omitted variable problem. We have omitted a variable, u, when we regress y on x. If we write (3) in the form

$$y=\beta_{1}x+\beta_{2}u+\varepsilon$$

and u were an observed variable, regressing y on x and u would produce unbiased estimates of β_{1} and β_{2} although they are not efficient ($\beta_{2}=-\beta_{1}$ is neglected). Both x and u affect y when they vary, but in a model that only x is the regressor, the effect of variation of u on y is transmitted through variation of x. Coefficient estimator is biased due to the omitted variable.

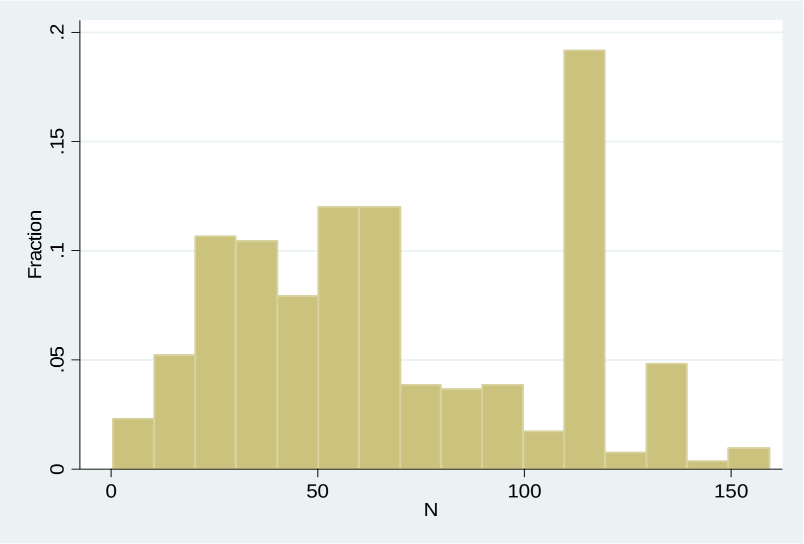
In a multiple regression model, not only is the coefficient on the poorly measured variable biased toward zero, the other coefficients are biased as well, but in unknown directions. Little can be said if more than one variable is measured with error in a multiple regression.

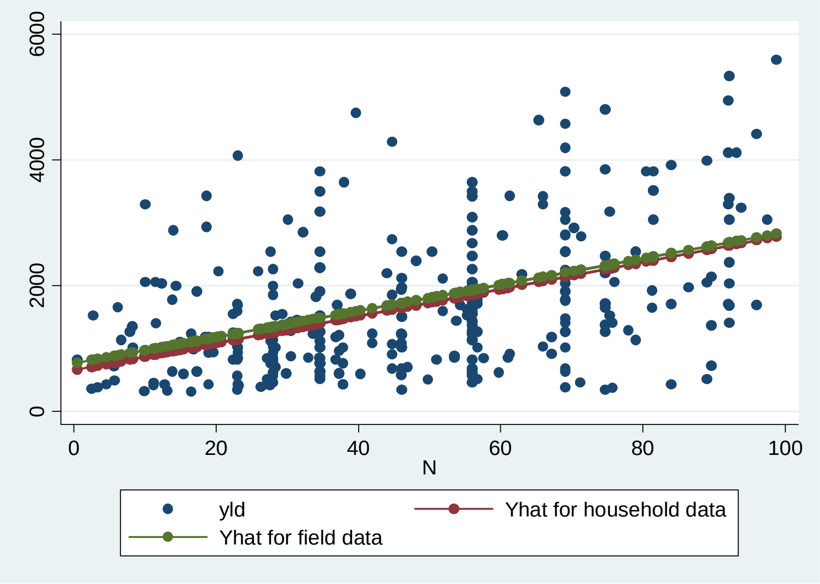
Appendix II

For the period 1996/97 to 1999/2000, only household level data are available from PHS. Household level data are essentially the aggregation of field level data and field level is the ideal level for estimating maize yield response functions. For the period 2000/2001, both household level and field level data are available which provides an opportunity to compare yield response to nitrogen at these two levels.

The coefficient estimates for two levels of data are essentially the same for the Acrisols which is consistent with out expectation because few households reported more than one field.

The first graph below is the histogram of nitrogen in Zones IIa and III for the soil group of Acrisols using field level data, and the second graph plots yield versus nitrogen at the household level and the regression estimates for both household level and field level data.





This result gives us some comfort in using household level data in previous years.