A theoretical derivation of the agricultural aggregate supply curve and research induced supply curve shift

Abstract

Conventional models of the aggregate supply of agricultural commodities—specifically those used in ex-ante impact assessments of policy and research shocks—do not endogenously model the supply response to technology shocks, but rather 1) assume the farm level supply response as exogenously given based on agronomic information, and then 2) assume, as a matter of fiat, the mechanism by which the farm level response manifests at the aggregate level (whether it is a shift or pivot of the aggregate supply curve, and what the adoption rate will be). Such models also do not account for farm heterogeneity—or, worse still, assume identical marginal farms—and do not account for farm entry/exit. For applications where these things matter, here a tractable aggregate supply model is formally derived from five weak axioms. The derived form is expressed in parameters that have clear economic meaning, including novel parameters such as price volatility and a term representing the mean probability of profit—which in turn endogenously accounts for farm entry/exit—thus broadening the range of policy questions that can be addressed in partial equilibrium analysis. Equations for the *endogenous* adoption rate, supply response, and producer welfare impact of technology shock are derived. As an illustrative use case, the model is fit to Kenya maize farm data. The fitted aggregate supply and price elasticity of aggregate supply curves agree with empirical estimates. The adoption rate, supply response, and producer welfare impact are calculated for a hypothetical technology shock.

1 Introduction

The conventional linear and constant elasticity of supply (CES) aggregate supply curves commonly used in ex-ante impact assessments of new agricultural technology (Alston et al. 1995; Alston et al. 2009) are not explicitly derived from microeconomic axioms, but rather specified primarily to facilitate calculations of consumer and producer surplus, with little consideration given to the plausibility of the theoretical underpinnings and empirical implications of this choice of functional form. In particular, it is unclear whether these forms are consistent with standard microeconomic axioms regarding homogeneity, convexity/concavity, profit maximization, and the slope of farm level supply and input demand curves.

That is to say, if we consider aggregate supply Q in its most basic definition as the definite integral of farm production function q^* times the production density $\rho(q^*)$ over the domain of profitable production $(\underline{q}^*, \overline{q}^*)$,

$$Q(P) = \int_{\underline{q}^*}^{\overline{q}^*} q^* \rho(q^*) \, dq^* \tag{1.1}$$

then it is not readily apparent from conventional linear and CES specifications of Q what the function q^* is, nor whether it is consistent with microeconomic theory. Critically, the conventional models also do not explicitly account for heterogeneously endowed farm populations (captured in the farm production density ρ)—or, worse still, assume a population of homogeneously endowed identical marginal farms; and they neglect to account for the domain of profitable production $(\underline{q}^*, \overline{q}^*)$, thus precluding any detailed analysis of supply response in terms of yield and entry/exit components.

Because these models are not formally derived from axioms, there is no clear mapping between their parameters and the relevant, measurable empirical phenomena (i.e., technology, input prices, agroclimatic conditions, demographic variables, etc.), such that it is not clear how a change in any of these phenomena might be accurately represented in terms of a change in parameter values, nor even which parameters might be affected by the change. This, in turn, precludes any endogenous modeling of economic decision making on the part of farmers in response to technology, or any other exogenous, shocks. Instead, the conventional models introduce farm level economic decision making (i.e., the economically rational supply/expenditure/input mix response to the exogenous shock) as a calibrated assumption based on agronomic data gathered from field trials, literature, and/or consultation with scientists involved in the research under evaluation.

This bears emphasis. The conventional aggregate supply models specifically designed for ex-ante impact assessment of new technology do not in fact model the supply response to new technology, but rather present a mathematical description of a presumed supply response that has been predetermined outside of the model, and then exogenously introduced into it. Model guidelines might offer some words of warning about the difference between economically and agronomically optimal yields/expenditures/input mixes, and about the glaring conflict of interest inherent in the expert opinion of scientists closely involved in the research under evaluation; but very little in the way of actionable advice is actually given. To redress conflicts of interest, e.g., Alston et al. (Alston et al. 1995) prescribe the creation of

> "an environment of peer review and, perhaps, a competitive process (e.g., Delphi methods) that will reduce the potential for personal incentives to bias estimates of technical parameters, and... an institutional setting in which scientists will be held accountable for systematic biases in their estimates of research impact".

Admirable aspirations, to be sure, but not actionable. If research institute directors are hard pressed to effect such deep cultural changes over the course of a multi-year tenure, how can the lowly economist tasked with the impact assessment—often hired by the scientists as a short term consultant—be expected to do so?

The subsequent step of translating the exogenously introduced farm level response into an aggregate supply response likewise involves no endogenous modeling of economic behavior, but rather exogenous guesswork translated into mathematical symbols. That is to say, after a long and inconclusive debate in the 1970s over whether the aggregate supply response should be represented as a pivot or parallel shift of the curve (Alston et al. 1995; Norton and Davis 1981; Voon and Edwards 1991), or some combination thereof, the current default choice of parallel shift in most impact assessments emerges not on any theoretical or empirical basis, but purely as a matter of fiat (Alston et al. 2009).

The final step of calculating net benefits in terms of discounted economic surplus, adjusted for probability of success, expected adoption rate, time lags in research and adoption, capital depreciation, and other exogenously collected parameters, bears many of the markings of a rigorous economic assessment. But insofar as the result of this calculation is prefigured by the initial, exogenously determined farm supply/expenditure response to technology change, and to its expression-by-fiat at the aggregate level, it is not so much an impact assessment as it is an impact assertion, i.e., an exercise in *petitio principii*.

Critics have long noted as much, if perhaps in a more understated way. Rose (1980), for example, notes that

> "[T]reatments have been concentrated on increasing the geometric sophistication of explanatory models and even that has been subject to fundamental errors. What is more important is to have an understanding of supply curves and the derivation of cost changes. Geometric manipulations are of little help at that level" (Rose 1980).

In other words, policy formation does not require mathematical descriptions of an exogenously assumed supply response to exogenous shocks, but rather endogenous modeling of the supply response to exogenous shocks. For those interested in such modeling, here an aggregate supply curve function is derived from the following five non-controversial economic axioms.

1) Diminishing marginal returns to input use

2) Farms maximize their expected net benefit subject to a budget constraint

3) Farms enter production only if they expect to make non-negative profit

4) Farm yield is lognormally distributed

5) Output price follows a geometric Brownian motion over time

In the course of the derivation, it becomes evident that five other desirable properties immediately follow from axioms 1 and 2 as corollaries, and need not be imposed as additional axioms: 1) The farm yield function is homogeneous (and thus homothetic) and 2) concave in inputs; 3) farm supply curves are upward sloping; 4) input demand curves are downward sloping; and 5) the derived yield function implies agreement with a consensus view of the economic rationale driving technology adoption, articulated by Lipton (2005).

On the other hand, two somewhat controversial corollaries also follow from this axiomatic approach: strictly decreasing returns to scale and constant marginal cost. The former contradicts the conventional wisdom handed down by Hicks (1963) and Solow (1957). The latter is well supported empirically (Lavoie 2014), but is eschewed in textbooks. Regardless of one's view on returns to scale and marginal cost curves, the very fact that it is possible to disagree with specific corollaries, and to trace the disagreement back to specific axioms, underscores a key strength of the derived model: its theoretical transparency. Under the theoretical opacity of the conventional approach, by contrast, it is not even possible to know if one agrees or disagrees with the premises, or whether the premises contradict each other mathematically, since those premises are stated in literary form (if they are stated at all), and the model is not explicitly derived from such premises.

All of the parameters appearing in the derived model have a straightforward interpretation in terms of relevant, measurable, empirical phenomena. This includes, in addition to the usual variables of interest (technology, agroclimatic variables, output price), farm expenditure population parameters and terms capturing the farm employment rate (i.e. farm entry/exit) and mean probability of profit. This facilitates the exploration of a range of distributional policy issues that conventional modeling approaches are admittedly ill suited to address (Norton, Pardey and Alston 1992). When stochasticity is taken into account, the aggregate supply curve is shown to have a dual expression in stochastic price parameter space, whereby questions of price volatility may be addressed.

In the Methodology section, immediately below, the model is derived from axioms 1-5. This includes a derivation of equations for assessing the endogenous supply response and producer welfare impact given an exogenous technology shock, together with an endogenous model of the adoption rate. The adoption rate model includes a term accounting for subsidies. The stochasticity of meteorology and output price is accommodated by axiom 5 at the end of the section. All of the salient differences between the present work and the conventional modeling framework stem from this explicit, careful accounting of the endogenous response to exogenous shock, and of the stochastic nature of meteorology and output price.

After the Methodology, the derived model is fit to Kenya maize farm data as an illustrative use case. The fitted maize supply model is vetted against empirical estimates of maize supply at different price points across seven years. The model's estimates of both supply and the price elasticity of supply agree closely with the empirical estimates. The model's explicit stochastic accounting of rainfall proves critical to this accuracy. A hypothetical, generic technology shock is then introduced in order to showcase the model's impact assessment capacity. The endogenously modeled aggregate supply response to the shock, and associated producer welfare outcomes, are evaluated at a range of different subsidy levels and assumed price shocks. Results are presented in a heatmap format to facilitate policy discussion.

The widespread use of conventional linear and CES modeling frameworks calls for a more detailed engagement with the literature than the cursory treatment offered above. To place the motivation of the present work on solid footing, this engagement is provided in Appendix A.

2 Methodology: Theoretically derived, endogenous modeling of aggregate supply and response to exogenous shock

Below, a deterministic model of aggregate supply and endogenous response to technology (or any other kind of exogenous) shock is derived from axioms 1-5. First, the farm production function is derived from axioms 1 and 2. Then a deterministic aggregate supply function is derived from the farm production function and axioms 3 and 4. (This is deterministic in the sense that it does not take into account the stochasticity of output price and weather.) An alternative deterministic form for aggregate supply is also derived based on the relation between conditional and unconditional means. Practical steps for evaluating curve parameters are briefly enumerated; and the importance of accounting for "stratified TFP"—i.e., heterogeneous total factor productivity—is addressed. The farm level endogenous yield and expenditure response to exogenous shock is then formalized—including a formula for calculating the endogenous adoption rate—and aggregated. Finally, stochasticity and axiom 5 are introduced to explicitly account for random fluctuations in weather and output price, resulting in the final functional form for expected aggregate supply.

2.1 Derivation of the farm production function from axioms 1 and 2

Throughout much of written history, astute observers have noticed that the usefulness of increments in the quantity of a good is inversely proportional to amount of the input or good or service already acquired (Kauder 1953). In 1738, Daniel Bernoulli put it this way:

> "In the absence of the unusual, the utility resulting from any small increase in wealth will be inversely proportionate to the quantity of goods previously possessed" (Bernoulli 1954).

Today, this is known as the "law of diminishing marginal returns". Bernoulli's statement is a consumer side formulation of the law. It can also has a producer side formulation: "In the absence of the unusual, the yield increase resulting from any small increase in a given input will be inversely proportionate to the quantity of the input already applied." With the aid of calculus (which was not available to Bernoulli), this may be formalized as follows.

$$\frac{\partial \ln(y)}{\partial x} \sim \frac{1}{x} \tag{2.1}$$

Where, to be clear, the term $\partial \ln(y)/\partial x$ is the percentage change in yield given a marginal increase in the quantity of good x, and "$\sim 1/x$", means "inversely proportionate to the quantity of good x already acquired". An expression $y(x)$ for the yield of good x can then be obtained by integrating Bernoulli's statement.

$$\begin{aligned} \int \frac{\partial \ln(y)}{\partial x} \, dx &= \int \alpha \frac{d \ln(x)}{dx} \, dx \\ \ln(y(x)) &= \alpha \ln(x) + k \\ y(x) &= e^k x^\alpha \end{aligned} \tag{2.2}$$

Where the proportionality constant $\alpha = \partial \ln(y)/\partial \ln(x)$ is the input elasticity of yield, i.e., the percentage change in yield given a 1 percent increase in input x.

In the multivariate case, an expression $y(\mathbf{x})$ for yield as a function of a vector of n goods $\mathbf{x} = [x_1, x_2, ..., x_n]$ can be obtained as follows.

$$\int \nabla_{\mathbf{x}} \ln(y) \cdot d\mathbf{x} = \int D(\boldsymbol{\alpha}) \nabla_{\mathbf{x}} \ln(\mathbf{x}) \cdot d\mathbf{x} \tag{2.3}$$

$$\begin{aligned} \ln(y(\mathbf{x})) &= \boldsymbol{\alpha} \cdot \ln(\mathbf{x}) + \alpha_0 \\ y &= e^{\alpha_0} \prod_{i=1}^{n} x_i^{\alpha_i} \end{aligned} \tag{2.4}$$

where $\boldsymbol{\alpha}$ is the vector of input elasticities corresponding to each x_i, the notation $D(\boldsymbol{\alpha})$ indicates a diagonal matrix with the elements of $\boldsymbol{\alpha}$ along its diagonal, and

$$\nabla_{\mathbf{x}} \ln(y) = \begin{bmatrix} \frac{\partial \ln(y)}{\partial x_1} \\ \frac{\partial \ln(y)}{\partial x_2} \\ \vdots \\ \frac{\partial \ln(y)}{\partial x_n} \end{bmatrix} \tag{2.5}$$

It is straightforward to see that equation 2.6 is a homogeneous (and thus also homothetic) function with degree of homogeneity h equal to $\sum_i \alpha_i$. There is no need, then, to impose homogeneity as an additional axiom. It follows from the axiom of diminishing marginal returns.

Empirical analysis involves exogenously given control variables (k_i), typically including demographic variables such as age, education, gender, household size, as well as agroclimatic variables such as temperature, rainfall, soil conditions, and so forth. Equation 2.4 is thus henceforth rewritten as follows.

$$y = g \prod_{i=1}^{n_k} k_i^{\kappa_i} \prod_{i=1}^{n_x} x_i^{\alpha_i} \quad ; \quad g = e^{\alpha_0} \tag{2.6}$$

Where the κ_i are the control variable elasticities of yield. Because the control variables k_i are non-discretionary, the corresponding elasticities κ_i may take positive or negative values. The input elasticities α_i, on the other hand, which correspond to discretionary management variables, are generally positive. For purposes of economic analysis, the degree of homogeneity is calculated only over the production elasticities corresponding to management variables ($h = \sum_i \alpha_i$, omitting the κ_i).

This is, of course, the Cobb-Douglas form. Outside of Economics, it would simply be referred to as a (multivariate) power law. Such forms are often good models of observed processes across a wide range of research contexts and disciplines. In Economics, the Cobb-Douglas form was widely used from the 1920s to 1950s on the basis of its good empirical fit, but fell into disrepute thereafter for its perceived lack a theoretical basis. Capalbo and Antle (2015) offer a succinct introduction to this vast discussion. The forgoing derivation of the Cobb-Douglas form from axiom 1 in some measure redresses the lack of a theoretical basis. A full defense against all of the criticisms levied against this form over the past seventy five years is a book-length project that will not be attempted here.

Interpretation of the constant g is as difficult as one wants to make it. It is, of course, usually called the "total factor productivity" (TFP), and understood to represent the yield per unit "all inputs"—although a more accurate description would be yield per unit "all input and control variables, raised to various powers". The constant g may in fact be a sum of binary endowments $b_i\beta_i$ that augment the TFP base constant (β_0)

on farms where the endowment is present (where $b_i = 1$). That is to say,

$$\ln(g) = \alpha_0 = \beta_0 + \sum_{i=1}^{n_b} b_i \beta_i \tag{2.7}$$

In the Kenya maize farm exercise farther below, for example, there are four ($n_b = 4$) binary endowment variables b_i (which take a value of 1 if the endowment is present and 0 otherwise): use of oxen in land preparation, use of tractor in land preparation, use of hybrid maize seed, and land tenure. This means that g can vary across the population. In the Kenya maize exercise, there are 12 unique values of g in the population, corresponding to 12 unique combinations of the four binary endowments plus the base endowment constant β_0.

The parameter g may also be arbitrarily defined in terms of a new parameter (u) combined with judiciously chosen reference terms such that the new parameter has a meaningful interpretation and/or results in a more useful expression for yield. For present purposes, g is defined as follows

$$g = \frac{u}{E[\prod_{i=1}^{n_k} k_i^{\kappa_i}] \prod_{i=1}^{n_x} E[x_i]^{\alpha_i}} \tag{2.8}$$

Where u is a constant, such that y may be rewritten

$$y = u \frac{k}{E[k]} \prod_{i=1}^{n_x} \left(\frac{x_i}{E[x_i]} \right)^{\alpha_i} \quad ; \quad k = \prod_{i=1}^{n_k} k_i^{\kappa_i} \tag{2.9}$$

This substitution will soon facilitate interpretation and manipulation of the optimal yield function. But first the optimal yield function must be derived from axiom 2 (profit maximization). Defining farm net benefit or profit ($\mathcal{L}$) as

$$\mathcal{L} = Py - \lambda(C - B) \tag{2.10}$$

Where P is the output price, C is expenditure ($C = \mathbf{w}'\mathbf{x}$, where $w_i \in \mathbf{w}$ is the i^{th} input price), B is the farm budget and λ is the farm budget shadow price—that is, the marginal benefit to the farm of a marginal increase in budget ($\partial\mathcal{L}/\partial B = \lambda$)—then the first order conditions for a maximum,

$$\begin{aligned}\nabla_{\mathbf{x}}\mathcal{L} &= P\nabla_{\mathbf{x}}y - \lambda\mathbf{w} = \mathbf{0} \\ &= PyD(\mathbf{x})^{-1}\boldsymbol{\alpha} - \lambda\mathbf{w} = \mathbf{0} \end{aligned} \tag{2.11}$$

(Where $\mathbf{0}$ is a vector of zeroes equal to the length of $\mathbf{x}$.)

can be rearranged into an expression for the i^{th} input demand.

$$x_i^* = \frac{y^* P\alpha_i}{\lambda w_i} \tag{2.12}$$

Where $y*$ is the optimal yield, i.e., the yield function in Equation 2.6 evaluated at the optimal input demand x_i^*. Multiplying this expression for x_i^* by w_i and summing both sides over i then gives an expression for optimal expenditure (C^*).

$$C^* = R^*\frac{h}{\lambda}; \quad R^* = y^*P \tag{2.13}$$

Such that the input demand may be rewritten

$$x_i^* = \frac{C^*\alpha_i}{hw_i} \tag{2.14}$$

Taking the mean of both sides and rearranging,

$$\frac{E[x_i^*]}{E[C^*]} = \frac{\alpha_i}{hw_i} \tag{2.15}$$

Combining this with the previous equation gives

$$\frac{x_i^*}{E[x_i^*]} = \frac{C^*}{E[C^*]} \tag{2.16}$$

Such that the optimal farm yield y^* (Equation 2.9 evaluated at x_i^*) may be expressed

$$y^* = u\frac{k}{E[k]}\left(\frac{C^*}{E[C^*]}\right)^h \tag{2.17}$$

The expression can be simplified still further by taking the mean and solving for u.

$$u = E[y^*] \frac{E[C^*]^h}{E[C^{*h}]} \tag{2.18}$$

And then substituting this for u in the yield expression.

$$y^* = E[y^*] \frac{k}{E[k]} \frac{C^{*h}}{E[C^{*h}]} \tag{2.19}$$

This expression for optimal farm yield—henceforth referred to simply as the farm yield—is particularly useful as an econometric yield model, as it only requires information on the control variables k_i and total cost C^*, and no information on individual input quantities, as would be required when estimating Equations 2.6 or 2.9. Given farm data, one specifies the model

$$y^* = vkC^{*h}e^{\epsilon} \;\; ; \;\; v = \frac{E[y^*]}{E[k]E[C^{*h}]} \tag{2.20}$$

(Where ϵ is the error term.) Such that the estimated coefficient on logged farm cost may be directly interpreted as the degree of homogeneity for the population. When estimating the model in Equations 2.6 or 2.9, by contrast, h must be tallied by summing the estimated coefficients α_i, and will likely vary across the population since not all farms necessarily apply all the inputs specified in the model. The new term v, moreover, retains meaning as a measure of TFP—and arguably a more instructive one since all inputs are expressed in identical units of cost. It is henceforth referred to as the "expenditure TFP", and is related to conventional TFP as follows.

$$\begin{aligned} v &= \frac{u}{E[k]E[C^*]^h} \\ &= g\frac{\prod_i^{n_x} E[x_i]^{\alpha_i}}{E[C^*]^h} \\ &= \frac{g}{h}\prod_i^{n_x}\left(\frac{\alpha_i}{w_i}\right)^{\alpha_i} \end{aligned} \tag{2.21}$$

Where the last line follows from the input demand expression in Equation 2.14.

2.1.1 Returns to scale, cancavity, price elasticities of farm supply and input demand

The second order condition for maximum profit is

$$
\begin{aligned}
\mathbf{x}'(\nabla^2_{\mathbf{x}}\mathcal{L})\,\mathbf{x} &= P\mathbf{x}'(\nabla^2_{\mathbf{x}}y)\,\mathbf{x} < 0 \\
&= Ph(h-1) < 0
\end{aligned}
\tag{2.22}
$$

Where Euler's Theorem has been invoked in the second line of the condition. Axioms 1 and 2 thus require that yield exhibit decreasing returns to scale ($0 < h < 1$), whereby it also follows that yield is globally concave in inputs. (Increasing returns to scale would imply minimum profit and convexity in inputs, while constant returns would imply the break even point and linearity in inputs.)

The price elasticity of farm supply resolves as follows.

$$
\begin{aligned}
\frac{\partial \ln(y^*)}{\partial P} &= h\frac{\partial \ln(C^*)}{\partial P} \\
&= h\left(\frac{\partial \ln(y^*)}{\partial P} + \frac{1}{P}\right) \\
\frac{\partial \ln(y^*)}{\partial P} &= \frac{h}{1-h}\frac{1}{P} \\
\frac{\partial \ln(y^*)}{\partial \ln(P)} &= \frac{h}{1-h} \equiv \eta
\end{aligned}
\tag{2.23}
$$

(Where Equation 2.13 is invoked in the second line.) Because returns to scale are strictly decreasing, then farm supply must be positively sloped. Moreover, farm supply is elastic for $h \in (1/2, 1)$, inelastic for $h \in (0, 1/2)$, and linear when $h = 1/2$. Note that increasing returns to scale would imply negatively sloped supply curves. In Appendix B.1, the input own and cross price elasticities of demand are likewise derived and shown to be strictly negative only when returns to scale are decreasing.

2.2 Aggregation of farm level supply based on axioms 3 and 4

The derivation of a compact, instructive expression for y^* above brings us a step closer to parsing the aggregate supply function defined in Equation 1.1. Denoting area k_a such that $q^* = k_a y^*$,

$$
\begin{aligned}
Q &= N \int_{\underline{q}^*}^{\bar{q}^*} q^* \rho(q^*) dq^* \\
&= N \int_0^{\infty} k_a \int_0^{\bar{y}^*} y^* \rho(y^*) dy^* \rho(k_a) dk_a \\
&= N E[k_a] v \int_0^{\infty} k \int_0^{\bar{C}^*} C^{*h} \rho(C^*) dC^* \rho(k) dk \\
&= N E[k_a] v E[k] \int_0^{\bar{C}^*} C^{*h} \rho(C^*) dC^* \\
&= N E[k_a] \frac{E[y^*]}{E[C^{*h}]} \int_0^{\bar{C}^*} C^{*h} \rho(C^*) dC^*
\end{aligned} \tag{2.24}
$$

The integral is now over the domain of expenditure $C^* \in (0, \bar{C}^*)$ instead of the domain of production. Moreover, letting $C^\#$ denote optimal unit expenditure (also commonly referred to as the "average cost", $C^\# = C^*/y^*$), note that C^* and y^* can be rewritten in terms of $C^\#$ as follows.

$$
\begin{aligned}
C^* = C^\# y^* &= C^\# v k C^{*h} \\
&\rightarrow C^* = (vkC^\#)^{\eta/h} \;\; ; \quad \eta = \frac{h}{1-h} \\
&\rightarrow y^* = (vk)^{\eta/h} C^{\#\eta}
\end{aligned} \tag{2.25}
$$

And note that taking the expected value of the last line and rearranging gives

$$
v^{\eta/h} = \frac{E[y^*]}{E[k^{\eta/h}] E[C^{\#\eta}]} \tag{2.26}
$$

Such that aggregate supply can also be written in terms of an integral over $C^\# \in (0, \bar{C}^\#)$,

$$
\begin{aligned}
Q &= N E[k_a] v^{\eta/h} E[k^{\eta/h}] \int_0^{\bar{C}^\#} C^{\#\eta} \rho(C^\#) dC^\# \\
&= N E[k_a] \frac{E[y^*]}{E[C^{\#\eta}]} \int_0^{\bar{C}^\#} C^{\#\eta} \rho(C^\#) dC^\#
\end{aligned} \tag{2.27}
$$

To parse further, the upper bound on unit expenditure $\bar{C}^\#$ and density function $\rho(C^\#)$ are resolved from axioms 3 and 4 below. The farm employment rate is also resolved as a byproduct.

2.2.1 Upper bound on expenditure, farm employment rate, deterministic aggregate supply

Given the introduction of $C^{\#}$, note that Equation 2.13 simplifies to

$$C^{\#} = \frac{P}{\lambda}h \tag{2.28}$$

Axiom 3 states that farms enter production only if they expect to make non-negative profit, i.e. if $C^* \leq R^*$, which simplifies to $C^{\#} \leq P$. (It follows from Equation 2.13 or 2.13 that another way of expressing axiom 3 is $\lambda \geq h$.) Note, in passing, that the farm employment rate ($\mathcal{E}$) may thus be expressed as follows.

$$\mathcal{E} = \int_0^P \rho(C^{\#})dC^{\#} \tag{2.29}$$

Assuming yield is lognormally distributed (axiom 4), then so must be C^* and $C^{\#}$; and this expression reduces to

$$\mathcal{E} = \Phi\left(\frac{\ln(P) - m_{\#}}{s_{\#}}\right) \tag{2.30}$$

Where $m_{\#} = E[\ln(C^{\#})]$, $s_{\#}^2 = Var[\ln(C^{\#})]$, and $\Phi()$ is the standard normal cumulative distribution function.

Now, to resolve the integral in the expression for aggregate supply (Equation 2.27), note that, for any lognormally distributed variable χ and upper bound ξ,

$$\begin{gathered} \int_0^{\xi} \chi\rho(\chi)d\chi = E[\chi]\Phi\left(\frac{\ln(\xi) - m_{\chi}}{s_{\chi}} - s_{\chi}\right) \ ; \\ m_{\chi} = E[\ln(\chi)] \quad s_{\chi}^2 = Var[\ln(\chi)] \end{gathered} \tag{2.31}$$

(See Appendix B.2 for a proof.)

Substituting $\chi = C^{\#\eta}$ and $\xi = P$,

$$\int_0^P C^{\#\eta}\rho(C^{\#})dC^{\#} = E[C^{\#\eta}]\Phi\left(\frac{\ln(P) - m_{\#}}{s_{\#}} - \eta s_{\#}\right) \tag{2.32}$$

Substituting the right hand side of this expression for the integral in Equation 2.27, the expression for aggregate supply simplifies as follows.

$$Q = NE[k_a]E[y^*]\Phi\left(\frac{\ln(P) - m_\#}{s_\#} - \eta s_\#\right) \tag{2.33}$$

To reduce notational clutter, the shorthand $\mathcal{P}$ is henceforth introduced,

$$\mathcal{P}(\xi; m_\chi, s_\chi) = \Phi\left(\frac{\ln(\xi) - m_\chi}{s_\chi} - s_\chi\right) \tag{2.34}$$

Such that aggregate supply can be written more compactly as follows.

$$Q = NE[k_a]E[y^*]\mathcal{P}(P^\eta; m_\#, s_\#) \tag{2.35}$$

Axioms 1-4 thus give us a provisional tractable expression for aggregate supply in terms of interpretable parameters. The derived form is provisional because it is a deterministic expression that does not capture the stochastic nature of meteorology and price movements affecting agricultural output. The stochastic dimension of aggregate supply is introduced along with axiom 5 at the end of the section.

2.2.2 Conditional vs. unconditional means, and an alternative form for aggregate supply

It is important to keep in mind that $E[C^*]$ and $E[y^*]$ are *unconditional* means. Because axiom 3 truncates the upper tail of the distribution of $C^\#$ at P, then sample parameters calculated from observed data must be interpreted as *conditional*. That is to say, the sample means of expenditure and yield do not represent $E[C^*]$ and $E[y^*]$, but rather the conditional means $E[C^*|C^\# \leq P]$ and $E[y^*|C^\# \leq P]$. The conditional and unconditional means of any lognormally distributed variable χ with upper bound ξ are related as follows.

$$E[\chi|\chi \leq \xi]\Phi\left(\frac{\ln(\xi) - m_\chi}{s_\chi}\right) = E[\chi]\Phi\left(\frac{\ln(\xi) - m_\chi}{s_\chi} - s_\chi\right) \tag{2.36}$$

Whereby it follows that

$$E[C^{\#\eta}]\mathcal{P}(P, m_\#, s_\#) = E[C^{\#\eta}|C^\# \leq P]\mathcal{E} \tag{2.37}$$

And

$$E[y^*]\mathcal{P}(P^\eta, m_\#, s_\#) = E[y^*|C^\# \leq P]\mathcal{E} \tag{2.38}$$

Such that aggregate supply may also be written

$$Q = NE[k_a]E[y^*|C^\# \leq P]\mathcal{E} \tag{2.39}$$

Which may be useful for point estimates of aggregate supply along the curve since $E[y^*|C^\# \leq P]$ can be estimated from the observed yield distribution, and because $\mathcal{E}$ has a straightforward interpretation as the farm employment rate. However, this form is inconvenient for purposes of drawing the entire aggregate supply curve, as the observed $E[y^*|C^\# \leq P]$ is valid only for the P at which it is measured. The unconditional mean yield $E[y^*]$, on the other hand, is constant for all P. For this reason, Equation 2.35 is preferred for aggregate supply curve assessment. Moreover, the term $\mathcal{P}(P^\eta; m_\#, s_\#)$ turns out to have a meaningful economic interpretation that will become clear when the stochastic character of price is taken into account farther below.

2.2.3 Evaluating supply curve parameters in practice

Tractable expressions for mean expenditure and yield follow from Equations 2.25 and from the lognormality of y^* and C^* (axiom 4).

$$E[C^*] = e^{m_c + \frac{s_c^2}{2}} \tag{2.40}$$

Where

$$\begin{aligned} m_c &= \frac{1}{1-h}(\ln(v) + m_k + m_\#) \;\; ; \quad m_k = E[\ln(k)] \\ s_c^2 &= \frac{1}{(1-h)^2}(s_k^2 + s_\#^2) \;\; ; \quad s_k^2 = Var[\ln(k)] \end{aligned} \tag{2.41}$$

And

$$
\begin{aligned}
E[y^*] &= vE[k]E[C^{*h}] \\
&= vE[k]e^{hm_c + h^2 \frac{s_c^2}{2}}
\end{aligned}
\tag{2.42}
$$

The conditional means can be expressed by replacing m_c, s_c, $m_\#$, and $s_\#$ with the sample parameters $\bar{m}_c$, $\bar{s}_c$, $\bar{m}_\#$, and $\bar{s}_\#$, defined

$$
\begin{aligned}
\bar{m}_c &= E[\ln(C^*)|C^\# \leq P] \\
\bar{s}_c^2 &= Var[\ln(C^*)|C^\# \leq P] \\
\bar{m}_\# &= E[\ln(C^\#)|C^\# \leq P] \\
\bar{s}_\#^2 &= Var[\ln(C^\#)|C^\# \leq P]
\end{aligned}
\tag{2.43}
$$

In practice, the sample parameters $\bar{m}_\#$ and $\bar{s}_\#$ can be calculated from the sample parameters $\bar{m}_c$ and $\bar{s}_c$ by rearranging Equations 2.41 into expressions for $\bar{m}_\#$ and $\bar{s}_\#$.

$$
\begin{aligned}
\bar{m}_\# &= (1-h)\bar{m}_c - \ln(v) - m_k \\
\bar{s}_\#^2 &= (1-h)^2\bar{s}_c^2 - s_k^2
\end{aligned}
\tag{2.44}
$$

This is preferred to calculating the parameters $\bar{m}_\#$ and $\bar{s}_\#$ directly from observed C^*/y^* because observed yield is harvest yield, whereas the y^* appearing in all of the equations above refers not to harvest yield but to expected harvest yield at planting time. This becomes clear when stochasticity is taken into account.

Because $C^\#$ is lognormally distributed, the unconditional parameters $m_\#$ and $s_\#$ can then be deduced from the conditional parameters $\bar{m}_\#$ and $\bar{s}_\#$ by solving the following system of equations.

$$
\begin{aligned}
\bar{m}_\# &= m_\# - s_\# \frac{\phi(\zeta)}{\Phi(\zeta)} \\
\bar{s}_\#^2 &= s_\#^2 \left(1 - \zeta\frac{\phi(\zeta)}{\Phi(\zeta)} - \left(\frac{\phi(\zeta)}{\Phi(\zeta)}\right)^2\right) \;\; ; \;\; \zeta = \frac{\ln(P) - m_\#}{s_\#}
\end{aligned}
\tag{2.45}
$$

Whence it is then possible to evaluate the unconditional parameters m_c, s_c, $E[C^*]$, and $E[y^*]$.

2.2.4 TFP strata

As noted earlier (Equation 2.7), binary endowments $\beta_i b_i$, and hence TFP, vary across the population. In the Kenya maize farm example explored farther below, there are 12 distinct values of expenditure TFP, corresponding to 12 distinct combinations of the binary endowments, henceforth referred to as "TFP strata" and indexed by ℓ. It follows that $E[C^*]$ and $E[y^*]$ can also be disaggregated by TFP stratum.

$$\begin{aligned} &E[C^*]_\ell = e^{m_{c\ell}+\frac{s_c^2}{2}} \\ &\text{Where} \;\; m_{c\ell} = \frac{1}{1-h}(\ln(v_\ell) + m_k + m_\#) \end{aligned} \tag{2.46}$$

And

$$\begin{aligned} E[y^*]_\ell &= v_\ell E[k] E[C^{*h}]_\ell \\ &= v_\ell E[k] e^{h m_{c\ell} + h^2 \frac{s_c^2}{2}} \end{aligned} \tag{2.47}$$

Parameters unindexed by ℓ refer to their values evaluated at the average expenditure TFP v.

$$v = \sum_\ell \frac{N_\ell}{N} v_\ell \tag{2.48}$$

Where N_ℓ/N is the fraction of farms in the ℓ^{th} TFP stratum. That is to say, $E[y^*]$ and $E[C^*]$ refer to the unconditional mean yield and expenditure evaluated at v.

2.3 An endogenous model of supply response to and welfare impact of exogenous shock

The agronomic yield and expenditure response to an exogenous shock observed at experimental stations does not translate into the economic yield and expenditure response likely to occur in real farm populations. This economically obvious fact of life (Herdt 1988; Balasubramanian, Bell and Sombilla 2000) bears reiteration because confusion on this matter unfortunately still pervades the conversation in many areas of agricultural research impact assessment (Sumberg 2012).

As noted in the Introduction, the authors of the conventional modeling framework are cognizant of the difference between economic yields and agronomic yields; and offer some words of warning on the matter

(Alston et al. 1995). But they provide no practical, formal method to distinguish between the two or to explicitly relate the former to the latter, such that their model effectively blurs the technology shock and supply response together as a single, nebulous exogenous shock.

In this subsection, equations for the evaluation of the farm level endogenous yield and expenditure response to exogenous shock are derived and aggregated based on the model derived from axioms 1-4 in the subsections above. This includes endogenous assessment of the adoption rate. As mentioned earlier, all of the salient differences between the theoretical framework proposed here and the conventional approach stem from these two steps.

More specifically, whereas the conventional approach assumes the farm level expenditure and/or yield response to exogenous shock based on agronomic data and expert opinion, the approach explained below endogenously models the farm level response to the shock based on the theoretical model derived from axioms 1-4 above. The farm level response model explicitly accounts not only for the shock itself, but also for any anticipated changes in input and output price, and for any subsidies included in the technology package.

And whereas the conventional approach exogenously assumes the adoption rate, again typically based on the opinion of experts closely involved in the technology under evaluation, here the adoption rate follows endogenously from the theoretical derivation, based on whether or not adoption increases farm net benefit.

2.3.1 The endogenous farm yield and expenditure response to exgenous shock

The impact of yield scaling technologies and/or management practices on farm supply and expenditure can be represented through perturbations of the TFP parameter (g_ℓ). This is tantamount to adding a binary endowment variable to the TFP sum in Equation 2.7. The four binary endowments included in the Kenya maize model example may be considered yield scaling interventions that increase the value of g_ℓ on farms where the endowment is present. The percentage increase in TFP given the introduction of a new intervention $\beta_5 b_5$ on adopting farms (farms where $b_5 = 1$) may be calculated as

$$\begin{aligned} \partial \ln(g_\ell) &\approx \frac{g'_\ell - g_\ell}{g_\ell} \;\; ; \;\; g'_\ell = g_\ell e^{\beta_5 b_5} \\ &\approx e^{\beta_5} - 1 \end{aligned} \tag{2.49}$$

Note that the percentage increase is the same across all adopting farms regardless of the starting value g_ℓ.

Ideally, the value of β_5 should be estimated directly from field trials. If this is not possible, then the estimate can be made on the basis of crop modeling and expert opinion. Regardless of the path taken in this exercise, it is critically important to keep in mind that the objective is to estimate the change in TFP (β_5 or $\partial \ln(g)$), not the change in economic yield itself ($\partial \ln(y^*)$). If information is gathered in terms of the change in agronomic yield and/or expenditure, then it must be converted into an estimate for $\partial \ln(g)$. The percentage changes in economic yield, expenditure, and expenditure TFP (v) then follow from Equations 2.25 as a function of the change in TFP as well as any change in output and input prices.

$$\begin{aligned} \frac{\partial \ln(y^*)}{\partial \ln(g)} &= \frac{1}{1-h}\left(1 - \boldsymbol{\alpha}' \frac{\partial \ln(\mathbf{w})}{\partial \ln(g)} + h \frac{\partial \ln(P)}{\partial \ln(g)}\right) \\ \frac{\partial \ln(C^*)}{\partial \ln(g)} &= \frac{1}{1-h}\left(1 - \boldsymbol{\alpha}' \frac{\partial \ln(\mathbf{w})}{\partial \ln(g)} + \frac{\partial \ln(P)}{\partial \ln(g)}\right) \\ \frac{\partial \ln(v)}{\partial \ln(g)} &= 1 - \boldsymbol{\alpha}' \frac{\partial \ln(\mathbf{w})}{\partial \ln(g)} \end{aligned} \tag{2.50}$$

(See Appendix B.3 for details.)

The direction of the expected change in prices depends upon the type of research product. Most agricultural research currently falls into one of two broad paradigms: "sustainable agricultural intensification" (SAI) or "agroecological intensification" (AEI) (Mockshell and Kamanda 2018). Widespread adoption of a new SAI technology typically results in increased output supply and input demand, and hence a lower output price and higher input prices. Widespread adoption of an AEI technology, on the other hand, can result in lower input demand and emphasizes yield stability over quantity, which could result in more nuanced impacts (e.g. lowering output price volatility but not necessarily the price itself). For illustrative purposes, in this article the new technology is assumed to be an SAI research product, such that $\partial \ln(\mathbf{w})/\partial \ln(g) > 0$ and $\partial \ln(P)/\partial \ln(g) < 0$.

2.3.2 The endogenous adoption rate

The farm net benefit function in Equation 2.10 evaluated at y^* gives the farm net benefit envelope.

$$\begin{aligned} \mathcal{L}^* = \mathcal{L}|_{y=y^*} &= R^* + \lambda(B - C^*) \\ &= \frac{\lambda}{h}((1-h)C^* + hB) \end{aligned} \tag{2.51}$$

The welfare effect on a given farm of adopting the new technology may be examined by differentiating $\mathcal{L}^*$

with respect to the change in TFP.

$$
\begin{aligned}
\frac{\partial \mathcal{L}^*}{\partial \ln(g)} &= \frac{\lambda}{h}\left((1-h)\frac{\partial C^*}{\partial \ln(g)} + h\frac{\partial B}{\partial \ln(g)}\right) \\
&= \frac{\lambda}{h}\left(C^*\left(1-\boldsymbol{\alpha}'\frac{\partial \ln(\mathbf{w})}{\partial \ln(g)} + \frac{\partial \ln(P)}{\partial \ln(g)}\right) + h\frac{\partial B}{\partial \ln(g)}\right)
\end{aligned}
\tag{2.52}
$$

Assuming farmers adopt only if $\partial\mathcal{L}^*/\partial\ln(g) \geq 0$ (this is implicit in axiom 2), then the adoption condition is

$$
\begin{aligned}
&\partial \ln(P) - \boldsymbol{\alpha}'\partial \ln(\mathbf{w}) + \partial \ln(g) + \frac{h}{C^*}\partial B > 0 \\
&\qquad \partial \ln(\nu) + \partial \ln(g) + \frac{h}{C^*}\partial B > 0 \ ; \quad \partial \ln(\nu) = \partial \ln(P) - \boldsymbol{\alpha}'\partial \ln(\mathbf{w})
\end{aligned}
\tag{2.53}
$$

Where the term ∂B represents the subsidies, costs, or any other exogenous mitigating circumstances, per unit area, entailed in the adoption of the technology package. In the absence of such mitigating circumstances, the condition simplifies to

$$
\partial \ln(g) > -\partial \ln(\nu) \tag{2.54}
$$

As mentioned above, for SAI technologies the expectation is that $\partial \ln(P) < 0$, $\partial \ln(\mathbf{w}) > 0$, and thus $\partial \ln(\nu) < 0$. In words, then, the adoption condition says that a farm stands to increase its net benefit through adoption of yield scaling technology only if the resulting percentage increase in TFP ($\partial \ln(g)$) is greater than the corresponding percentage decrease in output price minus the corresponding percentage increase in input prices scaled by their respective input elasticities (α_i). Note the resemblance between this and the "price/productivity tightrope" condition formulated by Lipton (2005).

> "New [SAI] science usually raises farm supply of outputs and demand for inputs. That makes outputs cheaper and inputs more expensive; hence the ratio of farm output prices to input prices falls. Do small and poor farmers gain? If, and only if, this science-induced fall in their relative farm prices is slower than the science-induced rise in their conversion ratio of physical inputs into physical outputs (that is, TFP)."

The only substantive difference between Lipton's condition and the condition in Inequality 2.54 is that the percentage increases in input prices in the latter are scaled by their respective efficiency parameters ($\boldsymbol{\alpha}$). And of course the more complete version of the adoption condition in Inequality 2.53 differs from Lipton by also including the exogenous net subsidy term.

If the efficiency adjusted output-input price ratio decrease is greater than the increase in TFP (g), then the margin by which it is greater is the net subsidy term $h/C^*\partial B$ required to make adoption profitable. Note that the required net subsidy ∂B will vary across farms based on C^*. The factor h/C^*, which may also be written λ/R^*, represents each farm's marginal valuation of the subsidy; and implies that the subsidy is relatively more valuable to farms with relatively lower expenditure. Solving Inequality 2.53 for C^* gives the maximum expenditure level for which adoption is profitable if the (absolute value of the) efficiency adjusted output-input price ratio decrease is greater than the increase in TFP ($|\partial \ln(\nu)| > \partial \ln(g)$).

$$C^* \leq \frac{-h\partial B}{\partial \ln(\nu) + \partial \ln(g)} \tag{2.55}$$

If $\partial B = 0$, then Inequality 2.53 implies that adoption will be either 0% or 100%. Empirically, this is almost never observed. Adoption rates tend to fall somewhere between 0% and 100%. The increased input demand associated with new technology adoption implies in and of itself that, without subsidies, $\partial B < 0$. And, of course, every major green revolution success story has involved subsidies of one kind or another to offset the required increase in input expenditure, and then some (Patel 2013; Pingali 2012). In his Nobel lecture, Norman Borlaug said that "equally as important as" the research product itself was the "sound governmental economic policy which would assure the farmer a fair price for his grain, the availability of the necessary inputs... and the credit with which to buy them" (1970). Within the theoretical framework developed here, this means that the net subsidy ∂B is generally positive.

Denoting the upper adoption bound on C^* (Equation 2.55) as $\hat{C^*}$, the adoption rate ($\mathcal{A}$) when $\partial \ln(g) < -\partial \ln(\nu)$ may be expressed as the integral of the density $\rho(C^*)$ over the domain $(0, \hat{C^*})$.

$$\begin{aligned} \mathcal{A} &= \int_0^{\hat{C^*}} \rho(C^*) dC^* \\ &= \Phi\left(\frac{\ln(\hat{C^*}) - m_c}{s_c}\right) \\ &= \Phi\left(\frac{\frac{-h\partial B}{\partial \ln(\nu) + \partial \ln(g)} - m_c}{s_c}\right) \end{aligned} \tag{2.56}$$

The aggregate net subsidy required to reach a given target adoption rate $\mathcal{A}'$ can then be computed as the subsidy per unit area times the total farm area times $\mathcal{A}'$.

$$\text{Aggregate net subsidy} = NE[k_a]\mathcal{A}'\partial B \tag{2.57}$$

The adoption rate within a single TFP stratum can be calculated by evaluating this expression at the v_ℓ of interest.

2.4 Aggregation of the endogenous farm supply response and change in welfare due to technology shocks

The total increase in aggregate supply due to a technology shock can be evaluated by integrating the change in farm production ∂q^* over the density of adopting farms $(0, \hat{y^*})$ where $\hat{y^*} = y^*(\hat{C^*})$.

$$\begin{aligned} \Delta Q &= N \int_0^{\hat{y^*}} k_a \partial y^* \rho(y^*) dy^* \\ &= N E[k_a] \int_0^{\hat{y^*}} y^* \partial \ln(y^*) \rho(y^*) dy^* \end{aligned} \tag{2.58}$$

From the expression for research induced yield change in Equations 2.50 it follows that

$$\partial \ln(y^*) = \frac{\ell}{1-h} \;\; ; \;\; \ell = \partial \ln(g) - \boldsymbol{\alpha}' \partial \ln(\mathbf{w}) + h \partial \ln(P) \tag{2.59}$$

Such that the integral reduces to

$$\begin{aligned} \Delta Q &= N E[k_a] \frac{\ell}{1-h} \int_0^{\hat{y^*}} y^* \rho(y^*) dy^* \\ &= N E[k_a] \frac{\ell}{1-h} v E[k] \int_0^{\hat{C^*}} C^{*h} \rho(C^*) dC^* \\ &= N E[k_a] v E[k] E[C^{*h}] \frac{\ell}{1-h} \Phi \left(\frac{\ln(\hat{C^*}) - m_c}{s_c} - h s_c \right) \\ &= N E[k_a] v E[k] E[C^{*h}] \frac{\ell}{1-h} \mathcal{P}(\hat{C}^{*h}; m_c, s_c) \end{aligned} \tag{2.60}$$

Where the penultimate line follows from Equation 2.31. Recalling from Equation 2.21 that $v = E[y^*]/(E[k]E[C^{*h}])$, this further reduces to

$$\Delta Q = N E[k_a] E[y^*] \frac{\ell}{1-h} \mathcal{P}(\hat{C}^{*h}; m_c, s_c) \tag{2.61}$$

The new aggregate supply curve (Q') can then be evaluated as

$$Q' = Q + \Delta Q \tag{2.62}$$

To assess the aggregate change in net benefit, note from Equation 2.52 that the farm level change in net benefit per unit area $\partial\mathcal{L}^*$ may be expressed

$$\begin{aligned} \partial\mathcal{L}^* &= \frac{\lambda}{h}(C^*\gamma + h\partial B) \ ; \quad \gamma = \partial g - \boldsymbol{\alpha}'\partial\ln(\mathbf{w}) + \partial\ln(P) \\ &= r(C^*\gamma + h\partial B) \ ; \quad r = \frac{\lambda}{h} = \frac{P}{C^{\#}} \end{aligned} \tag{2.63}$$

And recall from Equation 2.55 that $h\partial B = -\gamma\hat{C}^*$, such that the expression further reduces to

$$\partial\mathcal{L}^*(C^*, r) = r\gamma(C^* - \hat{C}^*) \tag{2.64}$$

The aggregate change in net benefit ($\Delta\psi$) can be evaluated by integrating $\partial\mathcal{L}^*(C^*, r)$ over the densities of farm expenditure ($C^* \in (0, \hat{C}^{*h})$), revenue to expenditure ratio ($r \in (1, \infty)$) and area (k_a). (The domain of r follows from axiom 3.)

$$\begin{aligned} \Delta\psi &= N\int_0^\infty\int_1^\infty\int_0^{\hat{C}^*} \partial\mathcal{L}^*(C^*, r)\rho(C^*)dC^*\rho(r)dr\rho(k_a)dk_a \\ &= NE[k_a]\int_1^\infty r\rho(r)dr\int_0^{\hat{C}^*}\gamma(C^* - \hat{C}^*)\rho(C^*)dC^* \\ &= \gamma NE[k_a]E[r]\Phi\left(\frac{\ln(P) - m_{\#}}{s_{\#}} + s_{\#}\right)\left(E[C^*]\mathcal{P}(\hat{C}^*; m_c, s_c) - \hat{C}^*\mathcal{A}\right) \\ &= \gamma NE[k_a]E[r]\mathcal{P}(P^{-1}; m_{\#}, s_{\#})\left(E[C^*]\mathcal{P}(\hat{C}^*; m_c, s_c) - \hat{C}^*\mathcal{A}\right) \end{aligned} \tag{2.65}$$

But note that, by Equation 2.36,

$$\begin{aligned} E[C^*]\mathcal{P}(\hat{C}^*; m_c, s_c) &= E[C^*|C^* \leq R^*]\mathcal{A} \\ E[r]\mathcal{P}(P^{-1}; m_{\#}, s_{\#}) &= E[r|r \geq 1]\mathcal{E} \end{aligned} \tag{2.66}$$

Such that the previous expression reduces to

$$\Delta\psi = \gamma N E[k_a] E[r|r \geq 1] \mathcal{A}\mathcal{E}(E[C^*|C^* \leq \hat{C}^*] - \hat{C}^*) \tag{2.67}$$

Which is both a more instructive and a more practical form, since it is expressed in readily interpretable terms $\mathcal{A}$ and $\mathcal{E}$, and in terms of conditional parameters ($E[C^*|C^* \leq \hat{C}^*]$ and $E[r|r \geq 1]$), which are calculable from observed sample data.

To calculate the change in aggregate supply and welfare for each TFP stratum (ΔQ_ℓ and $\Delta\psi_\ell$), replace v, N, with v_ℓ, N_ℓ in the equations above.

2.5 Accounting for stochasticity, derivation of the expected aggregate supply curve based on axiom 5

The deterministic aggregate supply function derived in section 2.2.1 may be a suitable form for sectors characterized by deterministic production processes, negligible delays between input decisions and output, and negligible output price volatility. However, it is not suited to the primary sector, where production is partly deterministic but also partly a function of stochastic meteorological variables; where there is a significant delay between input decisions and output (i.e., the cropping season); and where output price volatility looms large in any management decision. In order to accommodate these hallmarks of agricultural supply, it is necessary to be more explicit in the treatment of yield and output price as stochastic variables.

2.5.1 Accounting for meteorological stochasticity

To begin with, yield is partly a function of stochastic meteorological variables unknown at planting time. The farm yield y^* appearing in the equations above does not refer to harvest yield, but rather expected harvest yield at planting time. This is written $E[y^*(T)]|_{t=0}$, where T represents the number of time steps from planting ($t = 0$) until harvest. An expression for $E[y^*(T)]|_{t=0}$ can be derived by first considering the expression for observed yield at harvest.

$$\begin{aligned} y^*(T; k_m(T)) = y^*(T) &= E[y^*(T; k_m(T))] \frac{k_m(T)^{\kappa_m}}{E[k_m(T)^{\kappa_m}]} \frac{k'}{E[k']} \frac{C^{*h}}{E[C^{*h}]} \;\; ; \quad k' = k_m(T)^{-\kappa_m} k \\ &= v \frac{k_m(T)^{\kappa_m}}{E[k_m(T)^{\kappa_m}]} k' C^{*h} \;\; ; \quad v = \frac{E[y^*(T)]}{E[k'] E[C^{*h}]} \end{aligned} \tag{2.68}$$

Where the stochastic meteorological component $k_m(T)^{\kappa_m}$ has been factored out of k. The component $k_m(T)$ may be a single stochastic meteorological component such as total rainfall by harvest time, or it may represent a product of cumulative meteorological variables raised to their respective elasticities. For example, if the key meteorological variables are cumulative rainfall ($k_r(T)$) and solar radiation ($k_s(T)$), then $k_m(T)^{\kappa_m} = k_r(T)^{\kappa_r} k_s(T)^{\kappa_s}$.

Taking the stochastic expectation of Equation 2.68 then gives an expression for expected yield $E[y^*(T)]|_{t=0}$.

$$E[y^*(T;k_m)]|_{t=0} = E[\tilde{y}^*(T;k_m)]|_{t=0} \left[\frac{k_m(T)^{\kappa_m}}{E[k_m(T)^{\kappa_m}]}\right]\Big|_{t=0} \frac{k'}{E[k']} \frac{C^{*h}}{E[C^{*h}]} \;\; ; \;\; \tilde{y}^*(T;k_m) = E[y^*(T;k_m)] \quad (2.69)$$

Note that taking the population mean of both sides of this equation results in a useful expression.

$$E[E[y^*(T;k_m)]|_{t=0}] = E[\tilde{y}^*(T;k_m)]|_{t=0} \quad (2.70)$$

This result, which can also be reached by invoking Fubini's theorem, maps the mean expected yield to the expected yield on the mean farm—i.e., the farm where $k = E[k]$ and $C^{*h} = E[C^{*h}]$.

Farm data reports harvest yield $y^*(T)$, but expected yield at planting time ($E[y^*(T;k_m)]|_{t=0}$) is generally not observable. Nonetheless, the implicit observed mean expected yield can be calculated for a given mean expected meteorological condition from fitted expenditure TFP (v), the mean observed exogenous continuous (non-meteorological) parameters raised to their fitted coefficients ($E[k']$), and mean observed C^{*h}. To see this, first note that, by Fubini's theorem, the mean expected meteorological condition may be expressed

$$E[E[k_m(T)^{\kappa_m}]|_{t=0}] = E[\tilde{k_m}(T)^{\kappa_m}]|_{t=0} \;\; ; \;\; \tilde{k_m}(T)^{\kappa_m} = E[k_m(T)^{\kappa_m}] \quad (2.71)$$

Such that the implicit observed mean expected yield (i.e. the stochastically explicit version of Equation 2.42) can be written

$$\begin{gathered} E[\tilde{y}^*(T;k_m)|\theta]|_{t=0} = vE[\tilde{k}_m(T)^{\kappa_m}]|_{t=0}E[k']E[C^{*h}|\theta] \\ \text{Where } \theta = E[C^{\#}(T)]|_{t=0} \leq E[P(T)]|_{t=0} \end{gathered} \quad (2.72)$$

Note that the optimal farm expenditure per acre C^* remains a deterministic variable because it refers primarily to land preparation, seed, fertilizer and chemical pest control. The main cost components are thus typically known (executed or committed) at planting time.

2.5.2 Accounting for output price stochasticity

Absent a forward contract, the output sale price P is a stochastic variable unknown to the farmer at the time of planting. Wherever P appears in the equations above, then, what is really meant is $E[P(T)]|_{t=0}$—the value that the farmer expects P to take at harvest/sale, as evaluated at planting time. To parse $E[P(T)]|_{t=0}$, a suitable stochastic model must be assumed for $P(t)$. Within a given season, prices often follow a geometric Brownian motion (gBm). That is to say, changes in P over a small time interval Δt ($\Delta P = P(t+\Delta t) - P(t)$) are often accurately modeled by

$$\Delta P = P(t) m_p \Delta t - P(t) s_p \varepsilon \sqrt{\Delta t} \tag{2.73}$$

Where ε is a normally distributed random variable with mean 0 and variance 1, such that the arithmetic change $\Delta P/P$ is normally distributed with mean and variance

$$\begin{aligned} E\left[\frac{\Delta P(T)}{P(T)}\right] &= m_p \Delta t \\ Var\left[\frac{\Delta P(T)}{P(T)}\right] &= s_p^2 \Delta t \end{aligned} \tag{2.74}$$

Such that the expected value of $P(T)$ resolves as

$$E[P(T)]\big|_{t=0} = P(0) e^{m_p T} \tag{2.75}$$

(For details, see Hull (2015).)

The choice of gBm as the stochastic model for $P(t)$ is axiom 5. For what follows, it is important to note that the log change of the gBm $\ln(P(T)/P(\hat{t}))$ is defined over any finite time interval for some starting time $\hat{t} < T$, and is also normally distributed, with mean and variance

$$\begin{aligned} E\left[\ln\left(\frac{P(T)}{P(\hat{t})}\right)\right] &= \left(m_p - \frac{s_p^2}{2}\right) T \\ Var\left[\ln\left(\frac{P(T)}{P(\hat{t})}\right)\right] &= Var\left[\ln(P(T))\right] = s_p^2 T \end{aligned} \tag{2.76}$$

For those unfamiliar with stochastic modeling but familiar with statistics, the choice of stochastic model may be roughly thought of as the choice of size distribution of changes in $P(t)$ per small time increment Δt. Selecting gBm as the stochastic model of $P(t)$ is tantamount to assuming that such changes are lognormally distributed. Some examples of maize prices following a gBm are given in Figure 1, accompanied by histograms of their log changes per time step. The histograms offer visual confirmation that price changes per time step are convergent to a lognormal distribution, and that gBm is therefore a reasonable assumption.

It is important to keep in mind that the price trajectory considered here is over a cropping season—from planting to harvest. When considered over multiple cropping seasons, prices typically exhibit cyclic or mean reverting behavior, which requires a more sophisticated stochastic model.

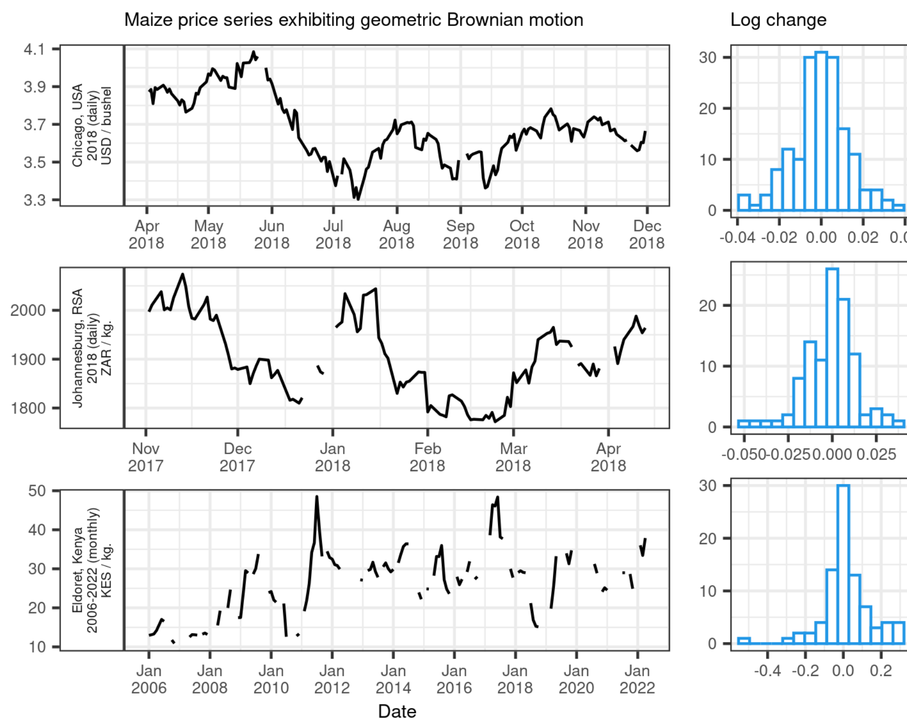

Figure 1: (Left) Maize price series from (top to bottom) Chicago, USA (CME, 2025), Johannesburg, RSA (JSE, 2025), and Eldoret market, Kenya (WFP, 2025). (Right) The corresponding daily (monthly, in the case of Kenya) log changes exhibit the normal distribution typical of geometric Brownian motion.

Farm cost per acre is deterministic, but optimal unit expenditure $C^{\#}$ is proportional to price, and must therefore also be stochastic. Rewriting Equation 2.28 in stochastically explicit terms,

$$
\begin{aligned}
E[C^{\#}(T)]\big|_{t=0} &= E[P(T)]\big|_{t=0}\frac{h}{\lambda} \\
&= P(0)e^{m_p T}\frac{h}{\lambda}
\end{aligned} \tag{2.77}
$$

Substituting the stochastic terms developed above for y^*, P, and $C^{\#}$ in the deterministic aggregate supply curve (Equation 2.35), gives the stochastically explicit aggregate supply curve.

$$
\begin{aligned}
E[Q(T)]|_{t=0} &= NE[k_a]E[\tilde{y}^*(T;k_m(T))]|_{t=0}\Phi\left(\frac{\ln(E[P(T)]|_{t=0}) - m_{\#}}{s_{\#}} - \eta s_{\#}\right) \\
&= NE[k_a]E[\tilde{y}^*(T;k_m(T))]|_{t=0}\mathcal{P}(E[P(T)^\eta], m_{\#}, s_{\#}) \\
m_{\#} &= E[\ln(E[C^{\#}(T)]|_{t=0})] \quad s_{\#}^2 = Var\ln(E[C^{\#}(T)]|_{t=0})
\end{aligned} \tag{2.78}
$$

Equation 2.38, relating the observed (conditional) mean and unconditional mean expected yield, can also be written in stochastically explicit form.

$$
E[\tilde{y}^*(T;k_m(T))]|_{t=0}\mathcal{P}(E[P(T)^\eta]|_{t=0}, m_{\#}, s_{\#}) = E[\tilde{y}^*(T;k_m(T))|\theta]|_{t=0}\mathcal{E} \tag{2.79}
$$

By which the stochastically explicit alternative version of Equation 2.39 follows.

$$
E[Q(T)]|_{t=0} = NE[k_a]E[\tilde{y}^*(T;k_m(T))|\theta]|_{t=0}\mathcal{E} \tag{2.80}
$$

2.5.3 Expected aggregate supply in stochastic price parameter space

In the remainder of this section, it is shown that the $\mathcal{P}(E[P(T)]|_{t=0})$ term has a meaningful economic interpretation. Specifically,

$$
\mathcal{P}(E[P(T)^\eta]|_{t=0}, m_{\#}, s_{\#}) = \text{Prob}(E[\tilde{C}^{\#}(T)^\eta]|_{t=0} \leq P(T)^\eta) \;\; ; \quad \tilde{C}^{\#}(T) = E[C^{\#}(T)] \tag{2.81}
$$

In words, the $\mathcal{P}$ term on the lefthand side equals the probability that the harvest output price (raised to η) will be greater than the mean farm's expected unit cost (raised to η). The term may thus be considered a reflection (up to the adjustment η) of the mean probability of profit. To see this, start by taking the population mean of both sides of Equation 2.77.

$$E[E[C^{\#}(T)]\big|_{t=0}] = E[P(T)]\big|_{t=0} hE[\lambda^{-1}] \tag{2.82}$$

But note also that

$$E[C^{\#}(T)] = P(T)hE[\lambda^{-1}] \tag{2.83}$$

Such that

$$E[E[C^{\#}(T)]]\big|_{t=0} = E[P(T)]\big|_{t=0} hE[\lambda^{-1}] \tag{2.84}$$

Then it follows that

$$E[E[C^{\#}(T)]\big|_{t=0}] = E[E[C^{\#}(T)]]\big|_{t=0} \tag{2.85}$$

Which is a result that can again also be reached by invoking Fubini's theorem. To be clear, the result states that the population mean of the optimal expected unit cost ($E[C^{\#}(T)]\big|_{t=0}$) is equal to the optimal expected unit cost on the mean farm, i.e. the farm with $\lambda^{-1} = E[\lambda^{-1}]$. It therefore maps the mean of all $C^{\#}(t)$ trajectories in the farm population to the mean of possible random trajectories on a single farm (the farm where $\lambda^{-1} = E[\lambda^{-1}]$). Letting $\tilde{C}^{\#}(T)$ denote the unit cost on the mean farm (i.e. $\tilde{C}^{\#}(T) = C^{\#}(T)|_{\lambda^{-1}=E[\lambda^{-1}]}$), an equivalent, perhaps clearer, way of writing the result is

$$E[E[C^{\#}(T)]\big|_{t=0}] = E[\tilde{C}^{\#}(T)]\big|_{t=0} \tag{2.86}$$

Where the righthand side has been rewritten to more clearly indicate future expectation on the single farm where $\lambda^{-1} = E[\lambda^{-1}]$.

Switching from population parameter space to stochastic price parameter space also requires that the variance likewise switches from a population variance to a stochastic variance. This means replacing $Var[\ln(E[C^{\#}(T)]|_{t=0})]$ (i.e. $s^2_{\#}$) with $Var[\ln(\tilde{C}^{\#}(T))]|_{t=0}$.

To parse $Var[\ln(\tilde{C}^{\#}(T))]|_{t=0}$, note first that the optimal expected unit cost $E[C^{\#}(T)]\big|_{t=0}$ is an abstract optimal target that farmers aim for, but that evolves before and after the actual farm expenditure that is "locked in" at planting based on the price information $P(0)$ available at that time. As an abstract target, then, optimal unit cost may be expressed

$$C^{\#}(t) = P(t)\frac{h}{\lambda} \tag{2.87}$$

Because its evolution is proportional to output price, then, by Ito's Lemma, $C^{\#}(t)$ must also be a gBm with the same drift and volatility parameters as $P(t)$. (A proof of this is given in Appendix B.4.) Therefore,

$$Var[\ln(C^{\#}(T))]\big|_{t=0} = s_p^2 T \tag{2.88}$$

This is true for any value of λ^{-1}, including $\lambda^{-1} = E[\lambda^{-1}]$. Thus, when moving from farm population space to stochastic price space, $s_{\#}^2$ becomes $s_p^2 T$.

As for $m_{\#}$, recall that the $m_{\#}$ in the $\Phi()$ term in Equation 2.78 comes from the definition $E[C^{\#\eta}] = e^{\eta m_{\#}+\eta^2 s_{\#}^2/2}$ such that $m_{\#} = \ln(E[C^{\#\eta}]) - \eta s_{\#}^2/2$. The results in Equations 2.86 and 2.88 mean that $m_{\#}$ also has the following dual expression in stochastic price space.

$$m_{\#} = \ln(E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0}) - \eta\frac{s_p^2}{2}T \tag{2.89}$$

Substituting the dual expressions for $m_{\#}$ and $s_{\#}^2$ in the $\Phi()$ term of Equation 2.81, and substituting $P(0)e^{m_p T}$ for $E[P(T)]|_{t=0}$ (via Equation 2.75),

$$\begin{aligned}
\mathcal{P}(E[P(T)]|_{t=0}) &= \Phi\left(\frac{\ln(E[P(T)]|_{t=0}) - m_{\#}}{s_{\#}} - \eta s_{\#}\right) \\
&= \Phi\left(\frac{\ln(P(0)) + m_p T - \ln(E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0}) + \eta\frac{s_p^2}{2}T}{s_p\sqrt{T}} - \eta s_p\sqrt{T}\right) \\
&= \Phi\left(-\frac{\ln\left(\frac{E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0}}{P(0)}\right) - \left(m_p - \eta\frac{s_p^2}{2}\right)T}{s_p\sqrt{T}}\right) \\
&= 1 - \Phi(z)\ ;\quad z = \frac{\ln\left(\frac{E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0}}{P(0)}\right) - \left(m_p - \eta\frac{s_p^2}{2}\right)T}{s_p\sqrt{T}}
\end{aligned} \tag{2.90}$$

But recall from Equations 2.76 that $(m_p - \eta s_p^2/2)T$ and $s_p\sqrt{T}$ are the stochastic expectation and variance of $\ln(P(T)^{\eta})$, such that z is just the standard score of $\ln(P(T)^{\eta})$ evaluated at the logged mean farm unit expenditure $\ln(E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0})$.

$$
\begin{aligned}
z &= \frac{\ln\left(\frac{E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0}}{P(0)}\right) - \left(m_p - \eta\frac{s_p^2}{2}\right)T}{s_p\sqrt{T}} \\
&= \frac{\ln(E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0}) - \ln(P(0)) - \left(m_p - \eta\frac{s_p^2}{2}\right)T}{s_p\sqrt{T}} \\
&= \frac{\ln(E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0}) - E[\ln(P(T))]|_{t=0}}{s_p\sqrt{T}}
\end{aligned} \tag{2.91}
$$

Therefore $\Phi(z)$ is the probability that the output sale price will be less than or equal to the mean farm cost, and, conversely, $1-\Phi(z)$ is the probability that the output sale price will be higher than or equal to the mean farm cost.

$$
\mathcal{P}(E[P(T)]|_{t=0}) = 1-\Phi(z) = \mathrm{Prob}(E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0} \leq P(T)^{\eta}) \tag{2.92}
$$

Which proves Equation 2.81. Recalling from section 2.1 that the adjustment η is the output price elasticity of farm supply (PEFS), the probability $\mathcal{P}(E[P(T)^{\eta}]|_{t=0}, m_{\#}, s_{\#})$ is henceforth referred to as the "PEFS-adjusted mean probability of profit". Equation 2.90 maps the PEFS-adjusted mean probability of profit from farm population parameter space to stochastic price parameter space. The most appropriate parameter space to work in is that for which the higher resolution data is available. When farm data is available, evaluation in farm population parameter space is the best approach. In the absence of farm data, evaluation in stochastic price parameter space is appropriate if high resolution price series data is available.

Regardless of the parameter space in which calculations are made, note that Equation 2.79 may now be written

$$
\begin{aligned}
E[\tilde{y}^*(T)]|_{t=0} &= E[\tilde{y}^*(T)|\theta]|_{t=0}\frac{\mathrm{Prob}(E[C^{\#}(T)]|_{t=0} \leq P(T))}{\mathrm{Prob}(E[\tilde{C}^{\#}(T)^{\eta}]|_{t=0} \leq P(T)^{\eta})} \\
\text{Exp. uncond. mean yield} &= \text{Exp. cond. mean yield} \times \frac{\text{Farm employment rate}}{\text{PEFS-adj. mean probability of profit}}
\end{aligned} \tag{2.93}
$$

In the absence of both farm and price data, then, back of the envelope calculations can be made based on expert consensus regarding mean observed (conditional) expected yield, the farm employment rate, and the PEFS adjusted mean probability of profit.

2.5.4 Stochastically explicit expressions for price elasticity of aggregate supply, aggregate supply response, and producer welfare impact

The price elasticity of expected aggregate supply follows as

$$\begin{aligned}\frac{\partial \ln(E[Q(T)]|_{t=0})}{\partial \ln(P(0))} &= \frac{d\ln(\Phi(-z))}{d(-z)}\frac{\partial(-z)}{\partial \ln(P(0))}\\ &= \frac{\phi(-z)}{\Phi(-z)}\frac{1}{s_p\sqrt{T}}\end{aligned} \tag{2.94}$$

Whereby it is clear that, under axioms 1-5, price elasticity of expected aggregate supply changes with price, contrary to the assumption in many economic surplus analyses. It follows, moreover, that aggregate supply is elastic/inelastic if

$$s_p\sqrt{T} < / > \frac{\phi(-z)}{\Phi(-z)} \tag{2.95}$$

The stochastically explicit formulations for the research induced change in aggregate supply (Equation **??**) and producer welfare (Equation 2.67) derived in section 2.4 follow as

$$\Delta E[Q(T;k_m(T))]|_{t=0} = NE[k_a]E[\tilde{y}^*(T,k_m(T))]|_{t=0}\frac{\ell}{1-h}\mathcal{P}(\hat{C}^{*h};m_c,s_c) \tag{2.96}$$

And

$$\begin{aligned}\Delta\psi &= \gamma NE[k_a]E[\tilde{r}_\theta(T)]|_{t=0}\mathcal{A}\mathcal{E}(E[C^*|C^*\leq \hat{C}^*]-\hat{C}^*)\ ;\quad \tilde{r}_\theta = E[r|r\geq 1]\\ &= \gamma NE[k_a]E\left[\frac{P(T)}{E[C^{\#}(T)|\theta]}\right]\Big|_{t=0}\mathcal{A}\mathcal{E}(E[C^*|C^*\leq \hat{C}^*]-\hat{C}^*)\end{aligned} \tag{2.97}$$

3 An illustrative use case: Kenya's maize aggregate supply curve

3.1 Expected aggregate supply model parameter estimation and comparison to empirical estimates

As an illustrative example, the expected aggregate supply curve derived above is fit to 2007 Kenyan maize farm data and compared against empirical estimates of expected aggregate supply for the years 2006-2012.

The farm data used in this pedagogical exercise come from a panel covering the main and short rains seasons for multiple crops across Kenya for the years 1997, 2000, 2004, 2007, and 2010, originally gathered by the Egerton University Tegemeo Institute of Agricultural Policy and Development in partnership with Michigan State University, under the auspices of a USAID project (Institute 2007). For the present exercise, only the 2007 main rains season maize farm data is used. The main season in Kenya generally runs from March to August or September. Farms not reporting maize as their main crop or reporting zero maize production cost are excluded, as well as five farms reporting extremely high unit cost, resulting in a sample size of 937 maize farms across 20 districts (nowadays referred to as "counties"). The sample represents all major maize growing regions in the country. Full details regarding the data and its processing for the present exercise, including summary statistics, are provided in Appendix C.

The empirical estimates of aggregate supply for the years 2006-2012 are taken from commodity intelligence reports published in mid to late season—prior to harvest but after planted area and main season rainfall are known. These estimates, summarized in Table 1, cover some of Kenya's worst drought years in living memory.

Table 1. Empirical midseason forecasts of Kenya main season maize harvest 2006-2012

Estimate date	Estimate (million MT)	Reference	Comments
12 Jun 2006	2.7	FAO Global Information and Early Warning System 2006	Normal rainfall
28 Jun 2007	2.6	USDA Foreign Agricultural Service 2007	Normal rainfall
15 Jul 2008	2.16	FAO Global Information and Early Warning System 2008	Post-election violence, drought
11 Aug 2009	1.84	Minister of Agriculture 2009	Severe drought, crop failure
5 Aug 2010	2.6	FAO Global Information and Early Warning System 2010	Good rains that came early, but some crop loss due to flooding
July 2011	2.34	Kenya Food Security Steering Group 2011	Dismal rainfall except in west Kenya
16 Jul 2012	2.4	FAO Global Information and Early Warning System 2012	Rains started late, maize lethal necrosis outbreak

To estimate the parameters of expected aggregate supply curve, Equation 2.68 (in logs) is fit to the farm data by OLS regression.

$$\ln(y^*(T)) = \ln(v) + \sum_i^{n_k} \kappa_i \ln(k_i) + h\ln(C^*) + \epsilon \tag{3.1}$$

Where ϵ is the regression residual. Recall from the exposition in section 2.1 that the regression constant $\ln(v)$ captures the (logged) expenditure TFP, and is related to the (logged) TFP $\ln(g)$ as follows.

$$\ln(v) = \ln(a) + \ln(g) \;\; ; \;\; \ln(a) = \sum_{i}^{n_x} \ln\left(\frac{\alpha_i}{w_i}\right)^{\alpha_i} - \ln(h) \tag{3.2}$$

Where TFP may include binary endowment variables that augment the constant on farms where the endowment is present (Equation 2.7, replicated here for convenience).

$$\ln(g) = \beta_0 + \sum_{i=1}^{n_b} \beta_i b_i \;\; ; \;\; n_b = 4$$

As mentioned in section 2.1, in this exercise there are four binary endowment variables b_i (which take a value of 1 if the endowment is present and 0 otherwise) are: use of oxen in land preparation, use of tractor in land preparation, use of hybrid seed, and land tenure. There are 12 unique combinations of these four endowments in the data. The control variables (k_i) are planting density (seed kg/acre), adult equivalent household size, farm size, and rainfall. The meteorological variable is 2007 main season rainfall divided by an historical benchmark, and so can be interpreted as the rainfall anomaly or departure from the historical benchmark. The farm expenditure C^* includes the cost of synthetic fertilizer, land preparation, pest and disease control chemicals, and seed. The OLS regression results are presented in Table 2.

Table 2. 2007 Kenya farm yield regression results

Dependent variable: ln[yield (kg/acre)]	
ln[Cost/acre (KES/acre)]	0.146***
	0.019 (<0.001)
ln[Farm size (acres)]	-0.224***
	0.025 (<0.001)
ln[Planting density (kg seed/acre)]	0.521***
	0.049 (<0.001)
ln[Adult equivalent household size]	0.176***
	0.034 (<0.001)
ln[Rain anomaly]	0.151*
	0.060 (0.013)
Hybrid seed	0.379***
	0.047 (<0.001)
Tenure - owned with title deed	0.097*
	0.038 (0.012)
Landprep - oxen	0.183***
	0.050 (<0.001)
Landprep - tractor	0.171**
	0.062 (0.006)
Num.Obs.	937
R2	0.441
R2 Adj.	0.436

The fitted coefficient on logged cost per acre is statistically significant and indicates that the population yield degree of homogeneity is $h = 0.15$, by which it follows immediately that the price elasticity of farm supply is $\eta = h/(1-h) = 0.17$. The coefficients on the other regressors are also statistically significant and take values and signs that agree with intuition. Note that the coefficient on the farm size regressor is negative, affirming the inverse relation between farm size and yield found in numerous other studies (Muyanga and Jayne 2019; Larson et al. 2014; Wassie, Abate and Bernard 2019; Gautam and Ahmed 2019; Kimhi 2006).

The fitted binary endowment coefficients (β_i) indicate that hybrid maize seed is the most significant contributor to TFP, followed by land prep method and tenure. Interestingly, the oxen coefficient contributes slightly more than tractor. Based on these values, the expenditure TFPs are calculated for each of the 12 TFP strata (v_ℓ), as well as the average across all strata (v). The implicit mean expected yield ($E[\tilde{y}^*(T;k_r)|\theta]|_{t=0}$) is then calculated by Equation 2.72, for $\tilde{k}_r^{\kappa_r} = 1$ (indicating rainfall expectation in line with the historical benchmark). These parameters are reported along with the mean observed cost per unit area and harvest yield, by TFP stratum, in Figure 2.

Note, in passing, that the stratified view reveals interesting information. As one might expect, hybrid maize yields are generally higher in tractor TFP strata than they are in oxen TFP strata. But the oxen systems generally exhibit lower unit cost and higher expenditure TFP, suggesting higher cost effectiveness. And note that the implicit expected yield at planting is lower than the actual harvest yield obtained in the data year.

This makes sense because the implicit expected yield is calculated assuming rainfall in line with the historical trend; and rainfall largely surpassed the historical benchmark ($\tilde{k}_r^{\kappa_r} > 1$) in the data year (see Appendix C for a rainfall histogram).

	Expenditure TFP	Mean observed expenditure/area (KES/acre)	Mean implicit expected unit expend. (KES/kg)	Mean observed harvest yield (kg/acre)	Mean implicit expected yield (kg/acre)
Base + Hybrid + Tenure + Oxen	86.1	2666.6	2.3	1070.3	1026.1
Base + Hybrid + Tenure + Tractor	85	7461.7	5.6	1274.2	1201.2
Base + Hybrid + Oxen	78.1	2693.5	2.5	1079.7	936.2
Base + Hybrid + Tractor	77.2	7227.2	6	1199.4	1074.5
Base + Hybrid + Tenure	71.7	2511.2	2.6	1099.2	859.6
Base + Hybrid	65	2230.7	2.6	986.1	771.3
Base + Tenure + Oxen	58.9	1702.1	2.2	757.7	645.3
Base + Tenure + Tractor	58.2	4985.5	5.8	456.1	772.6
Base + Oxen	53.5	1123.7	1.7	526.3	556.5
Base + Tractor	52.8	3350	4.7	596.4	684.9
Base + Tenure	49.1	1508.5	2.4	710.5	530.4
Base	44.5	928.7	1.8	543.6	448.1
Sample	69.8	3510.3	3.5	994	849.3

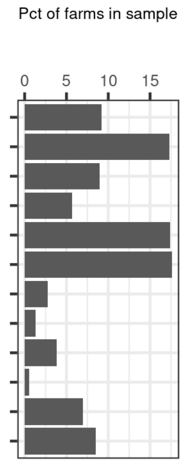

Figure 2: (Left) The expenditure TFP, observed (conditional) mean cost/area, implicit conditional mean unit expenditure, observed harvest yield, and implicit conditional expected yield for each of the 12 TFP strata present in the Kenya farm data. The implicit expected yield assumes rainfall equal to the historical benchmark. (Right) TFP strata prevalance in the sample.

To evaluate the expected unconditional mean yield term ($E[\tilde{y}^*(T)]|_{t=0}$) and parameters $m_\#$, $s_\#$ appearing in the expected aggregate supply expression (Equation 2.78), first the conditional parameters $\bar{m}_\#$ and $\bar{s}_\#$ are calculated from the fitted values v and h, and the sample parameters $\bar{m}_c$, $\bar{s}_c$, m_k, and s_k, as explained in section 2.2.3. The expected price at planting time ($E[P(T)]|_{t=0}$) for the year of the data is also required.

The Kenya farm data includes the harvest output price as observed at harvest $P(T)$, but not $E[P(T)]|_{t=0}$. Recall from Equation 2.75 that calculation of $E[P(T)]|_{t=0}$ requires the maize price at planting $P(0)$, the drift parameter m_p, and the time from planting to harvest T. In the present exercise, this information is taken from the World Food Program main season monthly price series plotted in Figure 3 (WFP 2025). Given the monthly time step of the data, the time horizon from planting in March to the August-September harvest is $T = 6$. The price at planting $P(0)$ is averaged over the main maize markets. The drift parameter m_p is then calculated as the mean log change from one month to another in all markets. Missing values are omitted from these calculations.

When calculating m_p, it is important to take note of price dynamics peculiar to the market in question. In

Kenya, the short rains maize crop is harvested about a month before the main season crop is planted. If the short rains harvest is good, then the maize price tends to remain flat during the main season or decrease a little. If the short rains harvest is poor, then the price tends to trend upwards during the main season and to be a little more volatile. This is reflected in Figure 3, where the price series in 2008, 2009, 2011, and 2012 are upward trending. It is known that the main season followed a poor short rains harvest in these years. In the other years, the price series are flat because the short rains harvest was normal. To accommodate this dichotomy, m_P is calculated separately for bad short rain years and good short rain years. For years beginning with a bad short rains harvest, the expected price is calculated using the "bad short rains" m_p, while expected price in years beginning with a good short rains harvest is calculated using the "good short rains" m_p.

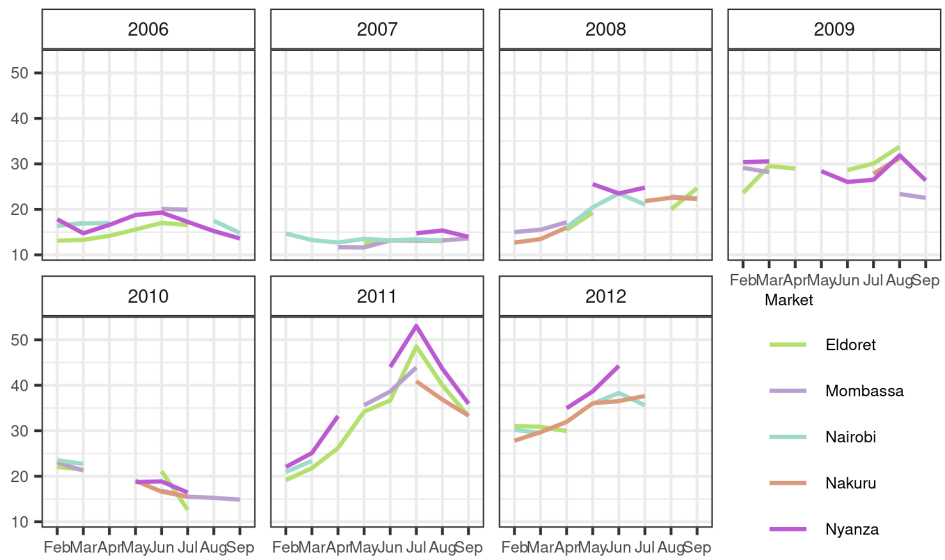

Figure 3: Maize monthly price series during main season (March-August) at primary markets in Kenya. Price data are from the World Food Program (WFP 2025).

The short rains harvest in 2007 was normal and so the "good short rains" estimate of m_p, and March 2007 price is used to compute $E[P(T)]|_{t=0} = P(0)e^{m_p T}$ (where $P(0)$ is the price at planting in March), which then feeds into the calculation of the unconditional parameters $m_\#$ and $s_\#$ by simultaneous Equations 2.45.

Once $m_\#$ and $s_\#$ are calculated, the unconditional parameters m_c, s_c are calculated by Equation 2.41, and then the expected unconditional mean yield term $(E[\tilde{y}^*(T; k_r(T))]|_{t=0})$ appearing in the expected aggregate supply function is calculated by Equation 2.72 at four different expected rainfall levels: $E[\tilde{k}_r(T)^{\kappa_r}]|_{t=0} = 1.5^{\kappa_r}, 1^{\kappa_r}, 0.5^{\kappa_r}, 0.1^{\kappa_r}$. These values correspond to expected rainfall that is 150% of the historical benchmark ("high"), in line with the historical benchmark ("normal"), 50% of the historical benchmark ("low"), and

10% of the historical benchmark ("dismal"), respectively.

The two remaining parameters required to evaluate expected aggregate supply in Equation 2.78 are number of farms N and mean farm size k_a. For mean farm size, the Kenya farm sample mean is used, giving $k_a =$ 1.85 acres. The number of maize farms is more difficult to assess. Nowadays, one sometimes sees the figure 3 million mentioned (e.g. Njeru et al. 2022), but this appears to be unsupported by hard evidence; and, to the extent that it is accurate, would imply a farm population far lower than 3 million in 2007. Here, a rough estimate is deduced by dividing planted area by mean farm size. The midseason reports examined for this exercise suggest that planted main season maize area during the period in question ranges between 3 and 4.5 million acres (KFSSG 2011). Divided by k_a, this gives a range between 1.56 and 2.34 million farms. The figure $N = 2$ million is used as a rough estimate.

With all the necessary parameters estimated, the theoretical expected aggregate supply curve developed in the sections above is plotted at the four expected rainfall levels, with empirical aggregate supply estimates overlaid, in the top panel of Figure 4. The empirical estimates for the major drought years agree closely with the theoretical expected supply curve evaluated at the "dismal" rainfall level, while the estimates in normal or good rainfall years generally agree with the theoretical curve evaluated at the "normal" and "high" rainfall levels. The curve overshoots the 2010 estimate somewhat, possibly due to crop loss from flooding that year (GIEWS 2010). The one major empirical divergence from the theoretical curve is 2009, an historic drought year (Ruto 2009).

The price elasticity of expected aggregate supply curve is presented in the bottom panel of Figure 4. Empirical estimates of the maize price elasticity of supply in Kenya generally range from 0.1 to 0.365 (Abodi, Obare and Kariuki 2021; Onono, Wawire and Ombuki 2013; Olwande, Ngigi and Nguyo 2009), which is roughly the range of elasticities indicated by the theoretical curve for the years examined here.

A key message emerging from these results is that aggregate supply varies widely with meteorological conditions. Individual supply curves represent specific scenarios within a broad range of such scenarios. It may thus more instructive to speak of a supply range or "ribbon", bounded by supply curves on top and bottom representing best and worst case meteorological scenarios.

Finally, it is important to keep in mind that the expected aggregate supply and price elasticity curves parameterized in this exercise are "short run" curves in the sense that they are valid only insofar as the values of their fitted parameters—h, N, k_a, k, $m_\#$, $s_\#$, m_p—are accurate. Empirical estimates for years well before or well after the year of the data used to estimate the parameters can be expected to diverge from the curve.

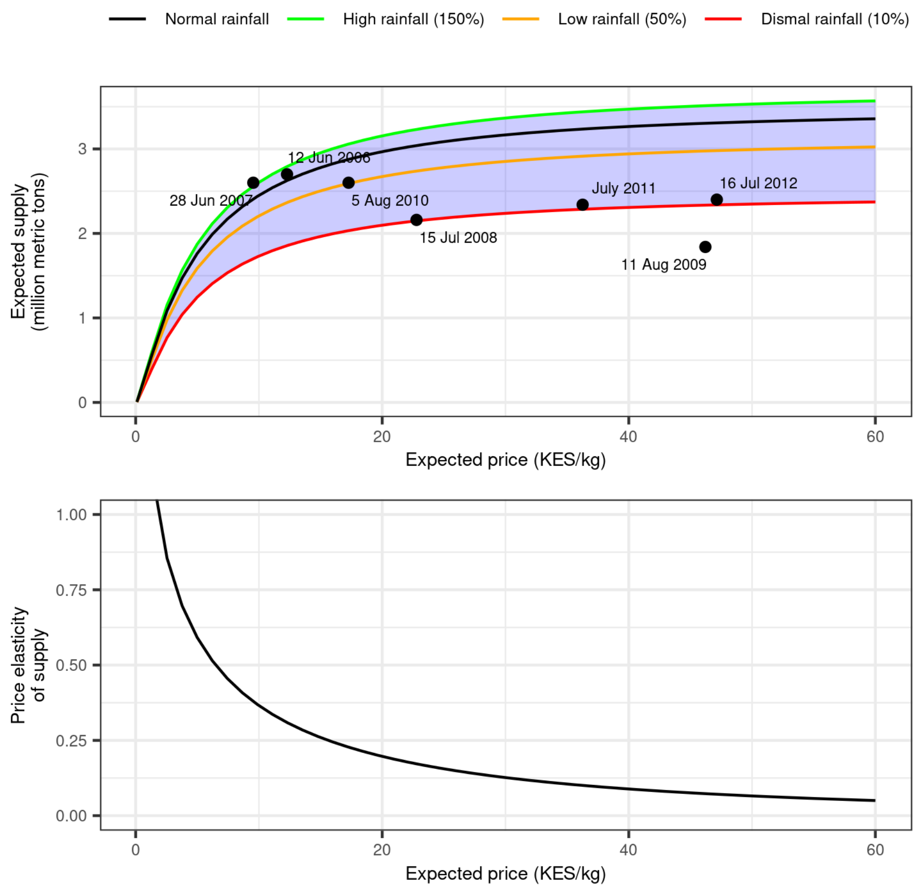

Figure 4: (Top) The theoretical expected aggregate maize supply curve fitted to 2007 Kenya long rains maize season data, plotted against empirical estimates of expected aggregate supply for the years 2006-2012. (Bottom) The corresponding theoretical price elasticity of expected aggregate supply curve.

3.2 The research induced supply ribbon shift

Say a new technology package is released in Kenya that results in a 30% increase in TFP on adopting farms. As explained in section 2.3.1, this is tantamount to adding another binary endowment $b_5\beta_5 = 0.26$ to the yield model. Say that the expected efficiency adjusted percentage increase in input prices ($\boldsymbol{\alpha}'\partial \ln(\mathbf{w})$) is 0.2, and the expected percentage decrease in output price ($\partial \ln(P)$) is -0.15, such that $\partial \ln(\nu) =$ -0.35, which is greater than the percentage increase in TFP. Adoption would thus result in a 10% increase in expenditure TFP (v), a 6% reduction in expenditure, and a 9% increase in yield on adopting farms (recall Equations 2.50). However, by the criterion in Inequality 2.53, adoption makes economic sense for the farmer only if a net subsidy is included in the technology package (i.e., only if $\partial B > 0$).

Say that the Ministry of Agriculture offers a net subsidy of $\partial B = 1000$ KES/acre to adopting farms, perhaps in the form of free seed and fertilizer. As explained in section 2.3.2, this then implies the adoption expenditure per acre ceiling of $\hat{C}^* = 2920$ KES (Equation 2.55), and the corresponding adoption rate of $\mathcal{A} = 49.1\%$ (Equation 2.56). The aggregate net subsidy required from the Ministry of Agriculture to achieve this rate follows from Equation 2.57 as 1820 million KES. The corresponding increase in expected aggregate supply and producer welfare follow from Equations 2.96 and 2.97 as 123.3 thousand metric tons and 870.5 million KES, respectively (assuming normal rainfall $\tilde{k}_r^{\kappa_r} = 1$). The breakdown by TFP stratum is given in the left panel of Figure 5, accompanied by a graphic of the corresponding supply ribbon shift in the right panel.

	Adoption rate (%)	Change in supply (1000 MT)	Change in aggregate producer welfare (million KES)
Base + Hybrid + Tenure + Oxen	42.4	12.5	79.8
Base + Hybrid + Tenure + Tractor	42.8	23.4	150.7
Base + Hybrid + Oxen	45.2	11.7	79.4
Base + Hybrid + Tractor	45.6	7.4	50.2
Base + Hybrid + Tenure	47.8	22	155.4
Base + Hybrid	50.6	21.3	157.3
Base + Tenure + Oxen	53.5	3.2	24.5
Base + Tenure + Tractor	53.8	1.5	11.3
Base + Oxen	56.3	4.2	33
Base + Tractor	56.7	0.6	4.6
Base + Tenure	58.8	7.2	57.6
Base	61.6	8.4	66.8
Population	49.1	123.3	870.5

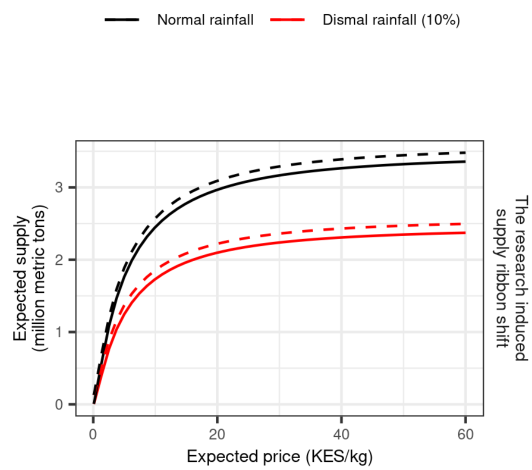

Figure 5: (Left) The adoption rate, change in welfare, and change in supply for each of the 12 TFP strata, for the technological intervention described in the text. (Right) The corresponding expected aggregate supply ribbon shift, where the dotted lines represent the new normal and dismal rainfall bounds of the aggregate supply ribbon.

For policy formation, it is helpful to consider prospective interventions in a sensitivity heatmap format, whereby multiple farm subsidy and price change scenarios can be examined at a single glance (Figure 6). These "policy decision maps" plot the adoption rate, supply change, and producer welfare outcome over a variety of different policy assumption "coordinates". In this case, the maps here are expressed in coordinates of assumed efficiency adjusted price ratio change along the x-axis, and assumed subsidy per acre along the y-axis. The map contour lines represent outcome isoquants, in the same sense that contours on a topographical map represent lines of constant elevation.

This facilitates evaluation of, for example, the sensitivity of expected adoption rate/supply/welfare outcomes to inaccurate assumptions regarding the price shock associated with adoption. Policy makers can then also evaluate, at a glance, the additional subsidy required to stay on the same outcome isoquant in the event of such inaccuracies. Additional maps may be made with one or both policy assumptions swapped out for other assumptions, such as assumed rainfall level or price volatility.

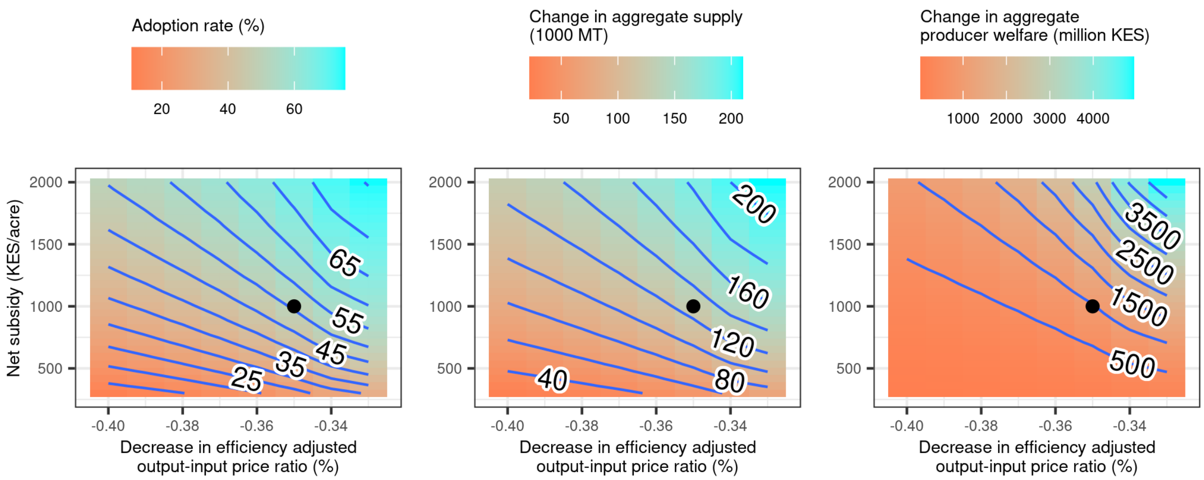

Figure 6: Kenya maize farm subsidy policy heat maps of the (left to right) adoption rate, change in aggregate supply, and aggregate producer welfare given release of the technology package described in the text. Each 'coordinate' on the map represents an assumed percentage decrease in the efficiency adjusted output-input price ratio and a net subsidy per acre. The scenario presented in the previous figure above is marked with a black dot. In these maps, the ratio of change in input price to change in output price is fixed at 0.75.

The illustrative example presented above is an exercise in comparative statics. That is to say, it is a comparison of the "before new tech" (or status quo) and "after new tech" situations, without accounting for the time it takes for release and adoption to occur. In a standard impact assessment, the static analysis would be converted into a temporally explicit assessment by introducing a sigmoidal adoption rate curve over time, as explained in Alston et al. (1995) and elsewhere. The welfare impact would be then be calculated as the sum of discounted yearly net impacts out to a relevant cutoff year determined by the policy context.

4 Discussion and conclusion

A model of agricultural supply and of the producer side impact of new technology (or any exogenous shock) is useful to agricultural policy makers insofar as it can map out the plausible outcome space for the range of price assumptions and policy lever settings (subsidies) relevant to the policy discussion; and insofar as this mapping is transparently rooted in an empirically vetted, sensible theoretical foundation.

More specifically, this means that a useful model of aggregate supply must explicitly account for 1) the shape of the yield distribution, and hence also provide an endogenous accounting of farm entry and exit in response to price changes; 2) the endogenous farm level supply response to new technology and to any associated exogenous changes in input/output prices and subsidies; and 3) the aggregation of farm level response and welfare impacts up to the population level—i.e. the supply curve shift—in a theoretically sensible and transparent way, including endogenous modeling of the adoption rate.

As mentioned in the Introduction, the conventional approach to modeling aggregate agricultural supply and technology shocks provides no such explicit, theoretically grounded, endogenous accounting. Instead, the shape of the distribution is unceremoniously ignored (or identically endowed marginal firms are assumed), the supply response and adoption rate are assumed as exogenous inputs into the model, essentially synonymous with the technology shock itself; and the aggregation of farm level impacts up to the population level is neither empirically nor theoretically grounded, but rather introduced as a matter of fiat. As a result, the model is ill suited to answer such basic policy questions such as:

- Upon what input price assumptions is the assessment based? What is the assumed rainfall level? What is the assumed subsidy level?
- By how much will the adoption rate/producer welfare/aggregate supply change if subsidies are increased/lowered?
- By how much will the adoption rate/producer welfare/aggregate supply change if the change in TFP turns out to be 10 percentage points higher/lower than expected?
- By how much will the adoption rate/producer welfare/aggregate supply change if the change in input prices is 10 percentage points higher/lower than expected?
- Given a certain change in TFP and input/output prices, what subsidy level is required to push adoption over 50%?

To more squarely meet the needs of agricultural policy makers, here an alternative modeling framework was theoretically derived from five non-controversial axioms and empirically vetted against Kenya farm data. The proposed model can answer the usual set of policy questions, such as those just enumerated, as well as many others not often considered part of the purview of such models. In particular, the model explicitly controls for meteorological conditions and price volatility, and disaggregates results by TFP stratum. The model can thus be used to examine highly nuanced scenarios regarding technology packages tailored to specific farming systems under different weather and price volatility settings.

For example, on untenured farms, the technology package could be bundled with legal services to formalize a title deed, such that TFP on such farms is augmented not just by the new techology coefficient (β_{new}), but also by the tenure coefficient $\beta_{new} + \beta_{tenure}$. Policy makers can use the model to weight the benefits of such a strategy against the additional cost in legal services, under different output/input price, subsidy, and rainfall scenarios. Or, if the technology shock is a new seed variety, then the percentage change in TFP ($\partial \ln(g)$) on farms growing hybrid maize will be considerably less than on non-hybrid farms, as the new technology in such cases does not augment but rather replaces the existing seed productivity. The model would explicitly accommodate this nuance by replacing, as opposed to augmenting, the hybrid maize binary endowment coefficient β_{hybrid} with the new seed coefficient.

The Kenya example focuses on a hypothetical SAI technology shock; but the model is equally applicable to AEI shocks that typically prioritize resilience over genetic gain. This may include assessment of, for example, agroclimatic advisory services, risk mitigating cooperative social mechanisms, payment for economic services arrangements, and combinations thereof. Because the model explicitly accounts for the heterogeneous endowments and expenditure, which in turn are reflections of income distribution, then the supply and welfare impacts of changing income inequality, and related distributional questions, may also be examined.

The proposed model is adaptive to the available data. The illustrative example above represents an ideal case where good farm survey data is available. Under less auspicious circumstances, model parameters can still be estimated up to a respectable degree of precision because the model is expressed in terms of economic indicators and population parameters (the farm employment rate, PEFS-adjusted probability of profit, observed mean and standard deviation of yield and expenditure) that in many cases are published or readily deductible from available information. As explained in section 2.5, model parameters can also be estimated based on output price series, if available.

The proposed model is limited to the evaluation of aggregate supply and producer side welfare impacts. However, it would be straightforward to extend the model to include aggregate demand and consumer side impacts—thereby converting it into a full fledged partial equilibrium model—by pairing the derivation above with a theoretical derivation of the demand curve from the same set of axioms. While this lies outside the scope of the present effort, one may reasonably anticipate that the demand side derivation will involve reinterpretation of the input demand curve derived in section 2.1 as a processing plant (e.g. miller, starch factory) feedstock demand curve; and the aggregation of this curve over the processor population density. In many use cases, the number of processors may be quite small, resulting in a cartel or even monopsony demand curve, especially at subnational scales of analysis. The axiomatic approach would explicitly capture this and other demand side nuance, which is often glossed over in the conventional treatment. Once aggregate demand is included, the equilibrium output price becomes endogenous to the model, and consumer welfare outcomes can be evaluated. In many cases where the producer welfare alone does not offset the total subsidy cost to government (as in the Kenya example above), the addition of consumer side welfare would likely push the total welfare above the subsidy cost by a considerable margin. [can also address questions of how the same result might be achieved through interventions of other variables (Ethiopia bean an example), so as to compare which impact pathway is cheapest]

References

Abodi, M.A., G.A. Obare, and I.M. Kariuki. 2021. "Supply and demand responsiveness to maize price changes in kenya: An application of error correction autoregressive distributed lag approach." *Cogent Food & Agriculture* 7(1):1957318.

Adato, M., and R. Meinzen-Dick. 2007. *Agricultural research, livelihoods, and poverty: Studies of economic and social impacts in six countries.* Intl Food Policy Res Inst.

Alston, J., M. Marra, P. Pardey, and P. Wyatt. 2000. "Research return redux: A meta-analysis and the returns of r&d." *Australian Journal of Agricultural Economics* 44:1364–85.

Alston, J.M., G.W. Norton, P.G. Pardey, and others. 1995. *Science under scarcity: Principles and practice for agricultural research evaluation and priority setting.* Cornell University Press.

Alston, J.M., and P.G. Pardey. 2001. "Attribution and other problems in assessing the returns to agricultural r&d." *Agricultural economics* 25(2-3):141–152.

Alston, J.M., P.G. Pardey, J.S. James, and M.A. Andersen. 2009. "The economics of agricultural r&d." *Annu. Rev. Resour. Econ.* 1(1):537–566.

Alston, J.M., and M.K. Wohlgenant. 1990. "Measuring research benefits using linear elasticity equilibrium displacement models." In J. M. Alston and J. D. Mullen, eds. *The returns to the australian wool industry from investment in r&d, rural and resource economics report no. 10.* Sydney: New South Wales Agriculture; Fisheries.

Balasubramanian, V., M. Bell, and M. Sombilla. 2000. "Yield, profit, and knowledge gaps in rice farming:

Causes and development of mitigation measures." International Rice Research Institute (IRRI).

Bank, W. 2007. *World development report 2008: Agriculture for development*. The World Bank.

Berg, A. G., and J.D. Ostry. 2017. "Inequality and unsustainable growth: Two sides of the same coin?" *IMF Economic Review* 65(4):792–815.

Berg, A., J.D. Ostry, C.G. Tsangarides, and Y. Yakhshilikov. 2018. "Redistribution, inequality, and growth: New evidence." *Journal of Economic Growth* 23:259–305.

Berg, A., J.D. Ostry, and J. Zettelmeyer. 2012. "What makes growth sustained?" *Journal of Development Economics* 98(2):149–166.

Bernoulli, D. 1954. "Exposition of a new theory on the measurement of risk." *Econometrica* 22(1):23–36.

Borlaug, N.E. 1970. "The green revolution, peace, and humanity." Available at: https://www.nobelprize.org/prizes/peace/1970/borlaug/lecture/.

Bourguignon, F., and C. Morrisson. 1998. "Inequality and development: The role of dualism." *Journal of development economics* 57(2):233–257.

Brester, G.W., J.A. Atwood, and M.A. Boland. 2023. *Equilibrium displacement models: Theory, applications, & policy analysis*. University of Minnesota Libraries Publishing.

Byerlee, D., A. De Janvry, and E. Sadoulet. 2009. "Agriculture for development: Toward a new paradigm."

Capalbo, S.M., and J.M. Antle. 2015. *Agricultural productivity: Measurement and explanation*. Routledge.

CME. 2025. "Daily US maize futures price, Commodities and Mercantile Exchange (CME) Group, Chicago Board of Trade."

Easterly, W. 2007. "Inequality does cause underdevelopment: Insights from a new instrument." *Journal of development economics* 84(2):755–776.

Edwards, G.W., and J.P. Voon. 1997. "The calculation of research benefits with linear and nonlinear specifications of demand and supply: reply." *American Journal of Agricultural Economics* 79(4):1368–1371.

Elbasha, E.H. 1997. "The calculation of research benefits with linear and nonlinear specifications of demand and supply: comment." *American Journal of Agricultural Economics* 79(4):1363–1367.

Fan, S., C. Chan-Kang, and A. Mukherjee. 2005. "Rural and urban dynamics and poverty: Evidence from china and india."

Feleke, S., V.M. Manyong, T. Abdoulaye, and A. Alene. 2016. "Assessing the impacts of cassava technology on poverty reduction in Africa." *Studies in Agricultural Economics* 118(2):101–111.

Gallup, J.L., S. Radelet, and A. Warner. 1998. "Economic growth and the income of the poor." *Manuscript, Harvard Institute for International Development.*

Gautam, M., and M. Ahmed. 2019. "Too small to be beautiful? The farm size and productivity relationship in bangladesh." *Food Policy* 84:165–175.

GIEWS. 2010. "GIEWS country brief: kenya." Food & Agriculture Organization (FAO) Global Information; Early Warning System (GIEWS).

GIEWS. 2012. "GIEWS country brief: kenya." Food & Agriculture Organization (FAO) Global Information; Early Warning System (GIEWS).

GIEWS. 2006. "GIEWS updates 2006." Food & Agriculture Organization (FAO) Global Information; Early Warning System (GIEWS).

GIEWS. 2008. "GIEWS updates 2008." Food & Agriculture Organization (FAO) Global Information; Early Warning System (GIEWS).

Gollin, D., C.W. Hansen, and A.M. Wingender. 2021. "Two blades of grass: The impact of the green revolution." *Journal of Political Economy* 129(8):2344–2384.

Hazell, P., and L.J. Haddad. 2001. *Agricultural research and poverty reduction*. Intl Food Policy Res Inst.

Herdt, R.W. 1988. "Increasing crop yields in developing countries."

Hicks, J. 1963. *The theory of wages*. Springer.

Hull, J.C. 2015. *Options,futures, and other derivatives* Ninth. Pearson.

Hurley, T.M., P.G. Pardey, X. Rao, and R.S. Andrade. 2016. "Returns to food and agricultural r&d; investments worldwide, 1958-2015."

Hurley, T.M., X. Rao, and P.G. Pardey. 2014. "Re-examining the reported rates of return to food and agricultural research and development." *American Journal of Agricultural Economics* 96(5):1492–1504.

Hurley, T.M., X. Rao, and P.G. Pardey. 2017. "Re-examining the reported rates of return to food and agricultural research and development: reply." *American Journal of Agricultural Economics* 99(3):827–836.

Institute, T. 2007. "Household survey 2007."

Janvry, A. de, and E. Sadoulet. 2022. "Agriculture for development: Analytics and action." *Annual Review of Resource Economics* 14(1):1–16.

Janvry, A. de, and E. Sadoulet. 2020. "Using agriculture for development: Supply-and demand-side approaches." *World development* 133:105003.

JSE. 2025. "Daily white maize spot price, Johannesburg Stock Exchange (JSE)."

Kauder, E. 1953. "Genesis of the marginal utility theory: From aristotle to the end of the eighteenth century." *The Economic Journal* 63(251):638–650.

KFSSG. 2011. "The 2011 long rains season assessment report." Government of Kenya, Kenya Food Security Steering Group (KFSSG).

Kimhi, A. 2006. "Plot size and maize productivity in zambia: Is there an inverse relationship?" *Agricultural Economics* 35(1):1–9.

Kydd, J., A. Dorward*, J. Morrison, and G. Cadisch. 2004. "Agricultural development and pro-poor economic growth in sub-saharan africa: Potential and policy." *Oxford Development Studies* 32(1):37–57.

Larson, D.F., K. Otsuka, T. Matsumoto, and T. Kilic. 2014. "Should african rural development strategies depend on smallholder farms? An exploration of the inverse-productivity hypothesis." *Agricultural Economics* 45(3):355–367.

Lavoie, M. 2014. *Post-keynesian economics: New foundations*. Edward Elgar Publishing.

Lindner, R.K., and F. Jarrett. 1978. "Supply shifts and the size of research benefits." *American Journal of Agricultural Economics* 60(1):48–58.

Lipton, M. 2005. *The family farm in a globalizing world: The role of crop science in alleviating poverty*. Intl Food Policy Res Inst.

Martin, W., and J.M. Alston. 1997. "Producer surplus without apology? Evaluating investments in RD." *Economic Record* 73(221):146–158.

Mathur, S., D. Pachico, and A.L. Jones. 2003. *Agricultural research and poverty reduction: Some issues and evidence*. CIAT.

Mockshell, J., and J. Kamanda. 2018. "Beyond the agroecological and sustainable agricultural intensification debate: Is blended sustainability the way forward?" *International Journal of Agricultural Sustainability* 16(2):127–149.

Muyanga, M., and T.S. Jayne. 2019. "Revisiting the farm size-productivity relationship based on a relatively wide range of farm sizes: Evidence from kenya." *American Journal of Agricultural Economics* 101(4):1140–1163.

Njeru, F., S. Mwaura, P.M. Kusolwa, and G. Misinzo. 2022. "Maize production systems, farmers' perception and current status of maize lethal necrosis in selected counties in kenya." *All Life* 15(1):692–705.

Norton, G.W., and J.S. Davis. 1981. "Evaluating returns to agricultural research: A review." *American Journal of Agricultural Economics* 63(4):685–699.

Norton, G.W., P.G. Pardey, and J.M. Alston. 1992. "Economic issues in agricultural research priority setting." *American Journal of Agricultural Economics* 74(5):1089–1094.

Oehmke, J.F. 2017. "Re-examining the reported rates of return to food and agricultural research and development: comment." *American Journal of Agricultural Economics* 99(3):818–826.

Olwande, J., M. Ngigi, and W. Nguyo. 2009. "Supply responsiveness of maize farmers in kenya: A farm-level analysis."

Onono, P.A., N.W. Wawire, and C. Ombuki. 2013. "The response of maize production in kenya to economic incentives."

Ostry, J.D. 2015. "Inequality and the duration of growth." *European Journal of Economics and Economic Policies* 12(2):147–157.

Ostry, J.D., A. Berg, and S. Kothari. 2021. "Growth-equity trade-offs in structural reforms." *Scottish Journal of Political Economy* 68(2):209–237.

Partridge, M.D. 1997. "Is inequality harmful for growth? comment." *The American Economic Review* 87(5):1019–1032.

Patel, R. 2013. "The long green revolution." *The Journal of Peasant Studies* 40(1):1–63.

Pingali, P.L. 2012. "Green revolution: Impacts, limits, and the path ahead." *Proceedings of the national academy of sciences* 109(31):12302–12308.

Pingali, P.L., and M. Hossain. 1998. *Impact of rice research.* Int. Rice Res. Inst.

Ravallion, M. 1999. *When is gowth pro-poor?* World Bank Publications.

Ravallion, M., and G. Datt. 1996. “How important to india’s poor is the sectoral composition of economic growth?” *The World Bank Economic Review* 10(1):1–25.

Ravallion, M., and G. Datt. 2002. “Why has economic growth been more pro-poor in some states of india than others?” *Journal of development economics* 68(2):381–400.

Redclift, M. 1983. “Production programs for small farmers: Plan puebla as myth and reality.” *Economic Development and Cultural Change* 31(3):551–570.

Rose, R.N. 1980. “Supply shifts and research benefits: comment.” *American Journal of Agricultural Economics* 62(4):834–837.

Rosegrant, M.W., and P.B. Hazell. 2001. “Transforming the rural asian economy: The unfinished revolution.”

Ruto, W. 2009. “Food security situation, kenya ministry of agriculture.”

Scobie, G.M., J.D. Mullen, and J.M. Alston. 1991. “The returns to investment in research on australian wool production.” *Australian Journal of Agricultural Economics* 35(2):179–195.

Solow, R.M. 1957. “Technical change and the aggregate production function.” *The review of Economics and Statistics* 39(3):312–320.

Sumberg, J. 2012. “Mind the (yield) gap (s).” *Food Security* 4:509–518.

Thirtle, C., L. Lin, and J. Piesse. 2003. “The impact of research-led agricultural productivity growth on poverty reduction in africa, asia and latin america.” *World Development* 31(12):1959–1975.

Thorbecke, E., and H.-S. Jung. 1996. “A multiplier decomposition method to analyze poverty alleviation.” *Journal of Development Economics* 48(2):279–300.

Timmer, C.P. 1997. *How well do the poor connect to the growth process?* Harvard Institute for International

Development.

USDA. 2007. “Kenya corn update report.” U.S. Dept. of Agriculture (USDA) Foreign Agricultural Service.

Valdés, A., and W. Foster. 2010. “Reflections on the role of agriculture in pro-poor growth.” *World Development* 38(10):1362–1374.

Voon, J.P., and G.W. Edwards. 1991. “The calculation of research benefits with linear and nonlinear specifications of demand and supply functions.” *American Journal of Agricultural Economics* 73(2):415–420.

Wassie, S.B., G.T. Abate, and T. Bernard. 2019. “Revisiting farm size-productivity relationship: New empirical evidence from ethiopia.” *Agrekon* 58(2):180–199.

WFP. 2025. “Price of maize per kg in kenyan markets, World Food Program (WFP).”

Wohlgenant, M.K. 1997. “The nature of the research-induced supply shift.” *Australian Journal of Agricultural and Resource Economics* 41(3):385–400.

Zhao, X., J.D. Mullen, and G.R. Griffith. 1997. “Functional forms, exogenous shifts, and economic surplus changes.” *American journal of agricultural economics* 79(4):1243–1251.

Appendix A: Literature review

Is impact assessment insensitive to misspecification of the supply curve?

When pressed on the matter, proponents of the conventional aggregate supply functions make little attempt to defend their specifications on empirical or theoretical grounds. On the contrary, proponents of linear models generally admit that real world commodity supply and demand curves are almost certainly never linear (Brester, Atwood and Boland 2023). Proponents of the CES model generally argue that CES models are probably at least slightly more accurate than linear specifications (Voon and Edwards 1991), but likewise offer no formal derivation of their model from theory. Rather, the defense of conventional models centers on the argument that the functional form is irrelevant for purposes of impact assessment.

> "The pertinent question is whether the functional form used is an adequate approximation for the purpose. It turns out, empirically, that measures of total research benefits and their distribution between producers and consumers are quite insensitive to choices of functional form. They are much more sensitive to the related but separate choices concerning the nature of the research-induced supply curve shift and elasticities" (Alston et al. 1995).

Alston et al. (1995) cite three peer reviewed articles in support of this claim. One of the cited articles actually runs contrary to the claim, explicitly concluding that impact assessment results are highly sensitive to choices of functional form (Voon and Edwards 1991)—a conclusion subsequently reaffirmed in a comment (Elbasha 1997) and reply (Edwards and Voon 1997). Another cited article, by Lindner and Jarrett (1978), concludes that impact assessment results are sensitive to the nature of the supply curve shift (pivot versus parallel shift); but also acknowledges that these shifts are determined by the choice of functional form. In their words,

> "[P]ast studies have made no attempt to justify or explain the particular assumptions made about the type of supply shift. In fact, in some studies the impression is given that the type of supply shift was determined by the specification of the mathematical form of the supply curve, which in turn appears to have been chosen simply for computational convenience" (Lindner and Jarrett 1978).

By the transitive property, then, their conclusion regarding impact assessment sensitivity to the type of supply curve shift extends to functional form, again contrary to the claim that functional form is irrelevant. Similarly, the third cited article, by Norton and Davis (1981), asserts that functional form is "less important than the type and magnitude of the supply shift". But then, contrary to this assertion, the authors proceed to survey four examples where the mathematical form of the shift is clearly prefigured by that of the supply curve; and explicitly state that differences between the four examples in the assessment of net benefits are due to differences in both type/magnitude of shift and functional form. Alston et al. (1995) themselves point out that the practitioner is at liberty to choose between different types of shift only in the case of the linear specification—where the choice is between a pivot and a parallel shift, or some combination thereof. In the case of the CES specification, the practitioner is constricted to assume a pivot (a point later emphasized by Voon and Edwards (1991)). And elsewhere Alston writes that "the choice [of type of supply shift is] dictated by analytic convenience, given *a prior* choice of functional form for supply" (Martin and Alston 1997).

A fourth document is also cited (Alston and Wohlgenant 1990), but is not peer reviewed and not, at the time of writing, readily available from the publisher. A subsequent peer reviewed article by Zhao et al. (1997) cites the unavailable document as an annex to a technical report on R&D in the Australian wool industry. The main text of this report appears to have been published in a peer reviewed journal, but without the annex in question (Scobie, Mullen and Alston 1991). In their own analysis, Zhao et al. (1997) find that assessment of net benefits (economic surplus) is highly sensitive to misspecification of the functional form in the case of the CES specification, again contrary to the claim that impact assessments are insensitive to misspecification of the supply curve.

They also find that impact assessment is insensitive to misspecification error when 1) the linear specification is assumed together with 2) a parallel shift, and assuming 3) the shift is small (Zhao et al. 1997). Proponents and critics alike acknowledge that the empirical and theoretical grounds for assumptions 1 and 2 are narrow, at best (Lindner and Jarrett 1978; Alston et al. 2009; Voon and Edwards 1991; Rose 1980). If the real shift is a pivot or mixture of pivot and shift, then Zhao et al.'s finding of insensitivity does not apply. Moreover, it is questionable whether the parallel supply curve shifts assumed in most economic surplus assessments of returns to agricultural R&D qualify as "small" (Hurley, Rao and Pardey 2014; Alston et al. 2000). To know for sure, Zhao et al. (1997) develop a method to calculate the upper bound on misspecification bias; but the present literature review finds no instance where this method has actually been implemented. Finally, Zhao et al. (1997) acknowledge that their findings regarding linear misspecification are valid only insofar as farm populations consist entirely of identical marginal farms, a condition never actually observed in any real farm population.

In synthesis, then, it is misleading to characterize the cited literature as supportive of Alston et al.'s (1995) broad claim that impact assessments are insensitive to misspecification of the functional form of the supply curve. More accurately, the literature supports the much narrower claim that impact assessments are insensitive to misspecification when the linear specification is assumed together with a parallel shift, and when the real shift is also parallel; and when the shift is small, and when the farm population consists of identical marginal farms. Apart from this exceptionally implausible case identified by Zhao et al. (1997), the literature indicates that impact assessments are highly sensitive to functional form, whether directly, as in the case of the CES function, or by the transitive property, i.e., through a supply curve shift which is mathematically prefigured by the functional form of the supply curve.

Is it impossible to do any better?

As a final line of defense, proponents of conventional aggregate supply curves argue that there is no alternative to their cavalier approach of arbitrarily assuming a functional form for supply curves and supply curve shifts, however implausible or grossly inaccurate they might be. Alston et al. (1995) write:

> "What does economic theory tell us about the nature of . . . [research induced aggregate supply curve] shifts? Unfortunately, not very much. To be confident about this aspect of the problem would require either (a) precise econometric evidence or (b) detailed information on the effects on individual agents, details of industry structure including details on exit and entry of firms, and a complete theory of aggregation. This information is not available; assumptions are unavoidable".

The present work is effectively one long disagreement with this passage. The sweeping claim that economic theory does not tell us "very much" about the nature of aggregate supply curve shifts is especially perplexing. On the contrary, economic theory provides a set of microeconomic principles whence a firm level yield envelope is derived, whence "effects on individual agents", i.e., farm level response to exogenous shock, can be can modeled. Insofar as the phrase "complete theory of aggregation" refers to the specification of aggregate supply and demand curves, this too can be derived from theory by integrating individual effects over a suitable density function (i.e., Equation 1.1). The firm yield model implies bounds on the domain of optimal output, such that the "exit and entry of firms" is accounted for by integrating over the domain of optimal behavior.

To be clear, the "precise econometric evidence or. . . detailed information" that the authors refer to is, indeed, generally not available; but the suggestion that economic theory offers nothing in its stead is mistaken. One of the key purposes of having a theory is to draw inference when detailed information is unavailable or otherwise provides no clear answer to our questions. And, yes, "assumptions are unavoidable"—but in the sense that axioms 1-5 in the Introduction are unavoidable, not in the sense of superimposing an exogenously assumed supply response on top of said axioms. Not economic theory, then, but rather a descriptive mathematical formulation which, by construction, precludes endogenous modeling of the aggregate supply response to exogenous shocks, is what tells us "not very much" about the the aggregate supply response to exogenous shocks.

The limited policy relevance of conventional forms

Effective policy formation requires careful assessment not only of the magnitude of net benefits of a given intervention, but also of their distribution across producers and consumers (the "food price dilemma"), and across income levels. Proponents and critics alike acknowledge that conventional functional forms are ill suited to the assessment of such distributional issues (Alston et al. 2009; Alston et al. 1995; Lindner and Jarrett 1978; Rose 1980; Wohlgenant 1997).

More specifically, impact assessments based on a pivot of the linear or CES supply curve can result in consumer-producer tradeoffs, whereas assessments based on a linear parallel shift of the linear supply curve guarantee increases in both consumer and producer welfare (Alston et al. 2009). With no clear theoretical or empirical premise upon which to recommend one set of results over another, little confidence can be attached to any of them. This is remarkable considering that the primary motive for conventional supply curve specifications is to facilitate the assessment of changes in consumer and producer surplus. It is also noteworthy that, of all the possible arbitrary configurations of supply curve and type of shift, most agricultural R&D impact assessments assume the one that, by construction, results in the highest net benefits and avoids any food-price dilemma (Hurley et al. 2014; Alston et al. 2000; Alston et al. 2009).

As for the assessment of impacts on income inequality, conventional models are ill suited because, as noted earlier, they either do not explicitly account for heterogeneous farm endowments (in the case of the CES curve), or assume a population of identically endowed marginal farms (in the case of the linear specification). In some sectors, one might reasonably justify the assumption of homogeneously endowed firms by invoking competitive long run equilibrium. Such arguments are strained in the primary sector, however, where one typically sees an abundance of heterogeneously endowed, inframarginal farms in operation. By failing to account for such heterogeneity, the conventional models fail to account for farm entry/exit. (See Wohlgenant (1997) for a deeper treatment of this point.)

Proponents of the conventional approach even go so far as to expressly discourage the assessment of distributional equity vis-a-vis agricultural research, arguing that such issues lie beyond its scope.

> "For many economists. . . the only defensible justification for government intervention in agricultural research, and the only legitimate and achievable objective of research, is the pursuit of economic efficiency. This is not to say that other objectives (such as. . . income distribution and food security) are illegitimate, irrelevant, or unimportant but that there are likely to be alternative, less costly ways to achieve these other objectives than by biasing the agricultural research portfolio away from programs that will maximize total national net income" (Alston et al. 1995).

An underlying assumption here is that distributional equity and economic growth ("efficiency") are mutually exlusive objectives, such that the one cannot be pursued except at the cost of the other. Norton (1992) makes this explicit by actually drawing an equity-efficiency tradeoff frontier to illustrate the concept. Again citing “many economists”, he affirms that “agricultural research is a relatively poor instrument for achieving non-efficiency objectives”.

The authors provide no references to indicate who they mean by “many economists”; and it is not readily obvious who they might have in mind. On the contrary, many economists think and/or have demonstrated econometrically that the exact opposite is true: agricultural research generally results in agricultural productivity growth, which goes hand in hand with poverty reduction (Rosegrant and Hazell 2001; Ravallion and Datt 1996; Thorbecke and Jung 1996; Fan, Chan-Kang and Mukherjee 2005; Bourguignon and Morrisson 1998; Timmer 1997; Gallup, Radelet and Warner 1998; Ravallion 1999; Ravallion and Datt 2002; Byerlee, De Janvry and Sadoulet 2009; Valdés and Foster 2010; Kydd et al. 2004; Bank 2007; Janvry and Sadoulet 2020; Janvry and Sadoulet 2022; Gollin, Hansen and Wingender 2021). Publicly funded agricultural research since the 1960s has generally endeavored to align with pro-poor growth objectives, with mixed success (Thirtle, Lin and Piesse 2003; Adato and Meinzen-Dick 2007; Hazell and Haddad 2001; Lipton 2005; Redclift 1983; Feleke et al. 2016; Pingali and Hossain 1998; Mathur, Pachico and Jones 2003). Moreover, it is now well documented that a nation’s economic growth is generally more robust and sustainable in the measure that it is equitable (Partridge 1997; Easterly 2007; Berg, Ostry and Zettelmeyer 2012; Ostry 2015; Berg and Ostry 2017; Berg et al. 2018; Ostry, Berg and Kothari 2021). Equity and sustainable growth are not mutually exclusive, but rather “two sides of the same coin” (Berg and Ostry 2017).

Finally, conventional models also introduce awkward artifacts that further undermine their explanatory power. The CES function, by definition, implies that yield and entry/exit responses to price changes always offset each other precisely such that the elasticity of supply is constant. The linear aggregate supply curve implies the possibility of negative supply quantities and/or of farms producing for free. This either requires redefinition of the curve as a piecewise function—which further complicates interpretation in terms of Equation 1.1—or compels us to regard the function as a local approximation to some inscrutable global function, with validity restricted to a small domain of convergence around the starting equilibrium point.

Real consequences

There is a growing “disconnect” between the actual level of funding for agricultural research and the level of funding that would be justified by impact assessments conducted on the basis of conventional aggregate supply models (Hurley et al. 2016). And there is good reason to suspect that this disconnect is due to the implausibly high—and sometimes absurdly high—internal rates of return (IRR) reported in conventional impact assessments of agricultural R&D (Hurley et al. 2014; Alston and Pardey 2001; Hurley et al. 2016). The high reported returns are, of course, made more conspicuous by the aforementioned conflict of interest inherent in the expert opinion upon which these studies are typically based. In a meta-analysis of 492 such impact assessments, for example, Hurley et al. (Hurley et al. 2016) note that the average reported IRR implies that the USDA's 2000 R&D investment of $4.1 billion should be worth $56.3 quintillion in 2050, or 2.3 million times projected world GDP at that time. "[I]t is not difficult to see," the authors conclude, "how policy makers may question the credibility of such evidence" (Hurley et al. 2016).

Hurley et al. (Hurley et al. 2014) respond to the incredulity of policy makers by replacing IRR with their "modified" IRR or "MIRR", whereby implausibly high projected returns are deflated to plausible levels. This is a suitable response insofar as the implausibility in question is rooted in the discounting method (although see the debate between Ohemke (2017) and Hurley et al. (2017)). However, the theoretical opacity, dubious implications, and limited interpretability of these models, outlined above, are good reason to suspect that much of the incredulity is in fact aimed at the impact modeling itself, particularly the modeling of aggregate supply and the supply response to exogenous shocks. Insofar as this is true, MIRR is an insufficient response, and may even exacerbate the problem by masking the real sources of inaccuracy.

Appendix B

B.1 Price and cross price elsaticities of input demand

Decreasing returns to scale imply strictly downward sloping input demand curves. To see this, consider the derivative of input demand (Equation 2.12) with respect to input own price.

$$\frac{\partial \ln(x_i^*)}{\partial \ln(w_i)} = \frac{\partial \ln(y^*)}{\partial \ln(w_i)} - 1 \tag{B.1}$$

To resolve the yield derivative, note that a more conventional textbook form of y^* can be written by combining Equations 2.6 and 2.12.

$$
\begin{aligned}
y^* = y|_{x_i=x_i^*} &= g\prod_{i=1}^{n_k} k_i^{\kappa_i}\prod_{i=1}^{n_x} x_i^{*\alpha_i} \\
&= g\prod_{i=1}^{n_k} k_i^{\kappa_i}\prod_{i=1}^{n_x}\left(\frac{y^* P\alpha_i}{\lambda w_i}\right)^{\alpha_i} \\
\rightarrow y^* &= \left(g\prod_{i=1}^{n_k} k_i^{\kappa_i}\prod_{i=1}^{n_x}\frac{\alpha_i}{w_i}\right)^{1/(1-h)}\left(\frac{P}{\lambda}\right)^{h/(1-h)}
\end{aligned}
\tag{B.2}
$$

Hence the yield derivative with respect to input price works out to

$$
\frac{\partial \ln(y^*)}{\partial \ln(w_i)} = -\frac{\alpha_i}{1-h} \tag{B.3}
$$

And so the input own and cross price elasticity resolve, respectively,

$$
\frac{\partial \ln(x_i^*)}{\partial \ln(w_i)} = -\frac{\alpha_i}{1-h} - 1 < 0 \tag{B.4}
$$

And

$$
\frac{\partial \ln(x_i^*)}{\partial \ln(w_{j\neq i})} = -\frac{\alpha_j}{1-h} < 0 \tag{B.5}
$$

whereby it is clear that the domain of h implies negatively sloped own and cross price demand curves. It follows, moreover, that input own price elasticity of demand is strictly less than or equal to unity. Input cross price elasticity of demand is elastic if $h \in (1-\alpha_j, 1)$, inelastic if $h \in (0, 1-\alpha_j)$, and linear if $h = 1-\alpha_j$. Note that increasing returns to scale would imply strictly upward sloping cross price demand curves, and the possibility of upward sloping own price demand curves.

B.2 Proof of Equation 2.31

For lognormally distributed χ, the probability density function $\rho(\chi)$ is defined

$$
\rho(\chi) = \frac{1}{\sqrt{2\pi}s_\chi} e^{-\frac{(\ln(\chi)-m_\chi)^2}{2s_\chi^2}} \quad ; \quad m_\chi = E[\ln(\chi)] \quad s_\chi^2 = Var[\ln(\chi)] \tag{B.6}
$$

So for a given constant upper bound $\chi \leq \xi$,

$$\int_0^{\xi} \chi \rho(\chi) d\chi = \frac{1}{\sqrt{2\pi} s_\chi} \int_0^{\xi} \chi e^{-\frac{(\ln(\chi)-m_\chi)^2}{2s_\chi^2}} d\ln(\chi) \tag{B.7}$$

Introducing a change of variables,

$$\zeta = \frac{\ln(\chi) - m_\chi}{s_\chi} \; \rightarrow \; d\ln(\chi) = s_\chi d\zeta \tag{B.8}$$

Equation B.7 resolves as follows.

$$\begin{aligned} \int_0^{\xi} \chi \rho(\chi) d\chi &= \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\ln(\xi)-m_\chi}{s_\chi}} e^{(\zeta s_\chi + m_\chi)} e^{-\frac{\zeta^2}{2}} d\zeta \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\ln(\xi)-m_\chi}{s_\chi}} e^{-1/2[\zeta^2 - 2\zeta s_\chi - 2m_\chi]} d\zeta \end{aligned} \tag{B.9}$$

Where the argument of the exponential function resolves as follows,

$$\begin{aligned} -1/2[\zeta^2 - 2\zeta s_\chi - 2m_\chi] \\ &= -1/2[(\zeta - s_\chi)^2 - 2m_\chi - s_\chi^2] \\ &= -\frac{(\zeta - s_\chi)^2}{2} + m_\chi + \frac{s_\chi^2}{2} \end{aligned} \tag{B.10}$$

So the integral can be rewritten

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\ln(\xi)-m_\chi}{s_\chi}} e^{-\frac{(\zeta - s_\chi)^2}{2} + m_\chi + \frac{s_\chi^2}{2}} d\zeta \\ &= \frac{1}{\sqrt{2\pi}} e^{m_\chi + \frac{s_\chi^2}{2}^{2}} \int_0^{\frac{\ln(\xi)-m_\chi}{s_\chi}} e^{-\frac{(\zeta - s_\chi)^2}{2}} d\zeta \\ &= \frac{E[\chi]}{\sqrt{2\pi}} \int_0^{\frac{\ln(\xi)-m_\chi}{s_\chi}} e^{-\frac{(\zeta - s_\chi)^2}{2}} d\zeta \\ &= E[\chi] \Phi\left(\frac{\ln(\xi) - m_\chi}{s_\chi} - s_\chi \right) \end{aligned} \tag{B.11}$$

Where $\Phi()$ is the standard normal cumulative distribution function.

$$\Phi(\mathcal{Z}) = \frac{1}{\sqrt{2\pi}} \int_0^{\mathcal{Z}} e^{-\frac{\psi^2}{2}} d\zeta \tag{B.12}$$

■

B.3 Differentiation of expenditure TFP, yield, and expenditure with respect to TFP

The derivative of expenditure TFP (v) with respect to TFP (g) follows from Equations 2.21.

$$\frac{\partial \ln(v)}{\partial \ln(g)} = 1 - \boldsymbol{\alpha}' \frac{\partial \ln(\mathbf{w})}{\partial \ln(g)} \tag{B.13}$$

And the derivatives $\partial \ln(C^*)/\partial \ln(g)$ and $\partial \ln(y^*)/\partial \ln(g)$ then follow from the expressions in Equation 2.25.

$$\begin{aligned} \frac{\partial \ln(C^*)}{\partial \ln(g)} &= \frac{1}{1-h} \left(\frac{\partial \ln(v)}{\partial \ln(g)} + \frac{\partial \ln(C^*)}{\partial \ln(g)} \right) \\ \rightarrow \frac{\partial \ln(C^*)}{\partial \ln(g)} &= \frac{1}{1-h} \left(1 - \boldsymbol{\alpha}' \frac{\partial \ln(\mathbf{w})}{\partial \ln(g)} + \frac{\partial \ln(P)}{\partial \ln(g)} \right) \end{aligned} \tag{B.14}$$

And

$$\begin{aligned} \frac{\partial \ln(y^*)}{\partial \ln(g)} &= \frac{\partial \ln(v)}{\partial \ln(g)} - h \frac{\partial \ln(C^*)}{\partial \ln(g)} \\ &= \frac{1}{1-h} \left(1 - \boldsymbol{\alpha}' \frac{\partial \ln(\mathbf{w})}{\partial \ln(g)} + h \frac{\partial \ln(P)}{\partial \ln(g)} \right) \end{aligned} \tag{B.15}$$

B.4 Proof that optimal unit expenditure is a geometric brownian motion with the same parameters as output price

If $P(t)$ is a gBm with drift and volatility parameters m_p and s_p, and given the relation (Equation 2.87),

$$C^{\#}(t) = P(t) \frac{h}{\lambda}$$

Then Ito's Lemma states

$$\Delta C^{\#} = \left(\frac{\partial C^{\#}}{\partial t} + m_p P(t) \frac{\partial C^{\#}}{\partial P} + \frac{1}{2} s_p^2 P(t)^2 \frac{\partial^2 C^{\#}}{\partial P^2} \right) \Delta t + s_p P(t) \frac{\partial C^{\#}}{\partial P} \varepsilon \sqrt{\Delta t} \tag{B.16}$$

Where ε is a normally distributed random variable with mean 0 and standard deviation 1. (For a reference see Hull (2015).)

By Equation 2.87,

$$\frac{\partial C^{\#}}{\partial t} = 0 \quad \frac{\partial C^{\#}}{\partial P} = \frac{C^{\#}}{P} \quad \frac{\partial^2 C^{\#}}{\partial P^2} = 0 \tag{B.17}$$

Such that Equation B.16 resolves to

$$\Delta C^{\#} = m_p C^{\#} \Delta t + s_p C^{\#} \varepsilon \sqrt{\Delta t} \tag{B.18}$$

By which it follows that $C^{\#}$ has the same drift and volatility parameters as $P(t)$.

$$\begin{gathered} E[\Delta C^{\#}/C^{\#}] = m_p \Delta t \\ Var[\Delta C^{\#}] = s_p^2 \Delta t \end{gathered} \tag{B.19}$$

And thus

$$\begin{aligned} E[\ln(C^{\#}(T))]\big|_{t=0} &= \ln(C^{\#}(0)) + \left(m_p - \frac{s_p^2}{2} \right) T \\ Var[\ln(C^{\#}(T))]\big|_{t=0} &= s_p^2 T \end{aligned} \tag{B.20}$$

■

Appendix C: Kenya farm survey data details and preparation

The raw Kenya farm data files used in this article are open access and available from Tegemeo Institute upon request. The files detail short and long rains maize production (and that of any other crops grown), farm size, input (seed, synthetic and organic fertilizer, chemical pest control) use and cost, land tenure type, and land preparation type and cost, for the years 1997, 2000, 2004, 2007, and 2010. For the present article, only the 2007 long rains (i.e. main season) data is used. Data on production, area, seed, fertilizer, land preparation, and tenure are provided at field level. Fields where maize is not grown or is not the main crop are dropped. Green and dry maize production are reported separately. Green maize production is converted to dry maize equivalent by subtracting 15% of the production weight. The field level input and production files are then merged and aggregated up to farm level. Farm yield is then calculated as maize production (kg.) divided by maize area (acres). Chemical input use and cost are reported at farm level, indicating the main crop on the farm to which the input is applied. Chemical pest controls are counted as costs only if applied primarily to maize. Synthetic fertilizer consists of calcium ammonium nitrate (CAN), nitrogen phosphorus potassium (NPK), Urea, and diammonium phosphate (DAP). Different mixtures under these headings (e.g. NPK 22:11:11, NPK 18:12:12, NPK 17:17:0, etc.) are aggregated. Farm level cost is then calculated as the sum of seed, fertilizer, chemical pest control, and land preparation costs in Kenyan Shillings per acre (KES/acre).

Land preparation type (manual, oxen, tractor, or none) and land tenure type (owned by parent/relative, owned with title deed, owned without title deed, rented) are reported at field level. Dummy variables are created for land preparation types "oxen" and "tractor", and for the land tenure type "owned with title deed". When aggregating to farm level, the land preparation dummies are preserved if present on any field.

The field level data files also include details on seed type and method of acquisition. For this article, farms are considered adopters of hybrid maize seed if they use either purchased hybrid seed or a mixture of retained and purchased hybrid maize seed on any of their fields.

Main season rainfall is reported in millimeters at district level. (Prior to the 2010 constitution, counties were referred to as "districts".) The "rain anomaly" variable appearing in the regression is calculated as 2007 main season rainfall divided by 1997 main season rainfall. Rain anomaly values greater or less than 1 thus indicate departures north or south of the 1997 benchmark. The data indicate that 2007 main season rainfall exceeded the historical benchmark by a considerable margin for most of the sample (Figure 7). Summary statistics are presented in Table B1 for all variables in the yield model.

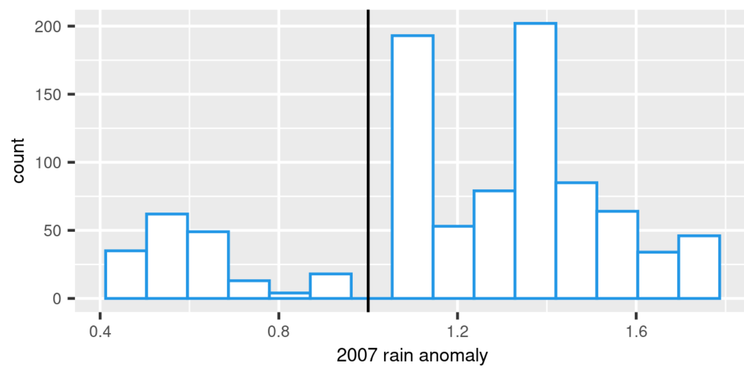

Figure 7: Histogram of 2007 rainfall anomaly. Values close to 1 (marked by the black vertical line) represent rainfall in line with the historical benchmark. The histogram indicates that rainfall exceeded this benchmark in 2007 for most of the farm sample.

Table C1: Summary statistics for the 2007 Kenya long rains farm level maize da

N = 937

	Mean	Stand. dev.	Max	Min
Yield (kg/acre)	993.993	687.01	9247	38.589
Cost/acre (KES/acre)	3051.197	3548.576	36626.667	9
Farm size (acres)	1.852	3.2	75	0.03
Planting density (kg seed/acre)	8.978	2.966	25.714	0.714
Adult equivalent household size	5.143	2.466	16.593	0.74
Rain anomaly	1.201	0.341	1.761	0.478
Tenure - owned with title deed (dummy)	0.549		1	0
Landprep - oxen (dummy)	0.248		1	0
Landprep - tractor (dummy)	0.248		1	0
Hybrid seed (dummy)	0.761		1	0

Figure 8: Summary statistics for the 2007 Kenya long rains farm level maize data.

–>

–>