What Drives Smallholders' Productivity in Pakistan's Horticultural Sector?

Shabbir Ahmad*
UQ Business School
The University of Queensland, Brisbane, Australia
E-mail: s.ahmad@uq.edu.au

Sriram Shankar
ANU Centre for Social Research and Methods & Research School of Economics,
Australian National University, Canberra, Australia
E-mail: Sriram.shankar@anu.edu.au

John Steen
UQ Business School
The University of Queensland, Brisbane, Australia
E-mail: j.steen@business.uq.edu.au

Martie-Louise Verreyne
UQ Business School
The University of Queensland, Brisbane, Australia
E-mail: m.verreyne@business.uq.edu.au

Abid Aman Burki
Department of Economics
Mushtaq Ahmad Gurmani School of Humanities and Social Sciences
Lahore University of Management Sciences, Lahore, Pakistan
E-mail: burki@lums.edu.pk

*Corresponding author: Shabbir Ahmad, UQ Business, School, The University of Queensland, Brisbane, Australia, Phone: +61 7 3346 8083 E-mail: s.ahmad@uq.edu.au

Acknowledgement: This paper was partially funded by a research grant from the Australian Institute of Business and Economics (AIBE), the University of Queensland, Australia and the Australian Centre for International Agricultural Research for the project ‘Farmer’s capabilities, productivity, and profitability: A case study of smallholders in selected agro zones in Pakistan’(ADP/2015/004). All errors remain our own.
What Drives Smallholders' Productivity in Pakistan's Horticultural Sector?

Abstract
Smallholders are indispensable to ensuring food security in the developing economies where they farm. Policy interventions often target smallholders to provide for example, input subsidies, extension services and access to credit, because increased total factor productivity (Hsieh & Klenow, 2009) can ensure that they are better placed to support food security. However, the impact of such interventions and the drivers of TFP growth are largely unknown due to lack of comprehensive data and appropriate methodology. To overcome these impediments, we propose an econometric estimation of the components of TFP growth in a Bayesian set-up and apply this to new farm-level survey data of smallholders from Pakistan’s horticulture sector. The results indicate large technical and mix efficiency differentials across agro-climatic zones and farm sizes. These disparities in technical and mix efficiency are due to suboptimal farm practices, potentially from limited access to and adoption of technology. Government policy makers, support agencies, farmer groups and other stakeholders have latitude in providing adequate education and training programs aimed at improving input-use efficiency and introducing innovative practices leading to TFP growth.

Key words: Scope economies, developing economy, aggregator function, mix efficiency, TFP

JEL Classification: D31

1. INTRODUCTION
The role of agricultural development in poverty reduction has been a source of continued debate. While agriculture continues to be the backbone of emerging economies, other factors such as lower prices for farm produce and persistent structural impediments to increased farm income have raised questions as to what degree agriculture can actually contribute to improved living standards (Christiaensen & Martin, 2018). Nonetheless, growth in agriculture is still two to three times more effective for poverty alleviation compared to other sectors of developing economies (Dorosh & Thurlow, 2018; Ligon & Sadoulet, 2018; Ravallion & Datt, 1996). Furthermore, agricultural sector productivity growth creates opportunities for the most economically marginalised communities in society (Christiaensen, Demery, & Kuhl, 2011; Ligon & Sadoulet, 2018).
While agriculture sector productivity presents itself as a development policy target for government, farmer groups and international agencies (e.g., FAO) there are challenges that arise from the predominance of small acreage farmers, or smallholders, in many developing economies (Lowder, Skoet, & Raney, 2016). Numerous factors have been identified that may affect the productivity of smallholder farms, such as adoption of technology, improved farming practices, and scale and scope economies (Mugera, Langemeier, & Ojede, 2016; Shenoy, 2017; Wainaina, Tongruksawattana, & Qaim, 2017). Policy-makers with limited resources must choose which factors to address because programs that target many thousands of small farmers are costly, especially those that involve time-intensive interventions such as training or technological support.

For policy analysts to make such evidence based choices it is important to understand the causes of low productivity. Productivity measurement and decomposition tends to be limited to technical and/or scale efficiencies. Further, compared to other agricultural sectors such as grain crops and livestock, the measurement of productivity and its decomposition in the horticultural sector has largely been overlooked. Understanding the drivers of productivity in this sector is especially important because the products from horticulture feed much of the population in emerging economies and have the potential to support further employment through value addition to create higher margin products (Gomez & Ricketts, 2013; Weinberger & Lumpkin, 2007). More sophisticated productivity analysis will be able to support policy targeting to address the wide variations in production efficiency within horticulture that has been previously noted in the literature. For example, policy initiatives to improve scale-mix efficiency may be different from programs to improve labour productivity. Within this paper we not only deploy novel methods for decomposing productivity, but also discuss how this can lead to better targeting of development support and policy initiatives.
This paper decomposes farm productivity into two components: technical efficiency (affected by farm management practices) and mix efficiency (influenced by farm resource allocation). The former may be improved by education, training, and the adoption of better practices, whereas the latter will be affected by changes in output or input mixes (i.e., scope economies) to achieve higher levels of productivity. Mix efficiency is defined as the potential improvement in productivity when input or output mixes are changed. Any change in output mix (e.g., balance of staple and vegetable crops) or input mix (e.g., machinery-to-labour ratio, fertilizer and pesticide) results in changes to productivity.

Significant variations are found in efficiency across horticultural produce (such as fruits and vegetables). In this paper we observe large disparities among smallholders across different landholding sizes as well as within agro climatic environments and this has important implications for the delivery of agricultural development programs. We apply econometric methods to estimate the components of productivity, which has not been attempted in the extant literature. For instance, O’Donnell (2012b) has proposed mix-efficiency measurement using non-parametric methods such as data envelopment analysis. Our paper makes an empirical contribution to the productivity policy literature by estimating input mix efficiency (a pure productivity concept) in a Bayesian framework. One of the advantages of the Bayesian stochastic frontier is that one can impose regularity restrictions (e.g., non-negative input shares), which is imperative for meaningful policy analysis.

The paper is organized into sections. Section 2 describes the agricultural sector of Pakistan and a snapshot of agricultural development policies, with particular attention on smallholders. Data and methodology are discussed in Section 3, discussion of empirical results is provided in Section 4, and concluding remarks follow in Section 5.
2. AGRICULTURAL LANDSCAPE AND SMALLHOLDER AGRICULTURAL POLICIES IN PAKISTAN

Agriculture is a significant contributor to Pakistan's economy, accounting for around one-fifth of the country's Gross Domestic Product (GDP) and employing more than two-fifths of the country's labour force (GoP, 2018). Smallholders play an important role in the agricultural GDP of Pakistan as they make up 78 percent of the rural households and provide the main source of livelihood for many rural poor through various farming activities (Mellor & Malik, 2017).

Pakistan has a diversified topography and varying land fertility across its regions. This results in changing crop cultivation patterns and huge differentials in their respective yields. Pakistan is classified into various cropping regions, spanning mostly arid and semi-arid agro-climatic zones (Mellor & Malik, 2017; Pinckney, 1989). Smallholders often diversify their farms by growing a variety of crops including grains, fruits and vegetables in these agro-climatic zones. Significant variations in crop yield and productivity within different cropping systems and agro-climatic zones can be attributed to varying farming practices and access to inputs and other services (Spielman, Malik, Dorosh, & Ahmad, 2016).

The land ownership structure in Pakistan has changed significantly over the last few decades (Mellor & Malik, 2017). A vast majority of smallholders hold fewer than five acres of land and this number is increasing as land holdings are passed on and divided between inheriting family members (see Figure 1). For instance, farmers holding under five acres of land increased from 47 percent in 1990 to 65 percent in 2010. In contrast, there has been a continuous decline in large landholdings. One of the obvious reasons for this shrinking landownership is the Muslim inheritance law by which land is distributed among the children of the deceased. In the wake
of declining farm size, smallholders face constraints including limited capital investment in farm improvement, slow uptake of innovative farm practices, and the challenge of linking small-scale farmers to efficient input services such as credit, extension and purchased inputs, which in turn hamper farm productivity.

Figure 1: Changing structure of agricultural landholding in Pakistan
Source: Adapted from Mellor and Malik, 2017

Since the independence of Pakistan in 1947, the policy environment has evolved. The earliest policies were geared towards enhancing production of cereals (wheat and rice), cash crops (cotton and sugarcane) and land distribution. This was followed by the policies during the “Green Revolution” from the 1970s to 1990s (Spielman et al., 2016). As a result, there was a substantial increase in the yield of staple crops. Following this, several other policy measures targeted increases in farm productivity, including the “National Agricultural Policy” and agriculture sector liberalization measures in the early 1980s. During the same period, Agricultural Prices Commission (APC) of Pakistan and the Pakistan Agriculture Research Council (PARC) were established, aiming to ensure commodity market stability (e.g., prices) and development of research and innovation to improve agricultural productivity. The National
Agriculture Policy also aimed to attain food self-sufficiency through improved irrigation and extension services, technology adoption and provision of better inputs. The main objective of the establishment of PARC was to conduct research to introduce improved varieties, pest management, and technology adoption to enhance farm productivity.

Heavy subsidies after independence created large distortions in both input and commodities markets. While these subsidies aimed to help the marginalized smallholders, most benefits were captured by the big landholders, which resulted in low productivity at significant cost to the government. Over most of this period, the main focus of government support was on four major crops: wheat, cotton, rice and sugarcane (Ali & Byerlee, 2002).

Government interest in the horticulture sector in Pakistan is more recent, and has emerged as an important contributor to GDP over the past few decades. Pakistan produces more than 28 types of fruit in four seasons throughout the year, and there are more than 30 different types of vegetables produced across its different geographical regions. In 2002, the Pakistan Horticulture Development and Export Board (PHDEB) was established under the auspices of the Export Promotion Bureau (EPB) to promote and develop horticulture exports. The PHDEB deals with product quality enforcement, certification and inspection of horticulture produce (CENTAD, 2009).

In recent years, there have been a number of developments in horticulture. The introduction of tunnel farming has boosted off-season production of vegetables, which has increased profitability for farmers. Some of the abundantly produced vegetables (e.g., potatoes and onions) have great potential for export. Floriculture, a relatively new industry in Pakistan, has been fast growing, yet most of its production is still consumed domestically. The horticultural
sector needs the development of government policies to support sustainable growth and promote export strategies for future profits. Measuring productivity and efficiency in horticulture is important for effectively targeting and evaluating policy and interventions. We address this challenge in the next section.

3. PRODUCTION TECHNOLOGY, AGGREGATOR FUNCTIONS AND MEASURES OF EFFICIENCY

3.1 Technology and its properties

The estimation of the firm performance\(^1\), characterized by various measures of efficiency, requires a well-defined production technology to transform inputs into outputs. Farmers make use of different technologies in the production of various crops (e.g., fruit & vegetables) in different environments (O'Donnell, 2016). However, they need to adjust input combinations in response to varying production environments (Hadley, Fleming, & Villano, 2013). Let \( q \) be a vector of \( m \) outputs, \( x \) be a vector of \( k \) inputs, and \( z \) the vector of \( l \) exogenous factors (e.g., environment). These output-input combinations can be described in a specific production environment as \( T = \{ (q, x) : x \text{ can produce } q \text{ in environment } z \} \). We assume that technology maintains the following standard properties: i) there is the possibility of producing zero output using given amount of input (i.e., no free lunch or inactivity); ii) output or inputs can be disposed of without incurring any cost (i.e., free disposability);\(^2\) iii) a finite amount of inputs can produce only limited amount of outputs (i.e., boundedness); iv) to strictly produce a positive output, it requires a strictly positive amount of at least one input (i.e., weak essentiality); and v) the set of production possibilities contains its boundaries (i.e.,

---

\(^1\) We use the term firm to refer an individual farm.

\(^2\) In other words, the same input vector that is used in the production of a particular output vector can also produce a scalar contraction of that output and vice versa (i.e., strong disposability in inputs and outputs).
boundedness). If all the above assumptions are maintained, the technology is said to be regular.

Alternatively, the technology can also be represented by distance, cost, revenue or profit functions (Balk, 1998; O'Donnell, 2012a). For instance, Shephard (1970) uses output and input distance functions. The input distance function measures the extent to which an input vector can be scaled down (by a factor) while holding an output vector fixed. The output distance function measures the extent to which an output vector can be scaled up (by a factor) while holding the input vector fixed. If the technology satisfies the above assumptions, then output and input distance functions satisfy the properties of i) linear homogeneity; ii) non-decreasing in output and inputs, respectively; and iii) the boundedness of input- and output–oriented technical efficiency.

3.2 Output and Input Aggregators
The estimation of meaningful TFP components require aggregator functions that are consistent with the index number theory. There are various price and technology based quantity aggregators that can lead to a range of TFP indexes. For example, price-based quantity aggregators include Laspeyres, Paasche, Fisher and Lowe, whereas technology-based aggregators include Malmquist (Caves, Christensen, & Diewert, 1982), Bjurek-Moorsteen (Bjurek, 1996) and Fare and Primont (O'Donnell, 2012a). Price-based quantity aggregators are commonly constructed using either the economic theoretic or the axiomatic approach, which rests on the assumptions of producer behaviour (e.g., cost minimization or revenue maximization) and which needs to satisfy certain properties of cost, revenue or profit functions.

---

3 The closeness of outputs and inputs plays an important role in the construction of production technology. It states that an output set that can be produced by a given vector of inputs includes all its points on the boundary as well as an input set that can produce a given output vector lies all points on its boundary.
In contrast, technology-based quantity aggregators do not require price or value information in their construction.

The output and input quantity aggregates of a firm $i$ can be represented by a function of output vector $(q_i)$ and input quantities $(x_i)$ such that $Q_i = Q(q_i)$ and $X_i = X(x_i)$, where, $X(.)$ is non-decreasing, nonnegative and linearly homogenous.\(^4\) Technology based input aggregator functions can be represented by a distance function $D_i(q_i, x_i)$, where $D_i(.)$ represents the input-oriented distance function.\(^5\) Using input distance function as an aggregator function leads to different quantity indexes. For example, a technology-based Malmquist input aggregator having these properties is the input distance function and the associated quantity index is the Malmquist input quantity index. However, Malmquist index is multiplicatively incomplete (except in special circumstances) and does not allow conducting a spatial comparison due to intransitivity (Balk & Althin, 1996). Transitivity is important in a cross sectional data setting where the data points are not in order (O'Donnell, 2012a, 2014, 2016). For instance, if vegetables (A) and fruit (B) producing farms are equally productive, and if farm A is twice as productive as a grain (C) producing farm, then transitivity shows that farm B is twice as productive as farm C. The property of transitivity has important policy implications as it allows multilateral comparison across firms/farms (O'Donnell, 2014).

3.3 Measures of Efficiency

---

\(^4\) Let $p_i$ and $w_i$ denote the associated output and input price vectors, which can be used to construct price-based aggregators (e.g., Laspeyers, Paasche). However, if for instance, $p_i = \bar{p}_i$ and $w_i = \bar{w}_i$ then it gives Lowe output and input aggregator.

\(^5\) A detailed discussion on input and output aggregators can be found in O’Donnell, 2012a.
The idea of economic efficiency and its decomposition into technical efficiency (TE) and allocative efficiency (AE) has been around since it was first formalized by Farrell (1957) in his seminal work, and built on the work of Debreu (1951) and Koopmans (1951). Excessive use of inputs for a given level of output or the production of less output from a given level of inputs results in technical inefficiency, while inappropriate use of the mix of inputs leads to allocative inefficiency. In this paper, we focus on mix efficiency instead of allocative efficiency. Mix efficiency is a relatively new but purely technical concept (i.e., associated with scope economies) as compared to allocative efficiency, which is a value-based judgment (e.g., cost minimization or profit maximization) (Kumbhakar, Ghosh, & McGuckin, 1991; Kumbhakar & Tsionas, 2005). Figure 1 illustrates input-oriented technical and mix efficiency in an aggregate input framework. Input-oriented technical efficiency is defined as the ratio of minimum possible input aggregate to observed input aggregate while holding input mixes fixed. This is explained in the graph using input space \( x_i = (x_{i1}, x_{i2}) \). The curve passing through \( a' \) and \( e' \) shows all technically efficient inputs to produce an output vector \( q_i \). The dashed line passing through points \( a' \) and \( d' \) is an iso-aggregate input, \( X_i \), which is a linear combination of input vector \( x_i \) such that \( X_i = \delta_1 x_{i1} + \delta_2 x_{i2} \). We call it an iso-aggregate input as the line maps all the points having the same aggregate input. While holding input mixes and output vector fixed, the firm operating at point \( a' \) can reduce its input aggregate to point \( b' \). By definition, input-oriented technical efficiency is the ratio of minimum possible aggregate input to observed aggregate input when both output vector and input mixes are held fixed. Thus, the input-oriented technical efficiency can be defined as \( \text{ITE}_i = \frac{\bar{X}_i}{X_i} \), where, \( \bar{X}_i \) (a scalar multiple of \( x \)) is the minimum aggregate input quantity when both input mixes and the output vector are held fixed. In contrast, mix efficiency occurs as a consequence of relaxing restrictions on input and output mixes. If input mixes are allowed to vary, by holding the output vector fixed the firm can further reduce its input aggregate to point \( e' \) (i.e., the minimum possible input...
aggregate). The input-mix efficiency can be defined as \( IME_i = \frac{\hat{X}_i}{\bar{X}_i} \), where, \( \hat{X}_i(\hat{x}_i) \) is the minimum possible aggregate input when input mixes are allowed to vary, but output vector is fixed. O’Donnell (2012b, p. 261) also calls it pure input mix efficiency, because input mixes are allowed to change while holding the output vector fixed. Thus, it measures the potential change in productivity when the restrictions on input mixes are relaxed. However, if we replace \( \delta_i \) with input prices, \( w_k \), where \( k = 1, 2 \) then it produces the firm’s allocative efficiency as discussed in Schmidt and Lovell (1979).

4. METHOD AND DATA

In this paper, we illustrate how to estimate mix efficiency using a linear input aggregator similar to Schmidt and Lovell (1979). Mix efficiency is a pure technical concept and an obvious consequence of input-mix efficiency is that farmers may avoid over-use of some inputs in response to substitution policies, which may increase the productivity of farms.

4.1 Analytical method of Mix Efficiency

Figure 1: Representation of technical and mix efficiency in aggregate input framework (O’Donnell, 2012)
We use the linear aggregator to derive the measures of input oriented mix efficiency of the following form:

\[ X(x_i) = \sum_{k=1}^{K} \gamma_k x_{ki}, \quad \gamma_k \geq 0 \]  

(1)

The technology is characterized by the homogenous production function as:

\[ h(q_i) = g(z_i) f(x_i)^\eta \exp(\xi_i) \]  

(2)

where \( h(q_i) \) is a function of output vector, \( f(x_i) \) defines production technology and \( g(z_i) \) represents factors beyond a farmer’s control (e.g., environment). More specifically,

\[ \ln h(q_i) = Q_i; f(x_i)^\eta = \prod_{k=1}^{K} x_{ki}^{\beta_i}, g(z_i) = \exp(\delta_0 + \sum_{l=1}^{L} \delta_l z_{li}); \text{ and } \xi_i = v_i - u_i, \text{ such that} \]

\[ \ln \left( \frac{f_k}{f_1} \right) = \ln \left( \frac{X_k}{X_1} \right) + \psi_k \text{ for } k = 2, \ldots, K. \]  

(3)

Where \( f_k / f_1 \) represents the ratio of marginal products and \( X_k / X_1 \) represents the ratio of first derivatives of aggregator function, respectively, and \( \psi_k \) represents errors that firms make in choosing an appropriate input mix. The presence of optimization errors (i.e. \( \psi_k \neq 0 \)) may increase or decrease the use of \( x_k \) (relative to \( x_1 \)) depending on value of \( \psi_k \). If \( \psi_k > 0 \), then the firm will under-utilize the \( k \)-th input mix with respect to input 1 and \( \psi_k < 1 \) indicates that the firm will over-utilize the \( k \)-th input mix with respect to input 1. Under- or over-utilization
of input mix leads to decreased productivity of a firm/farm. This minimization problem is similar to Schmidt and Lovell (1979), but our minimization problem has a different interpretation than the conventional cost minimization problem. We minimize a linear input aggregator subject to homogenous technology (i.e., Cobb-Douglas) to derive measures of input mix efficiency. Recently, Ahmad (2018) has proposed a non-linear aggregator to derive the measures of mix efficiency. To derive the first-order conditions for our minimization problem, the Lagrange function can be set up as:

\[
L = \sum_{k=1}^{K} \gamma_k x_{ki} + \lambda \left( \ln Q_i - \delta_0 - \sum_{l=1}^{L} \delta_l z_{il} - \sum_{k=1}^{K} \beta_k \ln x_{ki} - \varepsilon_i \right)
\]

where \(Q_i\) represents an aggregate output, \(x_i\) is a vector of inputs, and \(z_i\) is a vector of environmental variables, respectively; \(\gamma_k \geq 0, \beta_k \geq 0, \delta_0\) represents the constant term, \(\delta = (\delta_1, \ldots, \delta_L)\) is a vector of unknown parameters, and \(\sum_{k=1}^{K} \beta_k = \eta\). The error term, \(\varepsilon_i = v_i - u_i\), where \(v_i\) is an error term taking into account statistical noise due to factors such as droughts or floods and other errors of approximation and is i.i.d., whereas \(u_i\) is a one-sided error measuring the extent of firm technical inefficiency. Firms may not be able to achieve the maximum output and may remain below the frontier due to inefficient use of inputs. Together this represents the set of environmental variables where \(\varepsilon_i = v_i - u_i\).

If a firm is both technical and mix inefficient, then combining the first order conditions (3) with the production technology is described by the following factor demand equation:

\[
\ln x_{ki} = \ln b_k + \frac{1}{\eta} \ln Q_i + \ln \left( \beta_k / \delta_k \right) + \frac{1}{\eta} \sum_{k=2}^{K} \beta_k \psi_k - \psi_k \frac{1}{\eta} (v_i - u_i)
\]

where \(b_k = \beta_k \left( g(z_i) \prod_{k=1}^{K} \beta_k / \delta_k \right)^{\frac{\beta_k}{\eta}}\).
Combining factor input (5) with input aggregator function (1), we get the input aggregate as

\[ \ln X_i = \ln B + \frac{1}{\eta} \ln Q_i + \frac{1}{\eta} \sum_{k=2}^{K} \beta_k x_k + \ln \left( \sum_{k=1}^{K} \beta_k \exp(-\psi_k) \right) - \frac{1}{\eta}(\nu_i - u_i), \]

where \( \ln B = \sum_{k=1}^{K} \ln b_k \).

If a firm is fully technical and mix efficient (i.e., \( IME = ITE = 1 \)), it implies that \( u_i = 0 \) and \( \psi_2 = \psi_3 = \ldots = \psi_k = 1 \). If we set the random error term equal to zero (i.e., \( E(\nu_i) = 0 \)), then the aggregate input (6) reduces to

\[ \ln \hat{X}_i = \ln B + \frac{1}{\eta} \ln Q_i + \ln \eta \]

Now the ratio of actual input aggregate to minimum possible aggregate input can be expressed in two components due to mix and technical efficiencies as

\[ \frac{\hat{X}_i}{X_i} = \frac{\hat{X}_i}{\hat{X}_i} = IME_i \times ITE_i \]

The solutions to the minimization problem in equation (4) give the following expressions for input-oriented technical efficiency (ITE) and input-oriented mix efficiency (IME)

\[ ITE_i = \frac{\hat{X}_i}{X_i} = \exp(-u_i / \eta) \]

\[ IME_i = \exp \left[ \ln \eta - \left( \frac{1}{\eta} \sum_{k=2}^{K} \beta_k \psi_k + \ln \left( \sum_{k=1}^{K} \beta_k \exp(-\psi_k) \right) \right) \right] \]

The detailed derivation of equation (10) is provided in the Appendix A.

4.2 Empirical Setup

We estimate the conventional cross sectional stochastic frontier model as follows:

\[ \ln Q_i = \delta_0 + \sum_{l=1}^{L} \delta_l z_{il} + \sum_{k=1}^{K} \beta_k \ln x_{ki} + v_i - u_i \]
where \( Q_i \) is an aggregate output of firm \( i \) and \( x_{ki} = (x_{ki1}, x_{ki2}, x_{ki3}, x_{ki4}, x_{ki5}, x_{ki6}) \) represents the input vector. The vector \( z_{ii} = (z_{ii1}, ..., z_{iiL}) \) represents environment and innovation factors.

We estimate the Bayesian stochastic frontier model as follows:

\[
y = X\beta - v - u
\]

where \( y = \ln Q \) is a vector of aggregate output; \( X \) is an input matrix of order \( N \times (K + J + 2) \); \( \beta = (\beta_1, ..., \beta_K, \delta_0, \delta_1, ..., \delta_J) \) represents the vector of parameters to be estimated; \( v = (v_1, ..., v_N) \) is a vector of normal random errors representing the combined effect of measurement errors and errors occurring due to approximation of the functional form; and \( u \) is the vector in non-negative inefficiency effects. We assume that all elements in \( v \) are identically and independently distributed with joint probability density function \( v \sim N(0, h^{-1}) \), where \( h \) is a precision variable (i.e., the inverse of the variance \( \sigma^2_v \)); All the elements in \( u \) are also independently distributed and account for technical inefficiency.

The likelihood function is given by

\[
p(y|\beta, h, u) = (2\pi)^{-\frac{N}{2}} (h)^{\frac{N}{2}} \exp \left[ -\frac{1}{2} \left( y - X\beta + u \right)' \left( y - X\beta + u \right) \right]
\]

The elements of vector \( u \) are independent random variables, drawn from an exponential distribution that has a common unknown parameter \( \lambda \) that is \( p(u|\lambda^{-1}) = f_G(u|1, \lambda^{-1}) \). A more general distribution for \( u_i \) was used by van den Broeck, Koop, Osiewalski & Steel (1994) and Koop, Steel & Osiewalski (1995), but these distributions make it difficult to distinguish inefficiency errors, \( u_i \), from the normally distributed noise errors, \( v_i \) (Ritter & Simar, 1997). In addition, van den Broeck, Koop, Osiewalski & Steel (1994) discovered that an exponential

---

6 \( f_G(a|b,c) \) is used to indicate that \( a \) has a Gamma distribution with shape parameter \( b \) and scale parameter \( c \),
distribution is moderately robust to changes in priors. To sample from posterior density, it is helpful to use a Gibbs sampler with data augmentation.

The likelihood function given by equation (13) is very challenging computationally if approached via direct approximation of integrals as it will contain terms involving cumulative distribution functions, which are not available in closed form. Direct evaluation of the posterior densities are not only costly to compute, but their mathematical expressions are usually awkward and do not allow any analytic approximation. Since the required full conditional densities derive directly from the joint posterior, they are expensive and difficult to sample. Therefore, it is convenient to “augment” the observed data by drawing observations on \( u \). The Gibbs sampler with data augmentation necessitates drawing sequentially from the following conditional posteriors (see Appendix B for details).

4.3 Data and description of variables

The data come from a survey of horticulture farms in the Punjab province of Pakistan, conducted by the authors in September-November 2016. A multistage sampling strategy was adopted to select the sample of horticulture farms. In Stage 1, various vegetable and fruit farming clusters were identified in different agro-climatic zones and four horticulture districts were selected for data collection: Kasur, Lahore, Muzaffargarh and Sargodha. The selected districts were located in different agro-climatic zones with varying farming practices. For example, horticulture farms in Kasur district grow vegetables such as cauliflower, eggplant (brinjal), potato, tomato, okra, and turmeric (spice); farms in Lahore district grow cucumber, cauliflower, tomato and okra; Muzaffargarh district is known for its high-quality mangoes.

---

7 For a detailed account, see Gelfand, Adrian, & Lee (1992) and Tanner (1991).
8 For a classification of agro-climatic zones in Pakistan, see Pinckney (1989).
chilies, and tomatoes and also produces traditional wheat crops; and horticulture farms in Sargodha are famous for production of citrus and vegetables.

In Stage 2, a list of mouzas\(^9\) was used to randomly select villages. Finally, 730 horticulture farms covering 200 villages in the above mentioned agro-climatic zones were randomly selected from a list provided by the relevant local agriculture department offices. A group of 25 university graduates with agricultural sciences or social sciences degrees was selected to conduct the field surveys. These graduates belonged to the sampled districts and thus understood the local language. Prior to data collection, enumerators attended a rigorous two-day training workshop for data collection followed by a pilot survey. Enumerators were supervised by the faculty from five universities in Pakistan. Figure 2 presents the location of selected districts on the map of Punjab province.

![Map of Punjab Province showing location of four districts included in the survey](image)

**Figure 2: Map of Punjab Province showing location of four districts included in the survey**

All outputs were aggregated using the Lowe aggregator to construct a single output variable \((Q)\). All input quantities were aggregated with the Lowe aggregator to construct six input variables: ploughing

---

\(^9\) A mouza (sometime called Deh) is an administrative revenue estate within a district consisting of one or more than one village where farming communities inhabit in compact or scattered form.
(x_i), seeds(x_j), irrigation(x_k), fertilizers(x_l), pesticides (e.g., weedicides, herbicides, fungicide, and insecticides)(x_m); and labour (x_n).

Tunnel farming dummy variable (z_i) is included to control for the impact of technology adoption by the farmers. To capture the impact on farm productivity of agro-climatic zones that differ in soil and other geographic characteristics, we include three dummy variables: Muzaffargarh (z_2), Kasur (z_3) and Sargodha (z_4), while Lahore is the base category. As noted above, district Kasur is known for growing vegetables (e.g., cauliflower, potatoes and egg-plant), while farmers in districts Muzaffargarh and Sargodha are known for growing mangoes, chilies and citrus along with other vegetables and staple crops. Table 1 provides summary statistics of all the included variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output (Q)</td>
<td>2.52x10^6</td>
<td>1.64x10^7</td>
<td>5.35x10^4</td>
<td>4.42x10^8</td>
</tr>
<tr>
<td>Ploughing (x_i)</td>
<td>4.77x10^5</td>
<td>4.77x10^5</td>
<td>3.32x10^3</td>
<td>1.29x10^7</td>
</tr>
<tr>
<td>Seed (x_j)</td>
<td>1.67x10^6</td>
<td>1.67x10^6</td>
<td>4.28x10^3</td>
<td>4.48x10^7</td>
</tr>
<tr>
<td>Irrigation (x_k)</td>
<td>1.85x10^6</td>
<td>1.85x10^5</td>
<td>1.45x10^3</td>
<td>4.2010^6</td>
</tr>
<tr>
<td>Fertilisers (x_l)</td>
<td>2.46x10^6</td>
<td>2.46x10^6</td>
<td>7.76x10^3</td>
<td>6.66x10^7</td>
</tr>
<tr>
<td>Pesticides (x_m)</td>
<td>8.45x10^5</td>
<td>8.45x10^5</td>
<td>7.52x10^2</td>
<td>2.28x10^7</td>
</tr>
<tr>
<td>Labour (x_n)</td>
<td>827.71</td>
<td>827.71</td>
<td>32</td>
<td>7.89x10^3</td>
</tr>
<tr>
<td>Tunnel Farming (z_i)</td>
<td>0.06</td>
<td>0.34</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Muzaffargarh (z_2)</td>
<td>0.33</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Kasur (z_3)</td>
<td>0.20</td>
<td>0.40</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Sargodha (z_4)</td>
<td>0.33</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Sample size</td>
<td>730</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**5 EMPIRICAL RESULTS AND DISCUSSION**

We estimate the Bayesian stochastic Cobb-Douglas production function to compute input-oriented (ITE) technical efficiency and input-oriented mix efficiency (IME) as components of farm level productivity. These are important drivers of farm productivity and can help to draw policy implications to help smallholders. The estimation of the Bayesian stochastic frontier enables us to draw exact
inferences on efficiencies. In addition, it is relatively easy to incorporate prior information and regularity restrictions, and this method provides a formal treatment of parameters and model uncertainty through numerical integration methods for complex stochastic frontier models. We can draw the highest posterior densities (HPD) of farm specific efficiency measures. In Bayesian stochastic frontier analysis (SFA) the distribution of inefficiency components is determined using a posterior simulator (e.g., the Gibbs sampler), and (van den Broeck et al., 1994) provided the earliest estimation of the Bayesian stochastic frontier for cross section data. Later on, a number of researchers have applied these methods both on cross section and panel data sets (Burki, 1996; Griffin & Steel, 2007; Gary Koop, Osiewalski, & Steel, 1997; O'Donnell, 2014).

5.1 Bayesian Estimates of Production Function

We obtain these results from 10,000 draws using the MCMC algorithm, which can be implemented relatively easily in the Bayesian context. Figure 3 shows the distribution of posterior densities of all parameter estimates. To impose regularity conditions, in the case of the Cobb-Douglas production function, for example, input elasticities are assumed to be non-negative (i.e., $\beta \geq 0$) as implied by economic theory. We impose monotonicity restrictions on input coefficients, which guarantee that all output elasticity coefficients are economically feasible. If these conditions are not satisfied, then inferences about the efficiency estimates may be misleading.

Table 2 shows results for the Bayesian stochastic production function posterior means, standard deviations and 95 percent HPD intervals (labelled 2.5% and 97.5%, respectively) for all the estimates $(\beta, \gamma, \lambda, \sigma^2)$. We use the Bayes Factor (BF) to test the hypothesis whether the environmental variables, $(z_i - z_a)$ should be included in the stochastic production function (i.e., $\delta = 0$). We compute the odd ratios of posteriors of restricted and unrestricted production frontiers to obtain the BF. A positive value of BF (i.e., 1348.2) supports the hypothesis that the inclusion of environmental variables is preferred.
The first-order coefficients of the production function correspond to input elasticities. These estimates are plausible, showing a positive relationship between inputs and the aggregate farm output. For instance, the coefficient of seed (0.40) indicates that it contributes 40 percent to farm production, whereas fertilizers (0.31) contribute 31 percent to the production. Production technology exhibits decreasing returns to scale as the sum of input elasticities remains less than unity (i.e., 0.98), which may be attributed to slow technology uptake and inappropriate use in input mixes. This is not surprising as most of the farmers in the sample data set have low education and training.

Figure 3: Posterior densities of production function estimates
The technology and environmental variables are also included in the production function to examine if there are any significant differences in farm production due to use of technology and across different regions. We use district dummies to account for different agro-climatic zones as well as varying crop activities in these districts. The Lahore district is used as a base category in this analysis. Estimates of agro-climatic zones show significant differentials in farm production. For instance, mango (Muzaffargarh) and citrus (Sargodha) regions are more productive than the reference region (Lahore). Similarly, the tunnel farming dummy variable shows that farmers using tunnel cultivation are more productive, indicating that technology adoption contributes to smallholders’ production.

**Table 2: Bayesian Estimates of Stochastic Frontier**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Mean</th>
<th>SD</th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>$\beta_0$</td>
<td>3.12</td>
<td>0.42</td>
<td>2.29</td>
<td>3.94</td>
</tr>
<tr>
<td>Ploughing ($x_1$)</td>
<td>$\beta_1$</td>
<td>0.12</td>
<td>0.06</td>
<td>0.02</td>
<td>0.23</td>
</tr>
<tr>
<td>Seed ($x_2$)</td>
<td>$\beta_2$</td>
<td>0.40</td>
<td>0.05</td>
<td>0.30</td>
<td>0.49</td>
</tr>
<tr>
<td>Irrigation ($x_3$)</td>
<td>$\beta_3$</td>
<td>0.06</td>
<td>0.02</td>
<td>0.02</td>
<td>0.10</td>
</tr>
<tr>
<td>Fertilizers ($x_4$)</td>
<td>$\beta_4$</td>
<td>0.31</td>
<td>0.07</td>
<td>0.18</td>
<td>0.44</td>
</tr>
<tr>
<td>Pesticides ($x_5$)</td>
<td>$\beta_5$</td>
<td>0.03</td>
<td>0.02</td>
<td>0.00</td>
<td>0.08</td>
</tr>
<tr>
<td>Labour ($x_6$)</td>
<td>$\beta_6$</td>
<td>0.06</td>
<td>0.03</td>
<td>0.01</td>
<td>0.12</td>
</tr>
<tr>
<td>Tunnel ($z_1$)</td>
<td>$\gamma_1$</td>
<td>0.11</td>
<td>0.07</td>
<td>-0.02</td>
<td>0.25</td>
</tr>
<tr>
<td>Muzaffargarh ($z_2$)</td>
<td>$\gamma_2$</td>
<td>0.97</td>
<td>0.09</td>
<td>0.79</td>
<td>1.14</td>
</tr>
<tr>
<td>Kasur ($z_3$)</td>
<td>$\gamma_3$</td>
<td>-0.07</td>
<td>0.10</td>
<td>-0.27</td>
<td>0.11</td>
</tr>
<tr>
<td>Sargodha ($z_4$)</td>
<td>$\gamma_4$</td>
<td>0.60</td>
<td>0.09</td>
<td>0.42</td>
<td>0.79</td>
</tr>
<tr>
<td>Sigma Square</td>
<td>$h$</td>
<td>6.12</td>
<td>0.79</td>
<td>4.73</td>
<td>7.81</td>
</tr>
<tr>
<td>Lambda</td>
<td>$\lambda$</td>
<td>0.50</td>
<td>0.04</td>
<td>0.42</td>
<td>0.58</td>
</tr>
<tr>
<td>Log(Bayesian Factor)</td>
<td>BF</td>
<td>1348.2</td>
<td>215.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>N</td>
<td>730</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2 Estimates of Technical and Mix Efficiency

Based on the Bayesian estimates of the production function (Table 2), we computed farm level technical and mix efficiency scores, which are presented in Table 3. The estimates of mix and technical efficiency
are presented as 2.5% HPD, mean and 97.5% HPD. To illustrate, the 2.5% HPD, posterior mean and 97.5% HPD of input IME for the entire sample are 0.48, 0.65 and 0.83, respectively (Column 2-5 of Table 2). These results indicate that farmers could improve their productivity to about 35 percent of their mix efficiency. However, mean efficiency varies across regions. Farmers in Muzaffargarh district appear to be most mix efficient with an average score of 0.68. However, the average mix efficiency of Lahore is significantly lower than other districts. The observed differentials across agro-climatic zones may reflect problems with supply chain of farm inputs around the large city of Lahore where farmers may compete with industrial activities for common inputs like water and labour supply.

In Table 3 (Column 6-9), technical efficiency estimates show that a further increase in productivity would be possible by using the same input resources. Farmers in Sargodha and Kasur districts show the highest technical efficiency, indicating that farmers in these districts were able to produce the same level of output with fewer input resources. However, like mix efficiency, Lahore district remains the least technically efficient. These differentials in efficiency may be indicative of farmers’ innovation, adoption of new technologies and the use of input mixes due to changing farming operations. Even though the estimates of mean technical and mix efficiency are similar across most agro-climatic zones, they conceal large variations within each agro-climatic zone, which needs further analysis. Figure 4 shows a positive relationship between IME, ITE and farm production.

<table>
<thead>
<tr>
<th>Zones</th>
<th>IME Mean</th>
<th>STD</th>
<th>2.50%</th>
<th>97.50%</th>
<th>ITE Mean</th>
<th>STD</th>
<th>2.50%</th>
<th>97.50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Sample</td>
<td>0.65</td>
<td>0.09</td>
<td>0.48</td>
<td>0.83</td>
<td>0.64</td>
<td>0.15</td>
<td>0.36</td>
<td>0.93</td>
</tr>
<tr>
<td>Kasur</td>
<td>0.64</td>
<td>0.10</td>
<td>0.46</td>
<td>0.83</td>
<td>0.64</td>
<td>0.10</td>
<td>0.55</td>
<td>0.83</td>
</tr>
<tr>
<td>Lahore</td>
<td>0.55</td>
<td>0.09</td>
<td>0.39</td>
<td>0.74</td>
<td>0.60</td>
<td>0.15</td>
<td>0.34</td>
<td>0.90</td>
</tr>
<tr>
<td>Sargodha</td>
<td>0.66</td>
<td>0.08</td>
<td>0.51</td>
<td>0.80</td>
<td>0.66</td>
<td>0.08</td>
<td>0.45</td>
<td>0.80</td>
</tr>
<tr>
<td>Muzaffargarh</td>
<td>0.68</td>
<td>0.10</td>
<td>0.50</td>
<td>0.88</td>
<td>0.68</td>
<td>0.10</td>
<td>0.53</td>
<td>0.88</td>
</tr>
</tbody>
</table>
5.3 Disparities in Technical and Mix Efficiency

The numbers in Table 4 show that significant disparities in mix efficiency exist by farm size. Input mix efficiency monotonically increases as we move from the 10th to 90th percentile, however the mean efficiency varies much less, ranging from 0.63 to 0.67.

Efficiency estimates for farm sizes between two to five acres appear to have the highest range of efficiency, from 0.26 to 0.88, whereas the mean score is 0.66. Moreover, within the farm size of 7–12.5 acres, some farmers are as inefficient as those who own less than two acres.

Table 4: Distribution of farm-level input-oriented mix efficiency

<table>
<thead>
<tr>
<th>Farm Size</th>
<th>All Farms</th>
<th>Area&lt;2</th>
<th>Area 2-5</th>
<th>Area 5-7</th>
<th>Area 7-12.5</th>
<th>Area &gt;12.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.65</td>
<td>0.67</td>
<td>0.66</td>
<td>0.66</td>
<td>0.63</td>
<td>0.66</td>
</tr>
<tr>
<td>SD</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td>Min</td>
<td>0.19</td>
<td>0.50</td>
<td>0.26</td>
<td>0.42</td>
<td>0.19</td>
<td>0.47</td>
</tr>
<tr>
<td>Max</td>
<td>0.88</td>
<td>0.87</td>
<td>0.88</td>
<td>0.77</td>
<td>0.81</td>
<td>0.82</td>
</tr>
<tr>
<td>P10</td>
<td>0.56</td>
<td>0.56</td>
<td>0.57</td>
<td>0.57</td>
<td>0.53</td>
<td>0.57</td>
</tr>
<tr>
<td>P25</td>
<td>0.61</td>
<td>0.62</td>
<td>0.63</td>
<td>0.61</td>
<td>0.58</td>
<td>0.63</td>
</tr>
<tr>
<td>P50</td>
<td>0.66</td>
<td>0.69</td>
<td>0.68</td>
<td>0.66</td>
<td>0.63</td>
<td>0.68</td>
</tr>
<tr>
<td>P75</td>
<td>0.71</td>
<td>0.73</td>
<td>0.72</td>
<td>0.71</td>
<td>0.69</td>
<td>0.70</td>
</tr>
<tr>
<td>P90</td>
<td>0.74</td>
<td>0.76</td>
<td>0.75</td>
<td>0.75</td>
<td>0.73</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Large dispersion in mix efficiency is indicative of inappropriate use of input mixed within different farm sizes. This is not surprising since most small farmers are not able to access inputs when needed the most as they generally rely on local input providers and pay high prices for
low quality inputs. Our survey data also revealed that farmers face high energy and fertiliser costs that might have prevented them using an optimal input mix for their crops. Farmers shared their concerns about timely availability of inputs (including fertilisers and pesticides) in addition to the quality of available inputs (particularly pesticides). Almost 50 percent of the farmers in the sample survey expressed dissatisfaction with the timeliness and effectiveness of inputs. The cost and availability of quality seed was the most serious concern, with 70 percent of farmers dissatisfied with seed quality and cost.

Table 5 presents a similar picture of technical efficiency estimates as that of mix efficiency. However, the disparities in technical efficiency appear to be higher as compared to mix efficiency. Farmers holding fewer than two acres of land experienced the highest disparity in technical efficiency estimates. The bottom 10 percent of smallholders experienced low efficiency scores (i.e., 0.15), implying that they could increase their farm productivity significantly with improved use of the available inputs. In contrast, the top 10 percent of farmers holding two acres of land exhibited significantly higher technical efficiency (i.e., 0.83) compared with the bottom 10 percent. One of the reasons for the considerable differences in efficiency may be lack of access to information on appropriate use of inputs, which is almost non-existent for these smallholders.

Table 5: Distribution of farm-level input-oriented technical efficiency

<table>
<thead>
<tr>
<th>Farm Size</th>
<th>All Farms</th>
<th>Area&lt;2</th>
<th>Area 2-5</th>
<th>Area 5-7</th>
<th>Area 7-12.5</th>
<th>Area &gt;12.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.63</td>
<td>0.54</td>
<td>0.63</td>
<td>0.65</td>
<td>0.62</td>
<td>0.58</td>
</tr>
<tr>
<td>SD</td>
<td>0.17</td>
<td>0.22</td>
<td>0.18</td>
<td>0.16</td>
<td>0.17</td>
<td>0.21</td>
</tr>
<tr>
<td>Min</td>
<td>0.09</td>
<td>0.15</td>
<td>0.16</td>
<td>0.18</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td>Max</td>
<td>0.92</td>
<td>0.86</td>
<td>0.90</td>
<td>0.89</td>
<td>0.92</td>
<td>0.85</td>
</tr>
<tr>
<td>P10</td>
<td>0.38</td>
<td>0.24</td>
<td>0.42</td>
<td>0.45</td>
<td>0.40</td>
<td>0.24</td>
</tr>
<tr>
<td>P25</td>
<td>0.59</td>
<td>0.41</td>
<td>0.57</td>
<td>0.63</td>
<td>0.61</td>
<td>0.55</td>
</tr>
<tr>
<td>P50</td>
<td>0.73</td>
<td>0.67</td>
<td>0.72</td>
<td>0.73</td>
<td>0.73</td>
<td>0.72</td>
</tr>
<tr>
<td>P75</td>
<td>0.79</td>
<td>0.78</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>P90</td>
<td>0.83</td>
<td>0.83</td>
<td>0.84</td>
<td>0.84</td>
<td>0.83</td>
<td>0.81</td>
</tr>
</tbody>
</table>
These results clearly suggest that appropriate input resource allocation could significantly improve the productivity of smallholders. Similarly, farmers could improve productivity with improved extension services through better education, training programs and other supply chain interventions. Furthermore, the disparities in both technical and mix efficiency suggest that there are substantial opportunities to improve smallholders’ productivity by targeting training and extension services programs as well technology adoption within different agro-zones.

5.4 Kernel Densities

To further corroborate our findings, we also draw kernel densities for technical and mix efficiencies to explore if there are significant distributional differentials between different agro-climatic zones as well as technologies. Figure 5 depicts the distribution of technical and mix efficiency for the entire sample and across the agro-climatic zones. As seen in Figure 5(a-e), there is a relatively large dispersion in technical efficiency as compared with mix efficiency across all agro-climatic zones. Moreover, technical efficiency estimates are larger at zonal level and across different technologies. For instance, Panel (a) of Figure 5 shows densities of technical and mix efficiency based on the full sample. The results show clear differences in mix and technical efficiency for the entire sample. As can be seen from the graph, the median technical efficiency is significantly higher than mix efficiency.
Figure 5(a-e): Densities of technical and mix efficiency across different zones

We also plot separate densities for mix efficiency and technical efficiency to see if there are clear differences within each type of efficiency due to adoption of different technologies and
practices, which are depicted in Figures 6 & 7 (a-f). For instance, Figure 6(a) includes plots of mix efficiency for technology adopters vs. non-adopters and we note that the adoption of technology helps to improve mix efficiency of those farmers. Likewise, soil testing leads to improved farm mix efficiency. Figure 7(a-f) also presents density plots of technical efficiency against above mentioned categories. Farmers using tunnel farming and with more contacts with input dealers or other farmers experience clear differences in technical efficiency compared with their peer group with less or no contacts. These results show the importance of networks for diffusing practices and technology between farmers.

The empirical findings indicate that there exists increased variance in efficiency within different farming environments, which needs to be targeted to improve horticultural sector productivity. Our empirical analysis suggests that access to technology, adoption of best practices, appropriate use of input and better access to information through networks could help to improve farm technical and mix efficiency.
Figure 6 (a-f): Comparison of mix efficiency among different technologies
6. CONCLUSIONS

In this paper, we develop an econometric estimation of the components of TFP growth in a Bayesian framework to examine the drivers of TFP growth based on new farm-level survey data of smallholder producers from Pakistan’s horticulture sector. The model consists of
components of TFP, particularly a shortfall in productivity due to mix efficiency, which is a relatively new concept. The farm level survey data were collected from different agro-climatic zones of the Punjab province. We obtained the estimates of mix efficiency and technical efficiency by applying the Bayesian stochastic production frontier, which has the advantage of drawing exact statistical inferences by using finite sample properties. The imposition of the curvature restrictions ensures economically feasible results of the production technology and its associated efficiency measures. We computed the highest posterior densities and confidence intervals at both farm and zone levels.

The results from our model for mix and technical efficiency suggest that a significant gain in productivity can be realized through efficiency improvements. There is an approximate 35 percent shortfall in mix efficiency to suboptimal allocation of input mixes, which indicates that the farmers can increase productivity by using appropriate input mixes. Note however, that there are clear disparities in mix efficiency within agro-climatic zones and different farm-sizes, which may be attributed to access to technology and inputs. The large dispersion in mix and technical efficiency may be the consequence of lack of access to information as well as inappropriate use of input mixes. In other words, farmers are unable to make optimal choices about input utilization and farm practices.

This evidence strongly suggests that increased access to adequate education and training programs could help to reduce technical inefficiency gaps. Training programs may also partially address the need for contact with other supply chain actors, but dedicated programs aimed at broadening access to networks of farmers would help address the need for more diversification. Access to extension services would play an important role in adoption of technologies in the varying environments that can improve horticultural sector productivity.
Furthermore, timely availability of inputs (e.g., fertilizers) and access to financial capital could enable farmers to enhance productivity by using appropriate input mixes.

As mentioned above, the implication for policy is that programs that address input supply problems or improve farm practices need to be informed by productivity analysis because there is a wide variation in efficiency within the farming districts since not all farmers are operating inefficiently. The results for the farming area around Lahore district are interesting because there is a district level difference in efficiency compared to other regions in Punjab. While we have suggested that this might be due to farmers competing for inputs with industries in the city, it warrants further investigation.

A major weakness of this study is that it is not based on longitudinal data, so residual confounding is possible from unmeasured covariates. Moreover, the direction of causality is also uncertain. For example, while we have suggested that farmers with larger networks of suppliers and farmer contacts have better efficiency, it might be argued that more efficient farmers are able to broaden their networks due to wealth effects such as improved standing in the community. Panel data will be needed to clarify these causal effects.
References


CENTAD. (2009). *South Asian Yearbook of Trade and Development: Harnessing Gains from Trade: Domestic Challenges and Beyond*. Retrieved from Centre for Trade and Development:


Gomez, M. I., & Ricketts, K. D. (2013). Food value chain transformations in developing countries: Selected hypotheses on nutritional implications. *Food Policy, 42*, 139-150. doi:10.1016/j.foodpol.2013.06.010


Appendix A

Consider the minimization of input aggregator given in equation (1) subject to the homogenous production technology represented by equation (4), we setup the following Lagrangian:

\[ L = \sum_{k=1}^{K} \gamma_k x_{i_k} + \lambda \left( \ln Q_i - \ln(\delta, z_i) - \sum_{k=1}^{K} \beta_k \ln x_{i_k} - \epsilon_i \right) \]  

(A.1)

The first order conditions are

\[ \frac{\partial L}{\partial x_{i_k}} = \gamma_k - \lambda \frac{\beta_k}{x_{i_k}} = 0 \]  

(A.2)

\[ \frac{\partial L}{\partial \lambda} = \ln Q_i - \ln(\delta, z_i) - \sum_{k=1}^{K} \beta_k \ln x_{i_k} - \epsilon_i = 0 \]  

(A.3)

We take the logarithm of the ratio of first derivatives of the production technology with respect to their inputs (i.e., marginal products) and equate it with the ratio of first derivatives of the input aggregator as:

\[ \ln \left( \frac{x_{i_k}}{x_{i_l}} \right) = \ln \left( \frac{\beta_k \gamma_k}{\beta_l \gamma_l} \right) \text{ for } k = 2, \ldots, K \]  

(A.4)

This can be solved for the factor demand equations

\[ \ln x_{i_k} = \ln \omega_k + \frac{1}{\eta} \ln Q_i + \frac{1}{\theta} \left( \sum_{j=1}^{J} \beta_k \ln \gamma_j - \ln \gamma_i \right) - \frac{1}{\eta} \epsilon_i \]  

(A.5)

where \( \eta = \sum_{k=1}^{K} \beta_k \).

\[ \ln \omega_k = \left( \ln \beta_k - \frac{1}{\eta} \left[ \sum_{k=1}^{K} \beta_k \ln \beta_k \right] \right) \]  

(A.6)

The aggregate input can be obtained by combining equations (A.5 & A.6) and then substituting into equation (A.3) as

\[ \ln X_i = \ln \left[ \sum_{k=1}^{K} \omega_k \right] + \eta^{-1} \ln Q_i + \eta^{-1} \left( \sum_{k=1}^{K} \beta_k \ln \gamma_k \right) - \eta^{-1} \epsilon_i \]  

(A.7)

If there is mix inefficiency then (4) becomes
\[
\ln \left( \frac{x_i}{x_k} \right) = \ln \left( \frac{\beta_i \gamma_i}{\beta_k \gamma_k} + \psi_k \right) \quad \text{for } k = 2, \ldots, K \quad \text{(A.8)}
\]

This error is positive, negative or zero depending on whether the firm over-utilises, under-utilises or correctly utilises input 1 relative to input \( k \). A firm is regarded as being mix efficient if and only if \( \psi_k = 1 \) for all \( k \). The factor demand functions and aggregate input function can be derived from Schmidt & Lovell (1979).

The factor demand function is as follows
\[
\ln x_k = \ln \alpha_k + \eta^{-1} \ln Q_i + \left( \sum_{k=1}^{K} \frac{\beta_k}{\eta} \ln \gamma_k - \ln \gamma_1 \right) + \sum_{k=2}^{K} \frac{\beta_k \psi_k}{\eta} \quad \text{(A.9)}
\]

\[
\ln \hat{X}_i = \ln \left[ \sum_{k=1}^{K} \alpha_k \right] + \frac{\beta}{\eta} \ln Q_i + \frac{1}{\eta} \left( \sum_{k=1}^{K} \beta_k \ln \gamma_k \right) + (E_i - \ln \eta) \quad \text{(A.10)}
\]

where
\[
E_i = \frac{1}{\eta} \sum_{k=2}^{K} \beta_k \psi_k + \ln \left( \beta_i + \sum_{k=2}^{K} \beta_k \exp(-\psi_k) \right) \quad \text{(A.11)}
\]

Therefore,
\[
\hat{X}_i / X_i = \exp(\ln \eta - E_i - u_i / \eta) \quad \text{(A.12)}
\]

**Appendix B**

**Priors**

The following priors are used:
\[
p(\beta, h, u, \lambda) = p(\beta, h) p(u|\lambda^{-1}) p(\lambda^{-1}) \quad \text{(A.13)}
\]

\[
p(\beta, h) \propto h^{-1} \times I(\beta \in R) \times I(h > 0) \quad \text{(A.14)}
\]

\[
p(u|\lambda^{-1}) = \prod_{t=1}^{T} p(u_t|\lambda^{-1}) \quad \text{(A.15)}
\]
\[ p(\lambda^{-1}) = -\ln(\tau) \exp\{\lambda^{-1} \ln(\tau)\} \] (A.16)

where \( I(\cdot) \) is an indicator function that takes the value 1 if the argument is true and 0 otherwise, \( p(u_i | \lambda^{-1}) = f_g(u_i | l, \lambda^{-1}) \) and \( R \) is the region of the parameters space. The combination of the prior with the likelihood function generates the posterior density. The conditional posterior pdfs can be derived by using the likelihood function of equation (14) and combining it with the priors defined by equations (A.13) through (A.16), which are described as:

\[ \beta | h, u, \lambda, y, X \sim N \left( b, h^{-1}(X'X)^{-1} \right) \times I(R) \] (A.17)

\[ h | \beta, u, \lambda, y, X \sim G \left( \frac{T}{2}, \frac{2}{Ts^2} \right) \] (A.18)

\[ \lambda^{-1} | \beta, u, h, y, X \sim G \left( N + 1, \frac{1}{\sum_{i \in I} u_i - \ln(\tau)} \right) \] (A.19)

\[ u_i | \beta, \lambda, h, y, x_i \sim N^+ \left( x_i'\beta - y_i - (h\lambda)^{-1}, h^{-1} \right) \] (A.20)

where, \( b = (X'X)[X'(y + u)] \), \( s^2 = \frac{(y - X\beta + u)'(y - X\beta + u)}{N} \), \( G \) refers to Gamma distribution and \( N^+ \) refers to truncated normal distribution.