Can improved food legume varieties increase technical efficiency in crop production?

A Case Study in Bale Highlands, Ethiopia

Girma T. Kassie, Aden Aw-Hassan, Seid A. Kemal, Luleseged Desta, Peter Thorne, Mulugeta Yitayih
Outline

• Motivation
• The key questions
• Methodology
• Results and discussion
• Conclusion and more questions
Motivation

• Faba bean (broad bean), field pea, and lentil are very important legumes in the highlands of Ethiopia.
  • Ethiopia is the largest producer of faba bean in SSA.
    • In 2012/13, about 4.4 million smallholder farmers planted 574,000 ha of faba bean producing 0.9 million tons at an average productivity of 1.6 tons per ha.
  • Field pea is also an important source of protein in Ethiopia.
    • In 2012, the crop ranked fourth in area coverage with an acreage of 212,890 ha and annual production of 2.6 million tons (FAO, 2012).

• We all agree (it seems) that legumes are essential for the regeneration of nutrient-deficient soils.
  • They fix nitrogen!

• Bale highlands in south Eastern Ethiopia is known for mono-cropping production system: wheat and barley dominated.
Motivation

Mono-cropping:

• Growing one crop year after year on the same plot of land
• Non-diverse rotations – Only a single crop is grown at a time within a field.
• Associated with two problems:
  • Soil degradation
  • Increased vulnerability to risk
• Implies lower efficiency [broadly defined] compared to poly-cropping systems.
The key questions

• Our question(s)
  • How efficient are improved legume growers compared to non-growers?
  • If there is a considerable difference in efficiency, can we attribute this to the inclusion of improved legume crops?
  • Does crop productivity [crop output per unit of the most limiting input] vary between improved food legume growers and non-growers?

• Our objective
  • To empirically show whether the adoption of improved food legume varieties increases the technical efficiency of crop production.
Methodology

• Sampling
  • Multi-stage (mixed) sampling
    • 3 of the 4 major legume producing districts in Bale highlands were selected.
    • 4 peasant associations in each district were selected randomly.
    • 200 farm households from the 4 PAs/District selected using proportionate random sampling.
  • Total sample size 600 farm HHs.
Analytical framework

1. Efficiency analysis (SF Model)

\[ y_i = \alpha + x'_i \beta + \varepsilon_i, \quad i = 2, \ldots, N \]

\[ \varepsilon_i = \nu_i - u_i \]

\[ \nu_i \sim N(0, \sigma^2_v) \]

\[ u_i \sim F \]

\( y_i \) is log of total crop yield
\( x_i \) vector of (log of) inputs
\( \varepsilon_i \) composite error
\( \nu_i \) idiosyncratic error
\( u_i \) inefficiency (one-side disturbance)

The assumption about the distribution (F) of \( u_i \) term is needed to make the model estimable. Four options so far:

1. Half normal distribution \( u_i \sim N^+(0, \sigma^2_u) \) (Aigner et al., 1977).
2. Exponential distribution \( u_i \sim \varepsilon(\sigma_u) \) (Meesun and van den Broeck, 1977).
3. Truncated normal (Stevenson, 1980)
SF Model (2)

- **Two step estimation**
  
  1. estimates of the model parameters \( \hat{h} \) are obtained by maximizing the LL-function \( l(\hat{h}) \)
  
  2. point estimates of inefficiency can be obtained thru the mean (or the mode) of the conditional distribution \( f(u_i | \hat{\epsilon}_i) \)
  
  where \( \hat{\epsilon}_i = y_i - \hat{\alpha} - x_i' \hat{\beta} \)

- **Post-estimation procedures to compute efficiency parameters:**
  
  - Jondrow et al (1982): \( E = \exp(-E(s.u|\epsilon)) \)
  
  - Battese and Coelli (1988): \( E = E[\exp (-s.u|\epsilon)] \)
SFA models

- $Y =$ bread wheat equivalent in birr
  
  - Efficiency model 1
    - *Jondrow et al (1982)*
    - Stoc. frontier normal/tnormal model
  
  - Efficiency model 2
    - *Battese and Coelli (1988)*
    - Stoc. frontier normal/tnormal model
  
  - Efficiency model 3
    - *Jondrow et al (1982)*
    - Stoc. frontier normal/exponential model
  
  - Efficiency model 4
    - *Battese and Coelli (1988)*
    - Stoc. frontier normal/exponential model
Analytical framework

2. Impact analysis

\[ \delta_i = Y_i^A - Y_i^N \]

Where \( \delta_i \) is impact on individual ‘i’;

\( Y_i^A \) is potential outcome of adoption for individual ‘i’.

\( Y_i^N \) is potential outcome of non-adoption for individual ‘i’.

Let \( D \) denotes adoption decision (assumed to be binary) and takes the value 1 for adopters (A) and 0 for non-adopters (N).

Any problem in estimating this?

• YES! - B/C only one of the potential outcomes is observed for each individual \( i \). Missing data problem!!

• Therefore, estimating the individual trt effect \( \delta_i \) is not possible because it is unobservable.
  • Hence we concentrate on (population) average trt effects.
2. Impact analysis

\[ Y_i = \begin{cases} Y_i^A & \text{if } D_i = 1 \\ Y_i^N & \text{if } D_i = 0 \end{cases} \]

\[ ATET = E[\delta_i | D_i = 1] = E[Y_i^A | D_i = 1] - E[Y_i^N | D_i = 1] \]

\[ ATU = E[\delta_i | D_i = 0] = E[Y_i^A | D_i = 0] - E[Y_i^N | D_i = 0] \]

\[ ATE = E[\delta_i] = E[Y_i^A - Y_i^N] = ATT \times P(D = 1) + ATU \times P(D = 0) \]

\[ POM_D = E[Y_i] \]

ATET is identified only if E[Y^N|D=1]-E[Y^N|D=0]=0: That is the TEs of HHs from the adopter and non-adopter groups would not differ in the absence of the improved food legume varieties.
Assumptions for Matching Methods

• Identifying assumption (untestable) – selection on observables (conditional exogeneity)
  – Implies: all the relevant differences b/n treated and non-treated are captured in ‘X’
  – ATT: \( E[Y_{i}^{N} | X, D = 1] = E[Y_{i}^{N} | X, D = 0] \)
  – ATU: \( E[Y_{i}^{A} | X, D = 1] = E[Y_{i}^{A} | X, D = 0] \)
  – ATE: Both

• Common support
  – Implies: We observe adopters and non-adopters with the same characteristics
  – ATT: \( P(D=1 | X) < 1 \)
  – ATU: \( 0 < P(D=1 | X) \)
  – ATE: \( 0 < P(D=1 | X) < 1 \)
Treatment-effects estimators employed

• Adjustment and weighting
  – Regression adjustment [see: Lane and Nelder (1982); Cameron and Trivedi (2005, chap. 25); Wooldridge (2010, chap. 21); and Vittinghoff, Glidden, Shiboski, and McCulloch (2012, chap. 9).]
  – Inverse probability weighting [see: Imbens (2000); Hirano, Imbens, and Ridder (2003); Tan (2010); Wooldridge (2010, chap. 19); van der Laan and Robins (2003); and Tsiatis (2006, chap. 6).]
  – Inverse probability weighting with regression adjustment (IPWRA) [see: Wooldridge, 2007; Wooldridge, 2010]
  – Augmented inverse probability weighting (AIPW) [see: Robins, Rotnitzky, and Zhao (1995); Bang and Robins (2005); Tsiatis (2006) and Tan (2010).]

• Matching estimators
  – Nearest neighbor matching [see: Abadie et al. (2004); Abadie and Imbens (2006, 2011)].
  – Propensity score (treatment probability) matching [See: Rosenbaum and Rubin (1983); Abadie and Imbens (2012)].
Results and Discussion
Description of the sample population

- HHs – 90% male headed and 10% female headed.
- Average land holding/hhd: 2.81 ha (reported)
- On average 37% of the farm plot is allocated to faba bean and 12% for field pea by the sample households.
- Legume producers
  - Faba bean: 50.95%
  - Field pea: 31.37%
  - Faba bean or field pea: 67.76%
- Adopters of improved faba bean and field pea varieties:
  - Improved faba bean or field pea: 23.13%
## Efficiency results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Frontier</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incultland</td>
<td>0.182*** (3.36)</td>
<td>0.17*** (2.93)</td>
<td>0.182*** (3.36)</td>
<td>0.17*** (2.93)</td>
</tr>
<tr>
<td>lnhumanabor</td>
<td>-0.029 (-1.65)</td>
<td>-0.013 (-0.75)</td>
<td>-0.029 (-1.65)</td>
<td>-0.013 (-0.75)</td>
</tr>
<tr>
<td>Lnoxenlabor</td>
<td>0.308*** (7.41)</td>
<td>0.285*** (6.63)</td>
<td>0.308*** (7.41)</td>
<td>0.285*** (6.63)</td>
</tr>
<tr>
<td>Lntotureadapbiofe</td>
<td>0.027** (2.58)</td>
<td>0.034*** (3.25)</td>
<td>0.027** (2.58)</td>
<td>0.034*** (3.25)</td>
</tr>
<tr>
<td>Lnherbic</td>
<td>0.28*** (5.39)</td>
<td>0.27*** (5.28)</td>
<td>0.28*** (5.39)</td>
<td>0.27*** (5.29)</td>
</tr>
<tr>
<td>Lnfungic</td>
<td>0.18*** (4.9)</td>
<td>0.169*** (4.64)</td>
<td>0.18*** (4.9)</td>
<td>0.169*** (4.64)</td>
</tr>
<tr>
<td>Lnmachintime</td>
<td>0.067*** (7.05)</td>
<td>0.063*** (6.71)</td>
<td>0.067*** (7.05)</td>
<td>0.063*** (6.71)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.291** (-2.47)</td>
<td>-0.239* (-1.86)</td>
<td>-0.291** (-2.48)</td>
<td>-0.24* (-1.87)</td>
</tr>
<tr>
<td>( \mu )</td>
<td>-404.25 (-53.75)</td>
<td>-362.15 (-20.57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma_u )</td>
<td>4.76 (34.66)</td>
<td>4.54 (24.49)</td>
<td>-2.49 (-9.62)</td>
<td>-2.71 (-8.81)</td>
</tr>
<tr>
<td>( \sigma_v )</td>
<td>-2.19 (-15.51)</td>
<td>-2.08 (-15.4)</td>
<td>-2.192 (-15.46)</td>
<td>-2.07 (-15.38)</td>
</tr>
<tr>
<td><strong>Statistics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>575</td>
<td>575</td>
<td>575</td>
<td>575</td>
</tr>
<tr>
<td>LL</td>
<td>-332.62</td>
<td>-335.08</td>
<td>-332.61</td>
<td>-335.07</td>
</tr>
<tr>
<td>AIC</td>
<td>687.24</td>
<td>692.16</td>
<td>685.22</td>
<td>690.14</td>
</tr>
<tr>
<td>BIC</td>
<td>735.14</td>
<td>740.06</td>
<td>728.76</td>
<td>733.68</td>
</tr>
</tbody>
</table>
The outcome variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency measure 1</td>
<td>575</td>
<td>0.781</td>
<td>0.108</td>
<td>0.094</td>
<td>0.937</td>
</tr>
<tr>
<td>Efficiency measure 2</td>
<td>575</td>
<td>0.794</td>
<td>0.103</td>
<td>0.101</td>
<td>0.939</td>
</tr>
<tr>
<td>Efficiency measure 3</td>
<td>575</td>
<td>0.782</td>
<td>0.108</td>
<td>0.094</td>
<td>0.937</td>
</tr>
<tr>
<td>Efficiency measure 4</td>
<td>575</td>
<td>0.794</td>
<td>0.103</td>
<td>0.100</td>
<td>0.939</td>
</tr>
<tr>
<td>Total bread wheat eqvt (kcl/ha)</td>
<td>575</td>
<td>2.936</td>
<td>2.993</td>
<td>0.030</td>
<td>29.872</td>
</tr>
<tr>
<td>Total bread wheat eqvt (birr/ha)</td>
<td>575</td>
<td>3.047</td>
<td>3.137</td>
<td>0.030</td>
<td>31.648</td>
</tr>
</tbody>
</table>
Simple comparison of adopters and non-adopters

Kernel density estimate

Technical efficiency via $\exp(-u|e)$ (BC approach)

- Non-adopter
- Adopter

Kernel = epanechnikov, bandwidth = 0.0221

Kernel density estimate

Technical efficiency via $\exp(-E(u|e))$ (JLMS - approach)

- Non-adopter
- Adopter

Kernel = epanechnikov, bandwidth = 0.0241
Simple comparison of adopters and non-adopters

Kernel density estimate

Non-adopter
Adopter

Total bread wheat equivalent (kcal) per hectare

Kernel density estimate

Non-adopter
Adopter

Total bread wheat equivalent (birr) per hectare

Kernel density estimate

Non-adopter
Adopter

Kernel = epanechnikov, bandwidth = 0.4076

Kernel = epanechnikov, bandwidth = 0.4010
### Average trt effect and Average trt effect on the treated

<table>
<thead>
<tr>
<th>Efficiency measure 1</th>
<th>RA</th>
<th>IPW</th>
<th>AIPW</th>
<th>IPWRA</th>
<th>NNMATCH</th>
<th>PSMATCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATET</td>
<td>-.004 (-0.27)</td>
<td>-.003 (-0.25)</td>
<td>-0.006 (-.48)</td>
<td>-0.008 (-.51)</td>
<td>-0.010 (-.69)</td>
<td></td>
</tr>
<tr>
<td>ATE</td>
<td>.002 (0.16)</td>
<td>-.01 (-0.49)</td>
<td>0.002 (0.07)</td>
<td>0.006 (0.56)</td>
<td>0.016 (1.01)</td>
<td>-0.001 (-.04)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Efficiency measure 2</th>
<th>RA</th>
<th>IPW</th>
<th>AIPW</th>
<th>IPWRA</th>
<th>NNMATCH</th>
<th>PSMATCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATET</td>
<td>-.004 (-0.35)</td>
<td>-.003 (-0.3)</td>
<td>-0.006 (-.52)</td>
<td>-0.008 (-.54)</td>
<td>-0.009(-.69)</td>
<td></td>
</tr>
<tr>
<td>ATE</td>
<td>.001 (0.07)</td>
<td>-.01 (-0.52)</td>
<td>0.002 (0.08)</td>
<td>0.005 (0.50)</td>
<td>0.014 (0.95)</td>
<td>-0.002(-.09)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Efficiency measure 3</th>
<th>RA</th>
<th>IPW</th>
<th>AIPW</th>
<th>IPWRA</th>
<th>NNMATCH</th>
<th>PSMATCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATET</td>
<td>-.004 (-0.28)</td>
<td>-.003 (-0.25)</td>
<td>-0.006 (-.46)</td>
<td>-0.008 (-.51)</td>
<td>-0.010 (-.69)</td>
<td></td>
</tr>
<tr>
<td>ATE</td>
<td>.002 (0.16)</td>
<td>-.01 (-0.49)</td>
<td>0.002 (0.07)</td>
<td>0.006 (0.55)</td>
<td>0.016 (1.01)</td>
<td>-0.001 (-.04)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Efficiency measure 4</th>
<th>RA</th>
<th>IPW</th>
<th>AIPW</th>
<th>IPWRA</th>
<th>NNMATCH</th>
<th>PSMATCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATET</td>
<td>-.004 (-0.35)</td>
<td>-.003 (-0.3)</td>
<td>-0.006 (-.52)</td>
<td>-0.008 (-.54)</td>
<td>-0.009(-.69)</td>
<td></td>
</tr>
<tr>
<td>ATE</td>
<td>.001 (0.07)</td>
<td>-.01 (-0.52)</td>
<td>0.002 (0.08)</td>
<td>0.005 (0.49)</td>
<td>0.014 (0.95)</td>
<td>-0.002(-.09)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total bread wheat eqvt (kcl/ha)</th>
<th>RA</th>
<th>IPW</th>
<th>AIPW</th>
<th>IPWRA</th>
<th>NNMATCH</th>
<th>PSMATCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATET</td>
<td>.088 (0.33)</td>
<td>-.067 (-.16)</td>
<td>-2.298 (-.99)</td>
<td>0.337 (1.03)</td>
<td>-0.778 (-1.13)</td>
<td></td>
</tr>
<tr>
<td>ATE</td>
<td>-.109 (-.44)</td>
<td>-.392 (-1.50)</td>
<td>-.194 (-.72)</td>
<td>-.298 (-1.51)</td>
<td>0.117 (0.49)</td>
<td>-0.113 (-.31)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total bread wheat eqvt (birr/ha)</th>
<th>RA</th>
<th>IPW</th>
<th>AIPW</th>
<th>IPWRA</th>
<th>NNMATCH</th>
<th>PSMATCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATET</td>
<td>0.171(0.59)</td>
<td>-.046 (-.09)</td>
<td>-1.180 (-.55)</td>
<td>0.510 (1.59)</td>
<td>-1.034 (-1.08)</td>
<td></td>
</tr>
<tr>
<td>ATE</td>
<td>-.053(-.20)</td>
<td>-.364 (-1.21)</td>
<td>-.151 (-.49)</td>
<td>-.264 (-1.25)</td>
<td>0.215 (0.92)</td>
<td>-0.109 (-.26)</td>
</tr>
</tbody>
</table>
Conclusions and further questions

• Very low adoption of improved legume varieties – particularly faba bean and faba bean.
  • No relationship with efficiency no matter how the latter was measured.
  • No relationship with productivity per unit of limiting factor no matter what conversion [energy or price] was used.
• We observed that complementary inputs are not being used as per the recommendations.
Conclusions and further questions

• Human labor is not rewarding in crop production in the study area. Legumes are still dependent on human labor. Would they have any future in mechanized farming?

• Would simply disseminating the ‘improved varieties’ help? How?

• Bale highlands is known for farmers heavily dependent on machinery for their crop production. How will legumes – produced manually – fit into this system?
  • Are they meant to continue as break crops?
Thank You so much!!