Appendix 5: Regression analysis

Regression analysis

To illustrate the influence of rainfall on harvests, we regressed total harvest per household (kg, year 2014) on age and years of schooling of the household head, planted area (ha), rainfall (mm, year 2014) and the number of dry spells (>7 days without rainfall in the rainy season January-May 2014) (Hanisch 2015).

Results

The overall model was most highly significant ($P_{\text{ANOVA}} < 0.001$) and explained 65.4% of the variance in harvest data. Total planted area ($P = 0.001$), annual rainfall ($P = 0.005$) and numbers of dry spells ($P = 0.002$) were the only significant predictors of total harvest/yr/ha. Stepwise eliminating non-significant predictors yielded the same model. The standardised beta values of the coefficients confirmed that total planted area (0.717) had the strongest influence on harvests, followed by total rainfall in 2014 (0.166). The occurrence of dry spells had a significant negative effect on harvests (-0.155).

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{Model} & \text{R} & \text{R Square} & \text{Adjusted R Square} & \text{Std. Error of the Estimate} \\
\hline
1 & .808^a & .654 & .641 & 811.89605900230700 \\
\hline
\end{array}
\]

a. Predictors: (Constant), rainfall_mm, years of schooling, total planted area (ha), Age (HH head), dry spells (>7days)

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\text{Model} & \text{Sum of Squares} & \text{df} & \text{Mean Square} & \text{F} & \text{Sig.} \\
\hline
1 & Regression & 172809918,809 & 5 & 34561983,762 & 52,432 & .000^b \\
 & Residual & 91625353,577 & 139 & 659175,206 & & \\
 & Total & 264435272,386 & 144 & & & \\
\hline
\end{array}
\]

a. Dependent Variable: harvest (kg)
b. Predictors: (Constant), rainfall_mm, years of schooling, total planted area (ha), Age (HH head), dry spells (>7days)

1 In multiple linear regression, the single predictors contribute in an additive fashion to predict the independent variable. Fundamental agricultural production theory (law of the minimum) requires a non-linear, multiplicative model. In standard Cobb-Douglas production function analysis (Cobb and Douglas 1928), this is achieved by regressing the log of a harvest variable on the logs of socio-demographic, input and environmental variables. To illustrate the influence of rainfall on production in the most simple way, we opted for a simple regression analysis here following a reviewer comment.

2 Concerning the 2013/2014 cropping season, the rainy season started in December 2013 (Hanisch 2015). Unfortunately we don’t have rainfall data for that month.
<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>1 (Constant)</td>
<td>309,694</td>
<td>449,922</td>
<td>.688</td>
<td>.492</td>
</tr>
<tr>
<td>Age (HH head)</td>
<td>-3.125</td>
<td>5.041</td>
<td>-.033</td>
<td>-.620</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>32,061</td>
<td>22,435</td>
<td>.074</td>
<td>1.429</td>
</tr>
<tr>
<td>Total planted area (ha)</td>
<td>357,132</td>
<td>26,511</td>
<td>.717</td>
<td>13.471</td>
</tr>
<tr>
<td>Dry spells (&gt;7 days)</td>
<td>-171,977</td>
<td>60,208</td>
<td>-.155</td>
<td>-2.856</td>
</tr>
<tr>
<td>Rainfall (mm)</td>
<td>1.107</td>
<td>.359</td>
<td>.166</td>
<td>3.081</td>
</tr>
</tbody>
</table>

a. Dependent Variable: harvest kg

LITERATURE CITED
