MEASURING FOOD LOSSES

Session 6:
Loss assessment through modelling
Objectives of the presentation

• Provide guidance on conducting an assessment of losses through modelling

• Present the different models and issues
Introduction

1) Concepts and types of data

2) Regression analysis: general linear regression model

3) Example of Pakistan
Introduction

• Modelling approaches can be used to quantify PHL of certain crops or commodities at various stages of the supply chain

• The objective is to try to estimate what the researcher thinks are the determinants of post-harvest losses at various levels

• A model is a set of assumptions that describes the behaviour of a phenomenon

• It consists of:
  
  ➢ A set of equations describing the behaviour
  ➢ A statement about the errors in the observed values of the variables
  ➢ A specification of the probability distribution of the disturbances
1

Concepts and types of data
1. Concepts and types of data

• The researchers or investigators creating the model should aim to make it representative.

• It needs to be seen as continuous work of the other methods: variables can come from a survey.

• It should contain the salient features of the phenomena under study.

• Variables (parameters) thought to be relevant to explain the phenomenon should be explicitly included in the model.

• The choice of the variables is the most important part:
  - Expert judgements
  - Factor analysis
  - Machine learning methods (random forest, deep learning etc.)
1. Concepts and types of data

- **Times series data:** Information about the numerical values of variables collected over time from period to period.

- **Cross-section data:** Information on the variables concerning individual agents in the supply chain at a given point of time.

- **Panel data:** Data from a repeated survey of a single (cross-section) sample in different periods of time.

- **Dummy variable data:** When the variables are qualitative in nature, then data are recorded in the form of an indicator function.
Regression analysis: general linear regression model
2. Regression analysis: general linear regression model

- Statistical relationship between two or more variables so that one variable (quantitative) can be predicted from the other(s)

- In the case of PHL, Neter & Wasserman (1985) propose:

\[ Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_{p-1} X_{i,p-1} + \epsilon_i \]

- As independent variables are used for the loss assessment, one can choose, for example, the type of seed used, the area planted, the agricultural practices for harvesting (mechanical or traditional), etc. In the case of storage, the dependent variable could be the type of pesticide used, the storage facilities etc.
2. Regression analysis: general linear regression model

Estimation of the model:

Projection of the model:

\[ \widehat{PHL}_i = \hat{c} + \hat{\beta}_1 X_{i,1} + \cdots + \hat{\beta}_K X_{i,K} \]
Example of Pakistan
3. Example of Pakistan

Ahmed et al.

• **Goal:** Quantity PHL of Kinnow (citrus fruit) at various stages of the supply chain (farm, wholesale market and retail levels)

• Fruit but illustrate the modelling aspect

• **Data sources:** Survey from one district in Pakistan using all the sampling procedures
3. Example of Pakistan

Model at the farm level

\[ \ln L_1 = \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \beta_3 \ln X_3 + \beta_4 D_1 + \beta_5 D_2 + \varepsilon \]

Where

- \( L_1 \) = Post-harvest losses of Kinnow in kg
- \( X_1 \) = Education in years
- \( X_2 \) = Experience in years
- \( X_3 \) = Orchard size in Acres
- \( D_1 \) = Dummy variable for picking time (1 if picking time is morning, 0 if picking method is evening)
- \( D_2 \) = Dummy variables for picking method (1 if picked with scissor, 0 if picked manually)
- \( \varepsilon \) = Disturbance term
3. Example of Pakistan

Model at the farm level

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Std. Error</th>
<th>t-value</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>3.839</td>
<td>0.590</td>
<td>6.507</td>
<td>0.000</td>
</tr>
<tr>
<td>Ln X1 (education in years)</td>
<td>-0.211</td>
<td>0.138</td>
<td>-1.526</td>
<td>0.137</td>
</tr>
<tr>
<td>Ln X2 (experience in years)</td>
<td>-0.222</td>
<td>0.108</td>
<td>-2.057</td>
<td>0.048</td>
</tr>
<tr>
<td>Ln X3 (orchard size in acres)</td>
<td>0.214</td>
<td>0.074</td>
<td>2.878</td>
<td>0.007</td>
</tr>
<tr>
<td>D1 (dummy for picking time)</td>
<td>-0.276</td>
<td>0.143</td>
<td>-1.936</td>
<td>0.061</td>
</tr>
<tr>
<td>D2 (dummy for picking method)</td>
<td>-0.477</td>
<td>0.218</td>
<td>-2.187</td>
<td>0.036</td>
</tr>
</tbody>
</table>

Overall fitness:

- $R^2 = 0.406$,
- Adjusted $R^2 = 0.315$,
- $F$-value = 4.5 at 5% degree of freedom
3. Example of Pakistan

Model at the wholesale level

\[ \ln L2 = \alpha_0 + \alpha_1 \ln Y1 + \alpha_2 \ln Y2 + \alpha_3 D3 + \alpha_4 D4 + \alpha_5 D5 + \mu \]

Where

- \( L2 \) = Quantity of post-harvest losses in kg
- \( Y1 \) = Education in years
- \( Y2 \) = Experience in years
- \( D3 \) = Dummy variables for Infrastructure of transportation (1 if roads were metalled, 0 if roads were non-metalled)
- \( D4 \) = Dummy variable for loading method (1 if produce was loaded in boxes, 0 if produce was openly loaded)
- \( D5 \) = Dummy variable for storage place (1 if cold storage, 0 if normal storage)
- \( \mu \) = Disturbance term
3. Example of Pakistan

Model at wholesale level

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Std. Error</th>
<th>t-value</th>
<th>Sig.</th>
<th>Overall fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>4.808</td>
<td>0.461</td>
<td>10.437</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Ln Y1 (education in years)</td>
<td>-0.154</td>
<td>0.115</td>
<td>-1.335</td>
<td>0.191</td>
<td></td>
</tr>
<tr>
<td>Ln Y2 (experience in years)</td>
<td>-0.272</td>
<td>0.140</td>
<td>-1.944</td>
<td>0.060</td>
<td></td>
</tr>
<tr>
<td>D3 (dummy for Infrastructure of transportation)</td>
<td>-0.593</td>
<td>0.390</td>
<td>-1.521</td>
<td>0.137</td>
<td></td>
</tr>
<tr>
<td>D4 (dummy for loading method)</td>
<td>-0.555</td>
<td>0.273</td>
<td>-2.031</td>
<td>0.050</td>
<td>R2 = 0.68, Adjusted R2 = 0.63, F-value = 14.48 at 5% degree of freedom</td>
</tr>
<tr>
<td>D5 (dummy for storage place)</td>
<td>-0.562</td>
<td>0.293</td>
<td>-1.916</td>
<td>0.064</td>
<td></td>
</tr>
</tbody>
</table>
3. Example of Pakistan

Model at retail level

\[ \ln L3 = \gamma_0 + \gamma_1 \ln Z1 + \gamma_2 \ln Z2 + \gamma_3 D6 + \epsilon \]

Where

- \( L3 \) = Quantity of post-harvest losses in Kg
- \( Z1 \) = Experience in years
- \( Z2 \) = Unsold quantity on daily basis
- \( D6 \) = Dummy variable for type of retailer (1 if respondent was a shopkeeper, 0, if respondent was a hawker)
- \( \epsilon \) = Disturbance term
3. Example of Pakistan

Model at retail level

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Std. Error</th>
<th>t-value</th>
<th>Sig.</th>
<th>Overall fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>0.453</td>
<td>0.317</td>
<td>1.429</td>
<td>0.162</td>
<td></td>
</tr>
<tr>
<td>Ln Z1 (experience in years)</td>
<td>-0.080</td>
<td>0.108</td>
<td>-0.738</td>
<td>0.466</td>
<td>R2 = 0.62</td>
</tr>
<tr>
<td>Ln Z2 (unsold quantity in kgs)</td>
<td>0.259</td>
<td>0.135</td>
<td>1.921</td>
<td>0.063</td>
<td>Adjusted R2 = 0.59</td>
</tr>
<tr>
<td>D6 (dummy variable for type of retailer)</td>
<td>-1.320</td>
<td>0.266</td>
<td>-4.959</td>
<td>0.000</td>
<td>F-value = 18.74</td>
</tr>
</tbody>
</table>
Conclusion

• This presentation described the main approaches and strategies to model losses

• The most important aspect in modelling is the choice of the variables used to explain the phenomenon

• A model can be designed for all the stages of the supply chain if data are available
References


Thank You