Estimation and other considerations

Dimension compendium and a glance at project analysis
Outlook

- Validating data and model
- Time lag of micro databases
- Aging data
- Other use for model output
- The thing in common
Recap of sampling

- Creating a Sample
- Macro Modeling
- Micro Modeling
Validating data and model

**Data:**
- Variance tests cannot realistically be done on every variable.
- Imputation will be needed to help clean the data.
  - Simple tests such as: Min, Max, comparison of aggregate values over time, etc.

**Model:**
- Compare results on similar or dissimilar dimensions.
  - Example: Create average tax rates of small and large firms (assuming a different statutory tax rate).
Validating data and model

Example:

- One good way to clean the data is to sort the database on the variable analysed.

<table>
<thead>
<tr>
<th>CORP_NAME</th>
<th>WEIGHT</th>
<th>INDUSTRY</th>
<th>ASSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFS Corporation Inc</td>
<td>8.3</td>
<td>02. Manufacturing</td>
<td>999,962,344,024</td>
</tr>
<tr>
<td>QCP Corporation Inc</td>
<td>8.3</td>
<td>12. Oil,Gas</td>
<td>59,084,834</td>
</tr>
<tr>
<td>IKB Corporation Inc</td>
<td>8.3</td>
<td>10. Other Finance</td>
<td>51,473,460</td>
</tr>
<tr>
<td>GLE Corporation Inc</td>
<td>8.3</td>
<td>12. Oil,Gas</td>
<td>50,652,449</td>
</tr>
<tr>
<td>JXM Corporation Inc</td>
<td>8.3</td>
<td>13. Mining</td>
<td>46,333,786</td>
</tr>
</tbody>
</table>
Time lag of micro databases

- How to deal with the time lag that (inevitably) occurs when dealing with micro data?
  - Best alternative is to find ways to shorten the capture process
  - Updating data with current information (receipts)
  - If analysis is specific enough, get information from the industry or your revenue agency.
    - Sectoral updates give better results. Appropriate dimensions need to be available.
  - Do post adjustments of simulation results
Aging micro data

- Micro data can be aged prior to simulation
  - Inflation adjusted
  - GDP adjusted
  - Other
- Ideally, use disaggregated data if available from forecasting group (use of dimensions is key)
- Important to validate results on macro variables
Aging micro data

- Can use a mix of indexes to age data
  - GDP Economic activity
  - Price index on specific types of investment or sectors (again, *dimensions* are key)

- Aging micro data can be tricky and difficult, especially if many variables of many types are involved
  - Level, price, income, stock, etc
Forecasting tax changes

• Aging results using macro variables
  Doing post adjustments
  – Easier to do and understand (and thus explain)
  – Bias are easier to identify
  – Turnover time is short
  – **Dimensions** will help improve results
Other use for model output

• Evaluating Average Effective Tax Rates (AETR)
  – Impact of tax changes on AETRs
    • Profitable corporations (Dimensions)
  – Comparing domestic AETRs with another country (removing data bias - simulation on same database of different tax systems)

• Costing Tax incentives
Other use for model output

• Help identify input errors in database
  – Doing a simulation of current system on current data should help identify major errors

• Micro modelling allows doing more detailed modelling on identified sectors of the economy (e.g. Banks) - **Dimensions**
The thing in common

• DIMENSIONS
  – When creating a dataset, it is **crucial** to seriously look at the dimensions you will need to:
    • Validate your data
    • Validate your model
    • Compare with other source of information (e.g. OECD)
    • Do forecasting
Project analysis

• Industry
  – Evaluate rate of return on investment

• Policy makers
  – Analyse impact of fiscal parameters on specific projects
  – Measure attractiveness of national fiscal position for specific investments (international comparison)
Project analysis

- Most project analysis are done in a spreadsheet environment
- Can involve complex interactions
- Risks
  - Variability of parameters over time (such as prices)
  - Opaque coding language
Project analysis

• Minimizing spreadsheet risks
  – Documentation
  – Monte-Carlo Simulation

• Monte-Carlo Simulation
  – INPUT: Distribution pattern of uncertain parameters
  – PROCESS: Multiple runs (iterations)
  – OUTPUT: Mean and distribution profile