APPENDIX

APPENDIX A – Some Inequality and Poverty Indices

Here, we will give details of some inequality and poverty measures we have used during the analysis.

Gini index

The Gini coefficient is one of the most commonly used indicators of income inequality. It is defined as:

$$G = \frac{1}{2\mu N^2} \sum_i \sum_j \left| y_j - y_i \right|$$

where μ is the arithmetical mean of the incomes, N is the size of the population, and y_i and y_j are the incomes of agents *i* and *j*, respectively. Thus, the second factor at the right hand side represents the sum of the differences (in modulus) computed over all pairs of incomes. In the literature, however, we can also find different (although equivalent) definitions. In particular, it can be derived from the Lorenz curve, which plots the cumulative share of total income earned by households ranked from bottom to top (see below), in the following way:

$$G = 1 - 2\int_0^1 L(p)dp \, , \qquad \textit{(A.1)}$$

where $L(p)$ is the Lorenz curve. The previous formula thus measures the area that is laying between the curve and the diagonal as a fraction of the total area under the 45° line. In terms of Figure A.1 below, this means:

$$G = \frac{A}{A+B} = \frac{\frac{1}{2} - B}{\frac{1}{2}} = 1 - 2B \, .$$

If the Lorenz curve coincides with the 45° line, which represents the situation of perfect equality, then the integral in equation *(A.1)* will take the value of ½, and the Gini index will equal zero.

The Gini index can thus take values between zero (perfect equality) and one (maximum level of inequality, that is, when all the income in the economy is owned by only one individual: $y_{\max} = \mu N$). Thus, the smaller is the index, the smaller is the inequality in the economy. The Gini index is very useful because it allows the ordering of different income distributions according to their level of inequality.

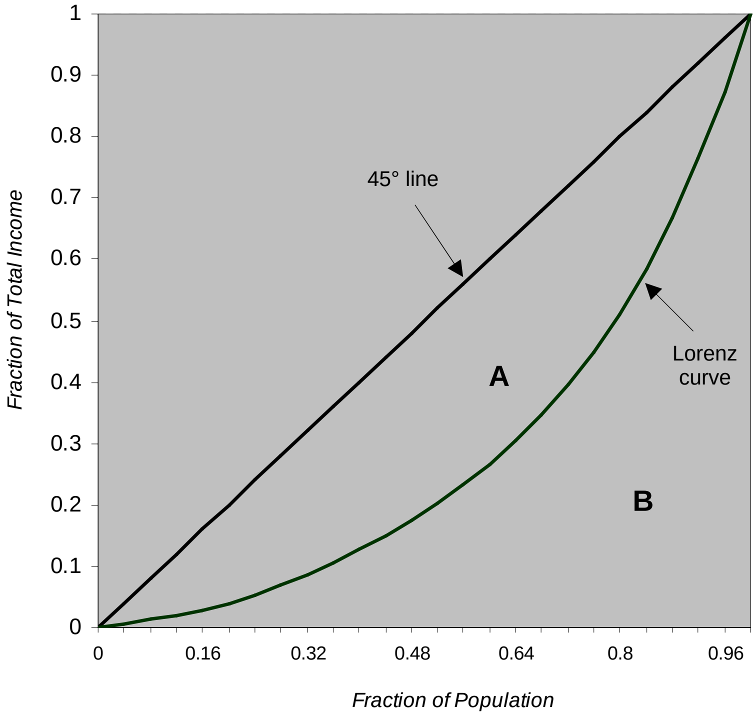

Figure A.1 – Lorenz Curve and Gini Coefficient

Lorenz curve

The Lorenz curve is defined according to the following expression:

$$L(p) = L(j/N) = \frac{\sum_{i=1}^{j} y_i}{\sum_{i=1}^{N} y_i}, \qquad \text{for } j = \{1, 2, \ldots, N\}.$$

In the continuous the expression becomes:

$$p = F(y) \Rightarrow L(p) = \int_0^y \frac{x \cdot f(x)}{\mu} dx ,$$

where y is the income in the point we want to compute the curve, p is the cumulated probability of income, and μ is mean income.

This curve therefore shows the relationship between a particular percentage of the population, say $j = h/N$, and the proportion of total income that it perceives. Based on an analysis of the stochastic dominance of Lorenz curves, one could eventually infer, among different economies, which one has a more equitable distribution.

Atkinson's index

Atkinson's index is one of the few inequality measures that explicitly incorporate normative judgments about social welfare (Atkinson, 1970). The index is derived by calculating the so-called equity-sensitive average income (y_e), which is defined as that level of per capita income which if enjoyed by everybody would make total welfare exactly equal to the total welfare generated by the actual income distribution. It is sometimes also called equally distributed equivalent income. It is given by:

$$y_e = \left[\frac{1}{N} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\mu} \right)^{(1-e)} \right]^{1/(1-e)},$$

where y_i is the proportion of total income received by individual i, and e is the so-called inequality aversion parameter, which measures the degree of society's inequality aversion. It indeed reflects the strength of society's preference for equality, and can take values ranging from zero to infinity. When $e > 0$, there is a social preference for equality (or an aversion to inequality). As e rises, society attaches more weight to income transfers at the lower end of the distribution and less weight to transfers at the top. $e \to 0$ implies neutrality with respect to inequality, so that inequality is not perceived as a problem. Suppose instead that $e \to \infty$, then it means that there are Rawlsian preferences in the society, that is, that individuals have a preference for perfect equality. Typically, in the literature the most common values that are used for e include 0.5 and 2.

The Atkinson index (I_e) is then given by:

$$I_e = 1 - \frac{y_e}{\mu},$$

where μ is the actual mean income. The more equal the income distribution, the closer y_e will be to μ, and the lower the value of the Atkinson index. For any income distribution, the value of I_e lies between 0 and 1.

<u>Coefficient of variation</u>

The coefficient of variation is a measure of the dispersion of data around the mean. It is defined as the ratio of the standard deviation to the mean, that is:

$$CV = \frac{\sigma}{\mu}.$$

The coefficient of variation is a dimensionless number that allows comparison of the variation of populations that have significantly different mean values. It is often reported as a percentage (%) by multiplying the above calculation by 100.

<u>Generalized Entropy coefficients</u>

The family of Generalized Entropy indices satisfies a desirable property for inequality indices, that is, all the indices belonging to this family can be decomposed into a within-group and a between group contribution. The formulas for the indices are:

Generalized entropy index: $$I(c) = \frac{1}{Nc} \cdot \frac{1}{(c-1)} \sum_{i=1}^{N} \left[\left(\frac{y_i}{\mu} \right)^c - 1 \right]$$ for $c \neq 0,1$

Mean Logarithmic Deviation: $$I(0) = \frac{1}{N} \sum_{i=1}^{N} \ln \left(\frac{\mu}{y_i} \right)$$ for $c = 0$

Theil coefficient: $$I(1) = \frac{1}{N} \sum_{i=1}^{N} \frac{y_i}{\mu} \cdot \ln \left(\frac{y_i}{\mu} \right)$$ for $c = 1$

Parameter c reflects different perceptions of inequality, with lower values indicating a higher degree of inequality aversion. A value of c greater than one means that differences at the high end of the welfare distribution are assigned more importance than those at the low end.

For the second index, known as Mean Logarithmic Deviation, a value of zero represents perfect equality and higher values denote increasing levels of inequality, within a given administrative unit. The parameter value 0 means that differences at the low end of the welfare distribution are assigned more importance than those at the high end.

Finally, Theil coefficient (or "information theory" measure) has a potential range from zero to infinity, with higher values (greater entropy) indicating more unequal distribution of income. If instead everyone has the same (i.e., mean) income, then the index equals 0. If one person has all the income, then the index is equal to ln(*N*). The parameter value 1 means that differences are equivalently treated at all points in the welfare distribution.

The Theil index has the advantage of being additive across different subgroups or regions in the country. Indeed, it is the weighted sum of inequality within subgroups. For example, inequality within the United States is the sum of each state's inequality weighted by the state's income relative to the entire country.

If the population is divided into m certain subgroups and s_k is the income share of group k, T_k is the Theil index for that subgroup, and μ_k is the average income in group k, then the Theil index of the population is:

$$T = \sum_{k=1}^{m} s_k \cdot T_k + \sum_{k=1}^{m} s_k \cdot \ln \frac{\mu_k}{\mu}$$

Therefore, one can say that a certain group "contributes" a certain amount of inequality to the whole.

Poverty indices

We will give details of the poverty indices we have used during the analysis.

Foster, Greer and Thorbecke (1984) have suggested a useful class of poverty indices that takes the following form:

$$P_\alpha = \frac{1}{N} \cdot \sum_{i=1}^{q} \left[\frac{(Z_p - Y_i)}{Z_p} \right]^\alpha \quad ,$$

where Z_p denotes the poverty line, Y_i the expenditure or income of the i-th poor household (or individual), N the total number of households and q the number of households whose expenditures or incomes are below the poverty line. Of course, the choice of the poverty line is of great importance in the determination of the index, and it may reflect different judgements about the researcher's choice for an appropriate level of welfare.

From the general formula above, one can compute different kinds of poverty measures by simply varying the value of α:

- If $\alpha = 0 \Rightarrow P_0 = \frac{q}{N}$

P_0 is also called "Headcount ratio", as it measures the incidence of poverty as the proportion of total population lying below the poverty line.

- If $\alpha = 1 \Rightarrow P_1 = \frac{1}{N} \cdot \sum_{i=1}^{q} \frac{(Z_p - Y_i)}{Z_p} = IP_0$

This index gives a good measure of the intensity of poverty, as it reflects how far the poor are from the poverty line. Indeed, it quantifies the extent to which the income of the poor lies below the poverty line. Hence the reason why it is also called "Income or Poverty gap ratio".

- If $\alpha = 2 \Rightarrow P_2 = \frac{1}{N} \cdot \sum_{i=1}^{q} \left[\frac{(Z_p - Y_i)}{Z_p} \right]^2$

This measure is also known as "Poverty Severity Index", as it gives an indication of the degree of inequality among the poor. The greater is the inequality of distribution among the poor and thus the severity of poverty, the higher is P_2.

APPENDIX B – Structure of Production and Foreign Sector

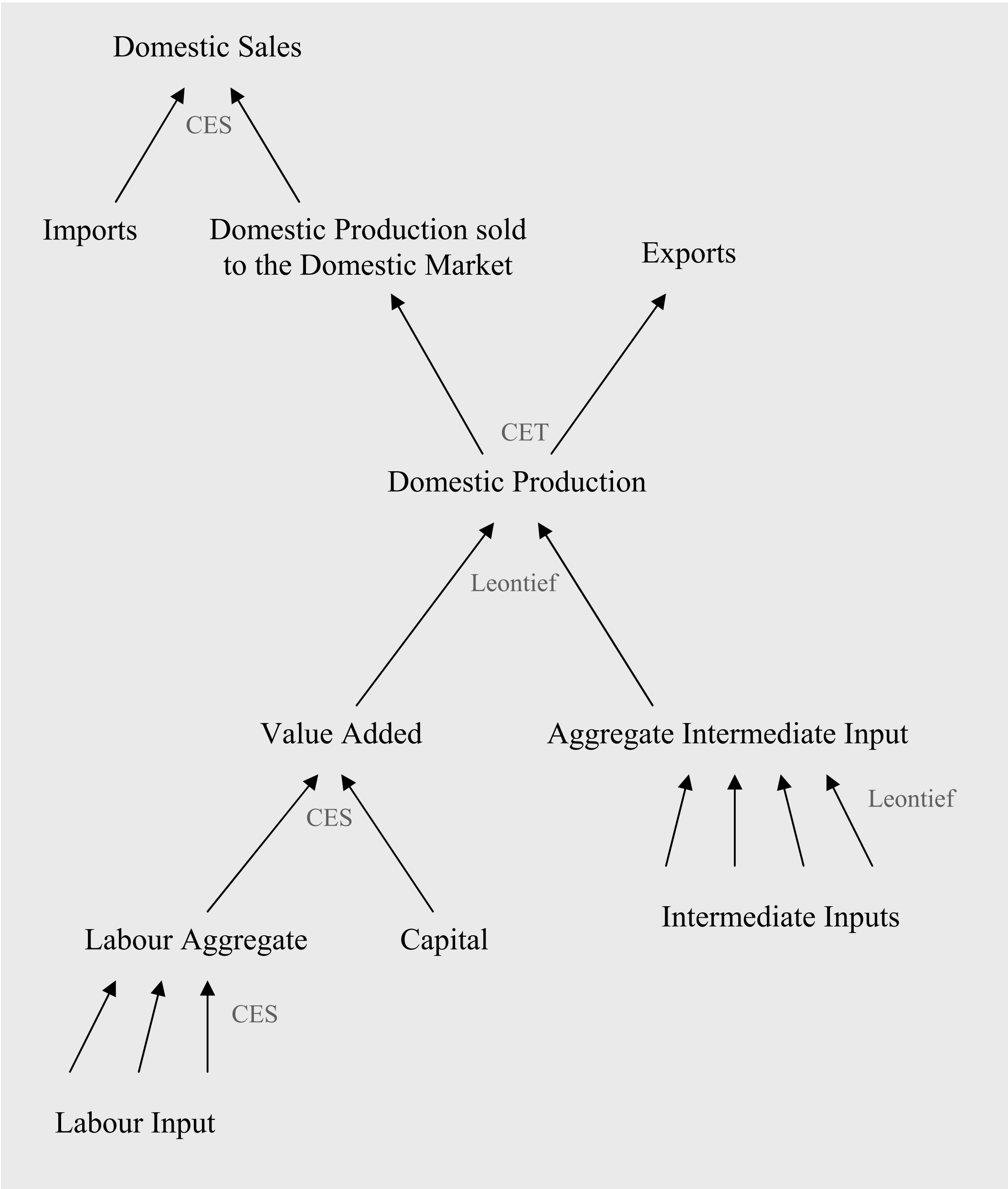

Figure B.1 – The Structure of Production and the Foreign Sector