

## Answers to A1 Exam of Social Economics and Public Policy 2022

Professor: Marcelo Neri

TA: Pedro Mencarini

**Time: 3 hours. Please handle in the questions sheet signed with your answers. In the last page there is a formula sheet. You don't need to use a calculator. Answers written in pencil are not subject to revision. The Exam has three main questions with different parts and plenty of choices, write only the number of items suggested. Please read and number your answers carefully. You may write your answers in Portuguese or in English. Correction Criteria: we point in the end of each item answer its strict correction criteria, Upward adjustments will be given on top of the final grades obtained.**

**Question I - Evaluate if each sentence is True or False. Don't need to justify your answers here.** Please answer 16 (and only 16) of the total 24 items below (A Total 56 points – 3.5 points for each selected item)

### Inequality

1. The Lorenz curve is a particular case of both the concentration curve and the generalized Lorenz curve. T
2. The minimum possible levels of the Gini Index and the concentration ratio coincide. F
3. Sources of mean income growth can be decomposed by income sources from changes in their respective concentration indexes. F
4. Inequality of years of schooling is not only high but also has been rising in Brazil in the last 40 years. F
5. Individual levels of years of schooling and Education of the parents are strictly related to inequality of circumstances in an inequality of opportunities framework. F
6. The inequality index associated with the bottom 40% income (shared prosperity) is consistent with mean and inequality decomposition and also with the Pigou-Dalton principle of transfers. F
7. Inequality of per capita household income tends to be lower than the one based on individual income for the same population. T
8. In general, the imputation of top Incomes in household surveys tends to increase income inequality, its mean and social welfare levels. T

### Poverty

9. The contribution of the indigenous population to aggregate poverty is higher than the one for the white population in Brazil. F

10. The incorporation of economies of scale in poverty analysis tends to raise poverty measures in comparison to measures based on plain per capita household income. F
11. Gender biases issues tends to appear more on individual earnings distribution than on per capita household income. T
12. The Mean Poverty Gap ( $P^1$ ) captures basically the distance of the poor income with respect to the poverty line. F
13. If the poverty line is raised the minimum cost of overcoming poverty rises more than proportionately. T
14. The counterfactual of inequality changes in Datt-Ravalion Poverty Decompositions is obtained by comparing initial poverty levels with the one found in later period distribution divided by the growth factor. T
15. Third order dominance implies that all poverty measures ( $P_0$ ,  $P_1$  and  $P_2$ ) for all poverty lines are always higher for one of the distributions involved. F
16. One advantage of an Universal Basic Income scheme is to minimize the cost of overcoming poverty. F

#### Social Targets, Polarization and Global Social Indicators issues

17. Poverty targets based on  $P^1$  (mean Poverty Gap) have difficulty in inducing actions aimed at the poorest in society. T
18. One of the strong aspects of the Multidimensional Poverty Index (MPI) is the weight structure attributed to different indicators. F
19. Relative economic classes segmentation that divides initially the society into equal sizes classes, has a greater explanatory power to total Brazilian inequality (using the Theil Index metric  $T_e / T$  contribution) than absolute economic classes segmentation using EGR (Esteban, Gradin and Ray) methodology. F
20. A rise on mean life expectancy with its inequality kept constant affects both the Human Development Index (HDI) and the Inequality Adjusted Human Development Index (IHDI). T
21. A system of targets based on international indicators such as SDGs (Sustainable Development Goals) allows to lengthen the planning horizons of policy makers and generate a common ground for actions between different levels of government. T
22. A system where the greater the initial poverty the greater the social transfer from the federal government to a region leads to an increase in non-social spending. T
23. In the presence of aggregate shocks one should use rankings of social indicators. T
24. Subjective indicators derived from questions on perceptions are subject to cultural influences and the problem of adaptation across time at the individual level. T

**Question II – Comment briefly the statement below. Justify shortly your answers also using graphs or formulas, if possible.** Please answer 4 (and only 4) of the 7 items below (Total 24 points – 6 points for each selected item). A: All are true

1. Falling unemployment rates, raising labor market participation rates and worked hours all contribute to rising mean labor earnings. So does raising hourly earnings. One can use this simple framework to also study since is multiplicative y earnings inequality trends.

**Answer:**



(1) Show the multiplicative form of classic labor ingredients in levels  $A \cdot B \dots Z = \text{Mean Earnings}$ , like the figure above, and (2) Show temporal changes (this in logarithmic approximation  $(\ln A_2 - \ln A_1) + (\ln B_2 - \ln B_1) + \dots + (\ln Z_2 - \ln Z_1) = \text{total change}$  or in rates of change without approximation  $(1+a) \cdot (1+b) \dots (1+z) = (1+\text{total change})$  where  $a$  is the rate of change of  $A \dots$ ). (3) This can be applied to different social welfare functions from Gini based to shared prosperity and so on. One can decompose social welfare into mean and inequality multiplication of labor ingredients levels and its rates of logarithmic changes as suggested below.

$$\mu_S = \mu(1 - I)$$

$$\ln(\mu_S) = \ln(\mu) + \ln(1 - I)$$

Which on taking the first difference gives:

$$\gamma^* = \gamma + g$$

To be sure each social welfare component can be decomposed in its respective mean and inequality components either in levels or rates of changes (exact or logarithmic approximation).

**Correction Criteria:**

- (1) 3 Points;
- (2) 2 Points;
- (3) 1 Point;

2. The R-Square of a log-linear regression can be used to calculate the gross and the net contribution of a particular variable to overall inequality.

**Answer:** (1) Explain how the R-Squared from a regression with constant and the variable of interest corresponds to the gross contribution to inequality. (2) While the difference of R-Squares of a full regression with all desired controls and the same regression without the variable of interest correspond to the net contribution to inequality of this particular variable. (3) Bonus: mention that it is Variance of Logs and the possibility of using R-Squared bar.

### Mincerian Model and the Variance of Logs\*

#### Gross Contribution to Income Inequality (%) – R<sup>2</sup> of an equation CTE + VAR\*

Ex: in the case of education:  $R^2$  of  $\ln(w_i) = \beta_0 + \beta_1 \text{Schooling}_i + \varepsilon_i$ .

#### Net Contribution to Income Inequality (%) \*

% difference between the R<sup>2</sup> of the full regression and the R<sup>2</sup> of the regression without the specific variable. That is:

$(R^2 \text{ of } \ln(Y_i) = \alpha + \beta S_i + x_i' \gamma + \varepsilon_i) - (R^2 \text{ of } \ln(Y_i) = \alpha + x_i' \gamma + \varepsilon_i)$

#### Correction Criteria:

- (1) 2.5 Points;
- (2) 2.5 Points;
- (3) 1 Point;

3. The use of interactive dummies in a regression framework allows to estimate difference in difference between treatment and control groups before and after an intervention.

**Answer:** Show the diff-in-diff estimator with the interactive variables in levels and in interaction.

#### Difference in difference estimator

In economics, vast research is done analyzing the so-called experiments or quasi-experiments. To analyze a natural experiment, it is necessary to have a control group, that is, a group that was not affected by the change, and a treatment group that was directly affected by the event of interest, both with similar characteristics. In order to study the differences between the two groups, pre and post-event data are needed for both groups. Thus, the sample is divided into four groups: the pre-change control group, the post-change control group, the pre-change treatment group, and the post-change treatment group. 1) The difference between the differences between the two periods for each of the groups is the difference in difference estimator, represented by the following equation:

$$g_3 = (y_{2,t} - y_{1,t}) - (y_{2,c} - y_{1,c})$$

2) Where each  $y$  represents the mean of the studied variable for each year and group, with the subscript number representing the sample period (1 for before the change and 2 for after the change) and the letter representing the group to which the data belongs ( $c$  for the control group and  $t$  for the treatment group).  $g_3$  is the so-called difference in difference estimator. Once the  $g_3$  is obtained, the impact of the natural experiment on the variable to be explained is determined.

3) Mathematically, we can represent this difference-in-difference estimator (D-D) used from equations in discrete or continuous variables (for example, in the case of logistic regressions or mincerian-type per capita income equations) with regression model like the one bellow.

$$Y_i = g_0 + g_1 d_2 + g_2 T_i + g_3 d_2 T_i + \text{other controls} + \varepsilon_i$$

Where  $T_i$  is a dummy that indicates if the individual was treated ( $T_i = 1$ ) or not ( $T_i = 0$ )  $d_2$  is a dummy that indicates the time (when  $d_2 = 1$  the time is after the treatment, when  $d_2 = 0$  the time is before the treatment),  $Y_i$  is an endogenous variable and  $\varepsilon_i$  is a

error term. Notice that the coefficient of the interaction (multiplication) between the treatment dummy and the time dummy ( $g_3$ ) gives us the difference-in-difference estimator.

**Correction Criteria:**

- (1) 2 Points;
- (2) 2 Points;
- (3) 2 Points.

4. A regression allows to isolate the partial correlations between the endogenous variable of interest and a set of explanatory variables. It is possible to capture non linearities such as increasing returns and externalities both in the household reference person years of schooling, for example.

**Answer:** (1) Increasing returns can be captured by a quadratic term of years of schooling (or a higher order polynomial). The mean labor earnings of years of schooling when one passes from 0 to 4 years of schooling raises 8% per year of schooling. While from 11 to 15 years of schooling raises 21% per year. The regression should show a positive second order term. (2) Externalities is captured by the mean years of schooling of a community. This last term captures the fact that people benefit from the proximity of better educated people as professors, nurses, social assistance workers etc.

*Mincerian Model:*  $y_i = \ln(Y_i) = \alpha + \beta S_i + GSS_i + LSC_j + x_i'\gamma + \varepsilon_i$

Where  $Y_i$  is the labour income of individual  $i$  (we change this metric below),  $S_i$  is the level of education of individual  $i$  measured by years of schooling,  $SS_i$  is the Squared level of education of individual  $i$  measured by years of schooling,  $LSC_j$  is the mean level of education of community  $j$  measured by years of schooling  $x_i$  is a vector of controls and  $\varepsilon_i$  is an error term.

**Correction Criteria:**

- (1) 3 Points;
- (2) 3 Points;

5. The Gini index based social welfare function gives more weight to the poor. While a pure logarithmic income specification assumes higher impacts at low income levels. Both approaches are not inconsistent with each other.

**Answer:** (1) Departing from a general social welfare function show Gini social welfare function weights:  $2(1 - F(x))$  and the pure logarithmic income specification  $u(x) = \ln(x)$ . Combine both in a specific social welfare function with both features nicknamed as Lini (mix of Logs individual social welfare function and Gini weights).

(2) *Kakwani et al. (2010) Lini Social Welfare Function:*

$$\log(x^*) = 2 \int_0^{\infty} [1 - F(x)] \log(x) f(x) dx$$

*Derived Inequality Measure from a log utility and Weights a la Gini = Lini:*

$$\log(I) = 2 \int_0^{\infty} [1 - F(x)] [\log(\mu) - \log(x)] f(x) dx$$

(3) Gini will fall in Brazil with any income increase below the 75th percentile, the Lini is more pro poor.

**Correction Criteria:**

- (1) 2 Points;
- (2) 3 Points;
- (3) 1 Points;

6. Total inequality movements are not always followed by polarization measures movements in the same direction.

**Answer:** (1) If the between groups Gini (of those above and below median income) raise so does the polarization (alienation aspect) but if the inequality within these groups raise (the identification aspect) polarization falls although overall Gini rises. The last formula below illustrate the point.

**Under Polarization:** *social welfare becomes:*

$$W_B = \int_0^{\infty} u(x)v(x)f(x)dx = \mu - (m_2 - m_1) + 2\mu G$$

$$\text{Or, } W_B = \mu - 2(G_B - G_W)$$

(2) And/or using an example: Consider a simple society with six people called A, B, C, D, E and F, with incomes of R\$ 6, 5, 4, 3, 2 and 1, respectively. Suppose that one Real is transferred from D to F and from A to C. Inequality indices that respect the so-called principle of transfers will necessarily decline. After these distributive changes, we will have a perfectly divided society in two internally homogeneous groups: an income of R\$ 2 for D, E and F and an income of R\$ 5 for A, B and C. Although less unequal, after these progressive transfers, society has become more polarized.

**Correction Criteria:**

- (1) + (2) 6 Points;
- (1) or (2) 5.5 Points.

7. The first Sustainable Development Goal (SDG) related to extreme poverty eradication (measured by the proportion of poor by 2030) presents pitfalls as a social target mechanism.

**Answer:** (1)  $P^0$  favors the least poor of the poor. The easiest way is to give an additional a small epsilon for those very close the poverty line. The fixed date could lead to a transfer only at the moment that the target is due (neither before nor after). (2) The ideal would be to use the present value of  $P^2$  along all the poverty path. (3) Bonus: Zero poverty is also unrealistic if people are entering and leaving poverty continuously (rotating poverty movements). In practice people use 3% poverty as reaching this minimum targeting point.

**Correction Criteria:**

- (1) 3.5 Points;
- (2) 1.5 Points;
- (3) 1 Point.

**Question III- Write a small essay on one (and only one) of the three topics below.** (Total 20 points):

1. What are the main challenges to measure income inequality in Brazil?

**Answer:** Go beyond gross income (that bases surveys as PNAD or PNADC): (1) incorporate the roles of direct and indirect taxes as does a microsimulation framework; (2) incorporate top incomes using Pareto type interpolations combining surveys and Personal Income Tax data; (3) using more merged employers and employees data sets such as Rais (this also allows to capture the role of firms) and so on. See [https://www.cps.fgv.br/cps/bd/curso/Drivers\\_IncomeDistribution\\_Neri\\_Brazil\\_Updated\\_GMD.pdf](https://www.cps.fgv.br/cps/bd/curso/Drivers_IncomeDistribution_Neri_Brazil_Updated_GMD.pdf)

**Correction Criteria:**

- (1) 7 Points;  
(2) 7 Points;  
(3) 6 Points.

2. What is your favorite poverty measure? Justify your answer comparing possible alternatives.

**Answer:** One simple approach is to argue in favor of poverty measures that are sensitive with respect to the inequality among the poor such as  $P^2$  which is obtained from the expression below with alpha equals 2:

$$P^\alpha = \frac{1}{n} \sum_{i=1}^q \left( \frac{Z - Y_i}{Z} \right)^\alpha$$

Other approach is to use (1) Sen measure  $P^S = P^0 \delta^P + P^1 (1 - \delta^P)$ . Where  $\delta^P$  is the Gini index among the poor. (2) The disadvantage here as opposed to  $P^2$  is that it is not exactly decomposable an impossibility inherited from the Gini. (3) The dual of the Theil -T would allow to circumvent this decomposition characteristic while keeping the polar cases of  $P^S$  measure:

If  $\delta^P = 0 \rightarrow P^S = P^1$

If  $\delta^P = 1 \rightarrow P^S = P^0$

An additional answer argues the importance of using the three traditional measures of the FGT poverty indicators as providing complementary answers to Key-Questions:  $P^0$  - How many are poor?;  $P^1$  - How serious is the problem? (the minimum poverty eradication cost); and  $P^2$  - Where should we start? (consistent targeting mechanism where, independent of the poverty line assumed, the poorest of the poor are first in line. While allowing the use of first, second and third order dominance concepts that allows to avoid poverty line always arbitrary choices).

Still other possibility is to defend the use of multidimensional poverty indexes (MPI) based on more dimensions in particular assets that are related to chronic poverty but subject to arbitrary choice of weights across dimensions and indicators and so on.

**Correction Criteria:**

All approaches have the same grade, that's 20 points:

- **First Approach (20 Points)**

- (1) 10 Points- Explain what's  $P^2$  and its sensitivity with respect to the inequality among the poor;
- (2) 10 Points- Compare  $P^2$  with other poverty measures.

- **Second Approach (20 Points)**

- (1) 10 Points;
- (2) 5 Points;
- (3) 5 Points.

- **Third Approach (20 Points)**

- (1) 8- Explain why using the 3 FGT poverty indicator;
- (2) 10 Points- Explain each FGT poverty indicator;
- (3) 2- Points- Talk about dominance and the arbitrary choices of poverty line.

- **Fourth Approach (20 Points)**

- (1) 12 Points- Explain Correctly the MPI;
- (4) 8 Points- Compare with other poverty measures.

3. How to measure the progress of societies? Highlight the main desired features.

**Answer:** One possibility is to work with income based poverty or MPI as last question suggest. Or could work with the Human Development Index (HDI) or the Inequality Adjusted Human Development Index (IHDI) as a framework.

Another more general answer is to incorporate desired results based properties such as prosperity, equity, sustainability and sensibility:

- Prosperity: growth in mean income (not only GDP but other National Accounts concepts. Also look to Household Surveys flows data);
- Sustainability: ability to maintain the standards of living achieved. Stocks of human, environmental, physical, cultural and social assets plus productivity;
- Equity: looking at the distribution among individuals and social groups of income flows, stocks of assets and rights;
- Sensibility: the last dimension is subjective, based on people's perception about the country, the public services and life quality.

Bonus: More formally you can measure the progress of societies with a General Social Welfare Function (discrete case):

**Single Period General Social Welfare Function:**

Integrated evaluation of Prosperity and Equality of a vector of income  $\tilde{x} \approx [x_1, x_2, \dots, x_N]$  into a single number:

- $$W = \sum_{i=1}^N u(x_i)v(x_i)\Pr(x_i)$$

Where  $\sum_{i=1}^N v(x_i)\Pr(x_i) = 1$ ,  $v(\cdot)$  is the weight and  $\Pr(x_i)$  is the probability function  $x_i$ .



**Social Welfare Function (Multi Period)**

Adds evaluation of Sustainability of the vector  $\tilde{x}_t \approx [x_{11}, \dots, \dots, x_{1T}, \dots, \dots, x_{N1}, \dots, x_{NT}]$ :

- Social Welfare Function (Multi Period)

$$W = \sum_{t=1}^T \sum_{i=1}^N u(x_{i,t}) v(x_{i,t}) Pr(x_{i,t})$$

**Add Subjective Sensibility**

- Can  $u(\cdot)$  above be measured directly? Happiness literature

**Correction Criteria:**

All approaches have the same grade, that's 20 points.

**First approach**

- (1) 20 Points-Explain the measure(s).

**Second approach (main):**

- (1) 4.5 Points for each desired results approach explained;
- (2) 2 Points- Bonus.

**Good Luck!**