First Exam (A1) of Social Economics and Public Policy 2023

Professor: Marcelo Neri
Teaching Assistant: Pedro Mencarini

Time: 3 hours. You may not use a calculator. Please handle in the questions sheet signed with your answers. Answers written in pencil are not subject to revision. The Exam has three main questions with different parts and choices. Please number your answers carefully and answer only the required number of questions in each question. See Formula sheet in the end of

Question 1 – Conceptual Exercises

Evaluate if each sentence is True or False (if any part is false the whole sentence is false). **YOU DO NOT NEED TO COMMENT WHY JUST Indicate** if each small letter is true or false. Answer only 25 (and only 25) of the 34 items from parts I to IV below **(0.28 points each)**:

I – Inequality and Social Welfare

a. The choice between national wide versus regional price indices in general affect income inequality measures. T
b. Income measurement error with zero mean does not affect social welfare measures. F
c. Lorenz dominance provide a valid comparison for all inequality measures. F
d. Lorenz curves can be seen as the general case of Concentration curves. F
e. Concentration ratios share the same upward bound as the Gini index. T
f. The advantage of the J-Divergence over the Theil T index is to allow decompositions between and within groups across variables such as education or gender. F
g. Inequality, mean and social welfare levels captured by household surveys such as PNAD are all underestimated by not capturing the income of the richest. The substitution of top incomes in PNAD by the ones extracted from Personal Income Tax data shows that in Brazil. T
h. The proportion of the richest 10% in income is an inequality measure that does not follow the principle of transfers (Pigou-Dalton). T
i. Income inequality measures captured by the share of the bottom 25% in income is consistent with Atkinson approach that derives inequality directly from a social welfare function. T
j. The growth of the proportion of active age population (PIA) in overall population provides a measure of the demographic bonus while the growth of years of schooling provides a measure of the educational bonus. T
k. The Theil-T Index dual can replace with some advantages the Gini index in the Poverty Indicator proposed by Amartaya Sen (1976). Being one improvement making poverty exactly decomposable. T
l. Increases in inequality within groups below and above the median, ceteris paribus, tend to increase both polarization and inequality (Gini). F
m. The alienation concept is related to inequality between groups. T

II - Poverty

a. Poverty targets based on P^1 (Poverty Gap) have difficulty in inducing actions aimed at the poorest in society. T
b. The minimum cost of overcoming poverty can be calculated from the poverty gap P^1. The rise of social benefits to its already beneficiaries is less costly than raising the poverty line by the same amount. T
c. The existence of pure economies of scale within households makes (linear) per capita household income to overestimate poverty. T
d. Social benefits fixed independently of household size can lead to fragmentation of families in the unified social register (CadÚnico). T
e. Datt-Ravallion decompositon allows to decompose poverty changes exactly into mean and inequality terms. F
f. The counterfactual of poverty measures with a distribution of year t and mean income of year t+1 is generated dividing the micro income data of year t+1 by the growth factor between the two years. F
g. If the proportion of poor in a given society A is always bigger than in society B so is the squared poverty gap. T
h. The poverty dominance analysis allows to reduce the arbitrariness derived from choosing a specific poverty line. T
i. If the Income Cumulative Distribution Function of society A is always above the one of society B, then we can ensure that all three FGT indicators (P^0, P^1 and P^i) are always higher in A than in B for any poverty line. T
j. If we adopt a social goal system based on the income poverty indicator known as the poverty head-count ratio (P^0) we will implicitly assume that priority is given to the least poor of the poor. T

III – Global Social Indicators and Social Targets

a. The standard Human Development Index (HDI), after incorporating the log of income as a component, is sensitive to inequality. T
b. The Inequality adjusted Human Development Index (IHDI) is irresponsive to mean changes in HDI components. F
c. The Multidimensional Poverty Index (MPI) is sensitive to the intensity of poverty across different dimensions (for example, child school enrollment) but not to the number of children not enrolled in school in a given household. T
d. A system of targets based on international indicators such as MDGs and SDGs allows to lengthen the planning horizons of policy makers; intermediate actions between different levels of government and creates automatically insurance against systemic (aggregate) shocks through published rankings. T
e. If politicians are less concerned with underrepresented groups in the electoral market, such as children, social targets can eliminate completely the distance of the treatment given to different groups. F
f. Ceará State uses municipality education performance to distribute the proceeds of its State Taxes (ICMS). This can not only improve local education performance, bring external resources but also increase the incentive to raise each municipality efforts to raise the education budget. T
g. Idiosyncratic shocks can be dealt with in a social targets framework through published rankings. F

IV - Empirical Methods Analysis
   a. In the context of a Mincerian (log-linear) earnings equation, the concept of the net contribution to inequality (isolating its effects from other variables) is measured by the R^2 of a regression with only one constant term and the variable that we want to evaluate the net contribution (education, for example). F
   b. The advantage of running bivariate regressions with a constant over using simple cross tabulations is to provide confidence intervals to test hypothesis. T
   c. In the context of a Stepwise Procedure, it is not possible to capture externality effects of infrastructure because the procedure is automatic and carried out without any theory. F
   d. Differences in differences estimator (DD or diff-in-diff) is captured by an interactive term in a regression. This work for continuous but not for discrete regression models. F

Question 2 – Conceptual and Discursive – Choose 1 (and only 1) of the 3 items below: **(1.5 points)**

I) i. What is your favorite income inequality index? ii. Justify your choice in practical, ethical and theoretical grounds using its formula and respective social welfare function specification. Illustrate its main features. iii. Describe decomposition methods applied to it. iv. (Bonus 0,5 points: What are the main challenges to measure income distribution in Brazil?).

Criteria:
ii) (0,5 Points): Justify your choice in practical, ethical and theoretical grounds using
             course contents;
             (0,25 Points)
             Show formula and respective social welfare function specification
             (0,25 Points).
iii) (0,5 Points): Describe decomposition methods applied to the inequality index.
iv) Bonus (0,5 points): What are the main challenges to measure income distribution
                       in Brazil?
Answer: Go beyond gross income (that bases surveys as PNAD or PNADC) i. incorporate the roles of direct and indirect taxes as does a microsimulation framework; ii. Incorporate top incomes using Pareto type interpolations combining surveys and surveys; iii) using more merged employers and employee's data sets such as Rais (this also allows to capture the role of firms) and so on.

II) i. What are the possible constraints imposed on the social welfare function below so that the Pigou-Dalton's Principle of Transfer is observed? Provide examples. ii. Write down the functional form of the Social Welfare function associated with the Gini Index from the equation below. Explain each component. iii. Sketch and explain the passage from the Social Welfare Function to the Gini inequality measure. iv. (Bonus 0,5 points: How to incorporate the temporal choice dimension in the equation below. Explain.)

$$u(x^*) = \int_0^\infty w(x)u(x)f(x)dx$$

Answer:

i.   (0,35 Points): We could observe the Pigou-Dalton Principle of Transfers if we assume utility functions u(x) with decreasing marginal utilities or if we impose bigger weights w(x) to the poorest. One special case is if we take u(x) = log(x) and w(x) = 2[1 − F(x)], where F(x) is the cumulative distribution function.

ii. (0,3 Points): The Social Welfare function associated with Gini Index corresponds to the particular case where the utility function is u(x) = x and the weights are: w(x) = 2 [1 − F(x)], where F(x) is the cumulative distribution function of the income.

iii. (0,35 Points): If u(x) = x and w(x) = 2 [1 − F(x)] then applying Atkinson certainty equivalent idea we have:

$$u(x^*) = \int_0^\infty w(x)u(x)f(x)dx = 2 \int_0^\infty x\,[1 - F(x)]f(x)dx = \mu(1 - G)$$

Where, μ is the mean income of the society and G is the Gini Index.

iv. Bonus (0,5 Points): A simple way to incorporate the temporal choice dimension in the equation above is from the formula bellow:

$$u(x^*) = \int_0^T \int_0^\infty w(x_t)u(x_t)f(x_t)dx_t\, dt$$

Besides aggregating individual welfare levels into social welfare in a moment of time, the first integral aggregates different instants of time (and also under uncertainty different states of nature).

III) i. What is your favorite poverty framework? Analyze its properties (uni x multidimensional; relative x absolute; subjective x objective indigence x poverty; etc) ii. Justify your choice. iii. (Bonus 0,5 points: Provide the main, historical trends and measurement issues associated comparing in broad terms the evolution of unidimensional (income-based poverty and multidimensional in Brazil).

Criteria:
i.   (0,5 Points): Properties (uni x multidimensional; relative x absolute; subjective x objective indigence x poverty; etc);
ii. (0,5 Points): Justify your choice using course contents.
iii. Bonus (0,5 Points): Provide the main, historical trends and measurement issues associated comparing in broad terms the evolution of unidimensional (income-based poverty and multidimensional in Brazil).

Question 3 – Quantitative Questions – Choose 1 (and only 1) of the 3 items below: **(1.5 points)**

I) i. Compare advantages and disadvantages of the incomebased poverty indicators known as P^0, P^1 and the Mean Squared Poverty Gap (P^2). Give the general and specific formulas and intuition. Sketch their relationship with poverty dominance concepts. ii) Calculate the Proportion of the Poor (P0), the Mean Squared Poverty Gap (P2), the Mean Poverty Gap (P1) and the mean cost of eradicating poverty per person for the 2 following periods and assuming a poverty line of 5 units: Period 1 = {2, 4, 6,8}; Period 2 = {3, 4, 7,10}. iii. Do we have first order dominance of the distribution in period 2 in relation to period 1? Consider the relevant range of poverty lines going up to 7. iv. As in the Datt-Ravallion decomposition, generate the contra factual distribution values of the mean in period 1 and distribution in period 2.

Answer:
i.   (0,4 Points):

(0,2 Points):
P^0: Advantage: it is a simple and easy-to-understand measure of poverty. Disadvantage: Don't take into account the inequality between the poor and don't measure how far the income is from the poverty line.
P^1: Advantage: Distinguishes different levels of poverty; measures how far, on average, it is from the poverty line. Disadvantage: does not take the inequality between the poor into account.
P^2: Advantage: Take into account inequality between the poor;

(0,2 Points):
The general formula of the 3 measure of poverty is given by:
$$P^n = \frac{1}{N}\sum_{i=1}^{Q}\left(\frac{z - x_i}{z}\right)^{\alpha}$$

Where x_i is the income of the individual i; z is the poverty line, Q is the population below the poverty line; N is the population; and α is the poverty aversion degree.

When $\alpha = 0$ we have the proportion of the poor (P^0) which is given by the formula below:

$$P^0 = \frac{1}{N}\sum_{i=1}^{Q}\left(\frac{z - x_i}{z}\right)^{0} = \frac{Q}{N}$$

 It answers the question: How many are poor in the society?

When $\alpha = 1$ we have the Mean Poverty Gap (P^1) which is given by the formula below:

$$P^1 = \frac{1}{N}\sum_{i=1}^{Q} \left(\frac{z - x_i}{z}\right)$$

It answers the question: How serious is the problem? This poverty measure measures how far, on average, the income is from the poverty line.

When $\alpha = 2$ we have the Mean Squared Poverty Gap (P^2) which is given by the formula below:

$$P^2 = \frac{1}{N}\sum_{i=1}^{Q} \left(\frac{z - x_i}{z}\right)^2$$

It answers the question: Where should we start? This poverty measure proportionately gives more weight to the poorest.

The relationship between the poverty measures and poverty dominance concepts is as follows: if a distribution A dominates another distribution B in 1st order (FOD) then $P_A^0 > P_B^0$ for every z; if a distribution A dominates another distribution B in 2nd order (SOD) then $P_A^1 > P_B^1$ for every z; if a distribution A dominates B in 3rd order (TOD) then $P_A^2 > P_B^2$ for every z.

ii. (0,4 Points):

(0,1 Points): P^0: Notice that half of the population income in both periods is below the poverty line of 5 units. Then, since P^0 is the proportion of poor, so P^0 50% in t1 and 50% in t2.

(0,1 Points): P^1: ¼ (5-2)/5 + ¼ (5-4)/5 = ¼ (3/5) + ¼ (1/5) = 3/20+1/20=4/20=20% in t1;

¼ (5-3)/5 + ¼ (5-4)/5 = ¼ (2/5) + ¼ (1/5) = 2/20+1/20=3/20=15% in t2.

(0,1 Points): P^2: ¼ ((5-2)/5)² + ¼ ((5-4)/5)² = ¼ (3/5)² + ¼ (1/5)² = ¼ 9/25+ ¼ 1/25= 9/100+1/100 =10% in t1;

¼ ((5-3)/5)² + ¼ ((5-4)/5)² = ¼ (2/5)² + ¼ (1/5)² = ¼ 4/25+ ¼ 1/25= 4/100+1/100 =5% in t2.

(0,1 Points): Mean cost of eradicating poverty per person: (3+1)/4=1 in t1 and (2+1)/4 = 0.75 in t2.

iii. (0,35 Points): A distribution A is said to dominates in first order (FOD) another distribution B if $P_A^0 > P_B^0$ for every poverty line z. So, for a FOD of period 1 in relation to period 2 we must have: $P_1^0 > P_2^0$ for every poverty line z. However, notice that in both periods for z=5 P^0 is the same then there's no FOD of the income distribution of period 1 in relation to the income distribution of period 2.

iv. (0,35 Points): The Datt-Ravallion decomposition, generate the contra factual distribution values of the mean in period 1 and distribution in period 2.

$$\mu_0 = 5; \; \mu_t = 6; \; \frac{\mu_t}{\mu_0} = \frac{6}{5}; \; \frac{\mu_0}{\mu_t} = \frac{5}{6}$$

Then the contra factual distribution values of the mean in period 1 and distribution in period 2 as in the Datt-Ravallion decomposition is given by:

d1: {2;4;6;8} $\Longrightarrow$ {12/5; 24/5; 36/5; 48/5}

d2: {3;4;7;10} $\Longrightarrow$ {15/6; 20/6; 35/6; 50/6}

II) Empirical Analysis of the Logistic regression: i. Discuss the level and the evolution of poverty in Brazil and in Rio de Janeiro State from the binomial logistic regression below. ii. How to interpret the two terms for education in the regression below? iii. What if the regression also displayed a negative coefficient for the mean education in the State. How would you interpret that? iv. What is the importance of restricting the sample to those with age 25 years of age or above?. v. What is the advantage of multivariate poverty analysis? vi. What is the difference between multivariate analysis of poverty and multidimensional poverty index?

Binomial Logistic Regression Poverty Line FGV CPS – For those with age 25 years of age or above.
INTERACTION STATE*YEAR OBS: Other State Categories are not displayed below
Also controls for gender age, city size and variables related to access to infrastructure are also not displayed.

Parameter	Category	Estimate	Standard Error	Chi-Squared	sig	Conditional Odds Ratio
YEARS OF EDUCATION		-0.0232	0.0001	25542.3	**	0.97703
(YEARS OF EDUCATION)^2		-0.0102	0.0000	728969	**	0.98983
STATE	RJ	0.0332	0.0010	1036.69	**	1.03371
STATE	zSP	0.0000	0.0000	.		1.00000
YEAR	a2015	-0.7293	0.0009	603648	**	0.48223
YEAR	z2004	0.0000	0.0000	.		1.00000
STATE*YEAR	RJ	-0.0661	0.0018	1411.80	**	0.93605
STATE*YEAR	RJ	0.0000	0.0000	.		1.00000
STATE*YEAR	zSP	0.0000	0.0000	.		1.00000

i.   (0,25 Points): Poverty incidence in Brazil decreased significantly between 2004 and 2015 (-.7293 coefficient with corresponding odds ratio below unity 0,482). People in Rio de Janeiro, had 3,37% higher chances of being poor than in São

Paulo considering the whole period. Rio de Janeiro citizens chances of being poor decreased 6,4% between 2004 and 2015.

ii. (0,25 Points): Education has a negative impact on poverty, the quadratic term show that this effect increases in module meaning that higher levels of education produce on the margin larger the effects on poverty.

iii. (0,25 Points): It means that a State with more average education has less poverty incidence besides the individual effect of education on poverty. That's probably due to positive externalities. For example, imply in better teachers or health professionals.

iv. (0,25 Points): Since the education cycle approximately ends before 25 years of age, this restriction avoids incorporating children and teenagers that would distort the analysis.

v. (0,25 Points): It controls for other observable factors, allowing to isolate the contribution of a particular factor, providing a more precise net estimate of the contribution of a particular factor on the variable of interest, keeping the other variables constant. For example, in the example we take into account educational, socio demographic and infrastructure differences when estimating the regional effects which may indicate possible policies paths.

vi. (0,25 Points): Multivariate analysis controls for other observable factors, allowing to isolate the contribution of a particular factor, providing a more precise estimate of the contribution of a particular factor on the variable of interest. In other words, it captures the relations between two variables keeping the remaining variables constant. In the other hand, the Multidimensional Poverty Index (MPI) identifies multiple deprivations at the household level in education, health and standard of living. It uses micro data from the same household surveys to construct a measure of poverty. However, it doesn't estimate the isolated contribution of a particular factor it takes a predefined weight in which indicator and condenses it into one poverty indicator. Other difference is that MPI, focus on education, health and standard of living, on the other hand, in multivariate analysis you can include other relevant variables and find the isolated effect of it on the variable of interest. In sum, multivariate analysis is related to the right hand side of the equation (or explanatory variables) while multidimensional poverty is related to the left hand side variables (or endogenous variables).

III) Imagine a set of four countries to be compared. In Table below you will see some of their respective social indicators, for each there is a minimum and maximum values to be considered for calculating the Human Development Index (HDI). In Table 2, you will see their loss due to inequality.

Table 1 - Indicator	Minimum	Maximum	Country's arithmetic mean
Life expectancy (years)	20	85	Brazil: 75.7
			Cuba: 79.9
			USA: 79.5
			China: 76.4
Expected years of schooling (years)	0	18	Brazil: 15.4
			Cuba: 14.0
			USA: 16.5
			China: 13.8
Mean years of schooling (years)	0	15	Brazil: 7.8
			Cuba: 11.8
			USA: 13.4
			China: 7.8
Gross national income per capita (2011 PPP $)	100	75000	Brazil: 13,755
			Cuba: 7,524
			USA: 54,941
			China: 15,270

Table 2 - Loss of HDI due to inequality

Brazil	23.9
China	14.5
Cuba	0
United States	13.8

Table 3 Below we already calculate the HDI and the Inequality Adjusted HDI for each country and rank them.

	HDI	Rank	InAdj HDI	Rank
Brazil	0.751924	3	0.572214	4
Cuba	0.777754	2	0.777754	2
USA	0.924204	1	0.796664	1
China	0.746589	4	0.638333	3

i. Summarize your take from the findings about Brazil posed on Table 3 above. ii.. Explain the methodology to construct the Inequality Adjusted Human Development Index (IHDI). iii. Provide the main advantages of IHDI compared to standard income based social welfare measures.

i. (0,5 Points): Brazil loses position and is behind China when adjusting HDI by inequality, since inequality in Brazil is higher than in China.

ii. (0,5 Points): Let x_j be the index of individual j (it can be heath index (h_j); education index (e_j) or income index (i_j); n be the population of the society. First, an inequality measure of the index x (A_x) is calculated by the formula bellow:

$$A_x = 1 - \frac{\sqrt[n]{x_1 + \cdots + x_n}}{\overline{x}}$$

Where $\overline{x}$ is the arithmetic mean of the index x.

Then Inequality-adjusted dimensions indices (I_{x^*}) are calculated by multiplying the HDI dimensions indices (I_x) by their corresponding inequality measure ($1 - A_x$):

$$I_{x^*} = (1 - A_x)I_x$$

Finally, the IDHI is obtained by the geometric mean of the three inequality-adjusted dimensions indices:

$$IDHI = \sqrt[3]{I_{h^*} \cdot I_{e^*} \cdot I_{i^*}} = HDI \cdot \left[\sqrt[3]{(1 - A_h)(1 - A_e)(1 - A_i)} \right]$$

iii. (0,5 Points): The welfare indexes based on a single monetary dimension are the most used with the, often unrealistic, assumption that individuals in a market economy are able to directly acquire goods and services that meet your needs. It is often possible to impute values of rent for those who own their homes, of commuting time lost evaluated at each hourly wage rate, of public education and health expenditures and to incorporate them directly into income-based welfare measure. However, it is interesting to have a distinction of welfare measures with many dimensions of those with a single dimension. IHDI, is one of those welfare measures with many dimensions that, besides of accounting for other aspects of social welfare beyond the income like the lack of access to other basic elements such as education and health, like HDI does, it also takes into account the inequality within each index that is an important component for social welfare measurement. It does that by discounting each dimension's average value according to its level of inequality.

Good Luck!