Spring 2017

ECON 133 "Global Inequality and Growth"
Final Exam

Exercise 1 (20 points): True False Statement/Questions

Explain your answer fully based on the material seen in lecture and section (no more than 5 lines per question). All the credit is based on the explanation. (2 points for each question.)

1. The Gini coefficient is very sensitive to inequality at the top-end of the distribution.
   - FALSE. The Gini coefficient puts a greater weight on the middle of the distribution, and not the extreme ends.
2. Modigliani's definition of the share of inherited wealth in total wealth is problematic because it fails to recognize that inherited wealth produces flow returns. However, Kotlikoff and Summers' definition is also problematic because using their definition, the share of inherited wealth in the economy can be larger than 100%.
   - TRUE. $W_B^M = \sum_{t-30 \leq s < t} B_s$, i.e., no capitalization of inherited wealth. This leads to artificially low values of inherited wealth ($\phi^M \approx 20\%$). However, ϕ^{KS} gives values close to or greater than 100%, and is extremely sensitive to the exact value of $r - g$.
3. An increase in $r - g$ typically raises the bequest flow because it leads to an increase in the ratio between the average wealth at death and the average wealth of the living population.
   - TRUE. μ_t tends to be high when $r > g$ because it makes it easier for capital owners to accumulate wealth.
4. If inequality came entirely from labor income, then capital taxation should be zero.
   - FALSE. Even if inheritance did not play a role, there are still practical reasons to tax capital. In practice, for instance, it is not easy to decompose income flows into pure L and K components (think Mark Zuckerberg). To avoid income shifting due to differential tax treatment of L and K, capital tax should be positive. In fact, in the extreme case where the government cannot distinguish at all between L and K, then the marginal tax rates on L and K should be equal.
5. Consider a corporation operating in Germany but fully owned by US shareholders. With a source-based corporate tax system it pays taxes in Germany, and with a residence-based corporate tax system it pays taxes in the United States.
   - TRUE. Full mark for explaining source- versus residence-based taxation.
6. The artificial shifting of profits to low-tax countries has contributed to a decline in the effective tax rate of top 1% income earners since the 1990s.
   - TRUE. The bulk of the corporate tax is borne by top 1% income earners (because it mostly falls on capital and capital is very concentrated in the US). The effective corporate tax rates on US corporate profits have dropped since 1980. Artificial profit shifting is the main factor behind this trend (responsible for about 2/3 of the effect).
7. The optimal bequest tax rate is high if there is a high degree of social mobility.
   - FALSE. It's the contrary: the optimal bequest tax rate is high if there is low social mobility (i.e., low $\bar{b}$, that is, zero bequest receivers are unlikely to leave large bequests to their kids).

8. Under sales-based formula apportionment, multinational firms would have fewer incentives to engage in profit shifting.

   - TRUE. Sales-based formula apportionment would put make your tax burden owed to a country be proportional to the amount of sales of your good or service in said country. Under this tax scheme, multinational firms would have fewer incentives to shift profits to lower-tax countries, as the location of *sales* will affect the ultimate tax burden rather than the multinational's source country.

9. The skill premium has been rising since the 1960s in the United States due to the rising demand for skilled labor (i.e., people with college degrees).

   - FALSE, for two reasons. First, the skill premium has not always risen - it stabilized in the 2000s. Second, the above statement improperly describes the Tinbergen model. The skill premium rises whenever the *rate of increase* in demand exceeds that of supply. This is *not* the same as saying that the skill premium is rising whenever demand is rising (or whenever demand exceeds supply).

10. Given what we know about the evolution of the capital share α and the wealth to income ratio β in the U.S., Cobb-Douglas production is a good approximation for the aggregate production function in the United States.

    - FALSE. The Cobb-Douglas function implies an elasticity of substitution of 1, which in turn, implies that the capital share remains constant over time (in this case, an increase in β would lead to an decrease in r of the same magnitude, keeping α constant). However, we have observed that both α and β have increased over time - this positive relationship between the two indicates that $\sigma > 1$. Thus, a more generalized CES function would be more appropriate production function.

Exercise 2 (10 points): Wealth Accumulation

Consider the wealth accumulation equation $W_{t+1} = W_t + s_t Y_t$.

1. What assumption does this equation make about capital gains? Is this assumption realistic? (0.5 points)

   - It assumes there are no price effects, i.e. only volume effects matter (0.25 points)
   - However in the short run, price effects matter. (0.25 points)

2. Express β_{t+1} in terms of β_t, s_t and g_t. (2 points)

   - Start out with the wealth accumulation equation: $W_{t+1} = W_t + s_t Y_t$. Then, divide both sides by $Y_{t+1} = Y_t(1 + g_t)$ to get $\beta_{t+1} = \frac{\beta_t + s_t}{1+g_t}$.

3. Now suppose $W_{t+1} = (W_t + s_t Y_t) \cdot (1 + q_t)$. What does the q_t term capture? Express β_{t+1} in terms of β_t, s_t, g_t, and q_t. (2 points)

   - Price effects matter with this wealth accumulation (1 point)
   - $\beta_{t+1} = \frac{\beta_t + s_t}{1+g_t} \cdot (1 + q_t)$ (1 point)

4. Show that if $q_t = 0$ and in steady-state, $\beta = s/g$. What is this formula called? Briefly interpret this formula (2 points).

   - Since in steady-state $\beta_t = \beta_{t+1}$, $s_t = s$ and $g_t = g$ (i.e. constant), we plug those in to get $\beta = s/g$. (1 point)
   - This is the Harrod-Domar-Solow formula (0.25 points)

- This formula states that, in steady state, the K/Y ratio is equal to net-of-depreciation savings rate over the growth rate (0.75 points)

5. If the savings rate s is 10% and the growth rate g is 2%, what is β equal to in the long-run? Assuming the capital share α is 30%, what is r, the long-run rate of return, equal to? (1 point)

   - $\beta = 10\%/2\% = 500\%$ (0.5 points)
   - If $\beta = 500\%$, then $r = \alpha/\beta = 30\%/500\% = 6\%$ (0.5 points)

6. If $\beta = s/g$ in the long run, do you expect the global wealth-to-income ratio to rise by the end of the 21st century? Why? (1 points)

   - Yes, β will rise. (0.5 points)
   - This is because long-run growth (g) = population growth (n) + productivity growth (h). Population growth will probably fall so that β might rise. Productivity growth may also fall or at least not compensate for the decline in population growth. (0.5 points)

7. Is it always the case that the capital share of income has to rise when the wealth-to-income ratio β rises? Explain. (1.5 points)

   - No. (0.5 points)
   - It depends on the value of the elasticity of substitution between capital and labor, σ. If $\sigma > 1$ then the capital share α and the wealth-income ratio β move in the same direction, and vice versa. (1.0 point)
   - Empirically, the capital share and the wealth-income ratio has moved in the same direction over the 20th century, suggesting $\sigma > 1$ (0.5 points)

Exercise 3 (10 points): Modigliani's life cycle model

According to Modigliani's life cycle model, individuals save during their worklife to consume during retirement, so as to fully smooth consumption across their life cycle.

1. Assume that everybody starts working at age 0, works for N years, dies at age L, and leaves no inheritance. Assume further that labor income Y is constant at $\bar{Y}$ during the working age period, and is 0 afterwards. Express individual annual consumption C as a function of $\bar{Y}$, N, and L. Interpret this equation. (2 points)

   - $C = (N/L)\bar{Y}$ (1 point)
   - There is full consumption smoothing: consumption is constant throughout an individual's lifetime. (1 point)

2. Express individual annual saving during worklife AND annual dissaving during retirement as a function of $\bar{Y}$, N, and L. What is the saving rate during worklife? Interpret. (1 points)

   - Annual saving during workage $S = \bar{Y} - C = \frac{(L-N)}{L} \cdot \bar{Y}$ (0.25 points)
   - Annual dissaving during retirement $S = -\frac{N}{L} \cdot \bar{Y}$ (0.25 points)
   - Saving rate during worklife is $\frac{(L-N)}{L}$ (0.25 points)
   - Saving rate increases with retirement duration (i.e., the longer an individual will be consuming without receiving income) and decreases with life expectancy (0.25 points)

3. Use a graph to represent income $\bar{Y}$, consumption C, saving S, and wealth A as a function of age. What is the amount of saving an individual will have accumulated at the time when he or she retires? (2 points)

   - Graph (1 point):

- 0.25 point for correct axes $T = 0$ to $T = L$ and marking $T = N$
- 0.25 points for straight $\bar{Y}$ line starting at $T = 0$ and ending at $T = N$
- 0.25 points for straight horizontal consumption line from $T = 0$ to $T = L$ below $\bar{Y}$ line
- 0.25 points for A triangle, with apex at N

- $A = S \cdot N = \frac{N(L-N)}{L}\bar{Y}$ (1 point)

4. Provide expressions for average wealth A/L and aggregate wealth A. Interpret. (1 point)

- $A/L = \frac{(L-N)N}{2L} \cdot \bar{Y} = \frac{(L-N)}{2L} \cdot Y$ (0.25 points)
- Average wealth increases with income and retirement duration, and decreases with life expectancy. (0.25 points)
- $A = \frac{(L-N)}{2} \cdot Y$ (0.25 points)
- Total wealth increases with income and retirement duration (0.25 points)

5. Give the formula for the wealth-to-income ratio A/Y according to this model. Interpret this formula. Calculate A/Y when $L = 80$ and $N = 70$. (1 point)

- $A/Y = \frac{L-N}{2}$ (0.25 points)
- The wealth-to-income ratio is not a function of income but purely of retirement duration, i.e., demographics. (0.5 points)
- $A/Y = \frac{80-70}{2} = 500\%$ (0.25 points)

6. The goverment has suddenly lowered retirement age from 70 to 60. All else equal, what will this reform do to the wealth-to-income ratio? Provide some intuition for your answer. (1 point)

- The wealth-to-income ratio increases to $A/Y = \frac{80-60}{2} = 1000\%$. (0.5 points)
- Now that individuals have longer retirement duration, they must save more during worklife. This means that wealth-to-income ratios must increase. (0.5 points)

7. Suppose now that $n = g = 0$ but $r > 0$. How would this affect C and S? Provide intuition. (1 point)

- If $n = g = 0$ and $r > 0$, C would increase and S would decrease. (0.5 points)
- Intuition: other things equal, the young need to save less for their old days (thanks to the capital income $Y_K = rW$) (0.5 points)

8. Does Modigliani's life cycle model accord well with the data? (1 point)

- No this model does not accord well with the data. (0.5 points)
- The model implies $\beta \approx 500\%$, which is close to the observed values. However, in the model wealth is as unequally distributed as labor income – but in reality it is more unequally distributed. The model focuses only on labor income, but capital income also matters. (0.5 points)

Exercise 4 (10 points): Optimal Taxation

Recall the optimal linear labor taxation problem: a benevolent social planner chooses the linear labor income tax rate τ and the amount of lump-sum per capita transfer G to maximize a social welfare function W subject to individual and government budget constraints.

1. Define the Rawlsian social welfare function. Explain intuitively why the optimal linear labor taxation problem is equivalent to maximizing G when W is the Rawlsian social welfare function. (1 point)

- The Rawlsian social welfare function $W = \min\{U_1, ..., U_n\}$, or the utility of the worst-off individual. (0.5 points)
- This individual solely relies on the government transfer G for revenue, so maximizing her utility is equivalent to maximizing the transfer she receives. Thus, in order to make the transfer as large as possible, we will want to maximize government revenue. (0.5 points)

2. Express the optimal tax rate τ^* in terms of the taxable income elasticity when the social welfare function is Rawlsian. Interpret the formula. (2 points)

   - $\tau^* = \frac{1}{1+e}$, where e is the taxable income elasticity. (1 point)
   - As e increases, τ^* decreases. Therefore, our optimal tax will be lower when we have an elastic labor supply. (1 point)

3. With a general social welfare function, can the optimal tax rate be greater than τ^*? Explain why or why not. (2 points)

   - No. (0.5 points)
   - All $\tau > \tau^*$ will be Pareto-inefficient meaning *every agent* will be worse off under higher taxes, regardless of the social welfare function. (0.5 points)
   - The government transfer will necessarily be lower, because for $\tau > \tau^*$, the behavioral effect will dominate, and the government will receive lower tax revenues. (0.5 points)
   - Additionally, every individual will have less post-tax labor income (lower G and lower z due to the behavioral effect), and thus, have lower utility. (0.5 points)

4. Suppose there were no behavioral responses to taxation. What is τ^* equal to under a Rawlsian social welfare function? (1 point)

   - $\tau = 1$ (0.5 points)
   - Explanation: if there were no behavioral responses, $e = 0$. In this case, $\tau = \frac{1}{1+0} = 1$. A logical explanation without math would also suffice. (0.5 points)

5. Consider now a piecewise linear labor income tax system, so there are now different marginal tax rates for different income groups. Assume that labor income is Pareto distributed for income earners in the top tax bracket. Assume the government wants to maximize the tax revenue it extracts from top earners. Will the optimal top marginal tax rate *always* be larger than the optimal linear income tax rate from question 2? (2 points)

   - No.
   - The optimal top marginal tax rate $\hat{\tau} = \frac{1}{1+a\cdot\hat{e}}$, where $\hat{e}$ is the elasticity for top-earners.
   - This means that $\hat{\tau} < \tau$ if and only if $a \cdot \hat{e} > 1$.
   - Therefore, $\hat{\tau} < \tau$ when $\hat{e}$ and/or a are large. This also means that the optimal top rate will be lower when *there is not much top end inequality and/or when the elasticity of top earners is relatively high.*

   (2 points, various points for partial credit, needed to discuss both a and $\hat{e}$ for full credit)

6. How have top marginal tax rates changed over the past century in the United States? Explain how top marginal income tax rates can affect the concentration of pre-tax income at the top. (2 points)

   - The top marginal tax rates were much higher in the years following World War II (the top rate was around 90%). The top marginal tax rate fell a bit in the mid-1960s and substantially in the 1980s. It has risen slightly since then - there have been some small changes in the past 20 years, but nothing as drastic. The current top rate is 39.6%. (1 point)

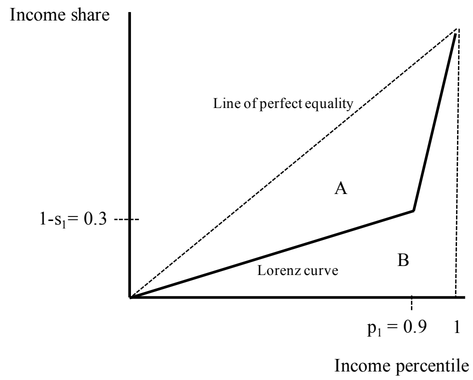
- When the top marginal tax rate is high, CEOs have less of an incentive to bargain for higher compensation (the reward is much lower). Thus, lower top tax rates can lead to greater pre-tax income concentration. (1 point)

Exercise 5 (10 points): Measuring Income and Wealth Inequality

Part A: Gini coefficients

Suppose in the nation JJ's World, there are 2 homogeneous income groups—a "low-wage" group and a "high-wage" group. That is, there are two possible wages in the economy, a low wage w_L and a high wage w_H, and everyone either earns the low wage w_L or the high wage w_H. 90% of the population belongs to the low-wage groups, and this group earns 30% of total income.

1. Draw a Lorenz curve depicting this distribution of income. Using this graph, show that the Gini coefficient in JJ's world is 0.6 (2 points)



(1 point)

- $G = s_1 + p_1 - 1 = 0.7 + 0.9 - 1 = 0.6$ (1 point)

2. How does the Gini coefficient in this economy compare to the empirical Gini coefficient for the world as a whole? (1 point)
   - As seen in the Milanovic reading, the world Gini coefficient is about 0.7, so JJ's world has a *smaller* Gini coefficient than the world. (1 point)

3. What is the top 10% income share in this economy? How does it compare to the top 10% pre-tax income share in the United States? (1 point)
   - The top share here is $100\% - 30\% = 70\%$. (0.5 points)
   - As seen in the Piketty, Saez, Zucman reading, the top 10% pre-tax income share is about 45% (alternatively, the top 1% pre-tax income share is about 20%). Thus, the top 10% share in JJ's world is *larger* than that of the United States. (0.5 points)

Part B: Wealth inequality

4. How can one use capital income, as reported on individual tax returns, to measure the distribution of wealth? What must we assume about rates of return on assets? (2 points)
   - We capitalize it! Divide these capital incomes from each asset type by the rate of return for each respective asset to measure the individual's wealth stock. (1 point)
   - The assumption is that within asset classes, the rates of return on assets do not vary across the wealth distribution. (1 point)

5. How does the concentration of wealth compare to the concentration of income in the United States? (1 point)

- Wealth is more concentrated than income: top 10% pre-tax income share around 50%, top 10% wealth share around 75%. (1 point)

6. One result found when estimating wealth inequality is that the rich tend to have higher saving rates. Is this finding consistent with the precautionary savings model seen in class? (1 point)

   - The precautionary savings model, which motivates savings due to insurance against "rainy days", implies that people with less money will have higher savings rates, since they are more likely to be negatively impacted by an adverse labor income shock. However, empirically, we find that the rich tend to save at a higher rate, which is completely opposite the implications of this model. (1 point)

7. What type of model can successfully explain the high degree of wealth concentration? Explain. (2 points)

   - Dynamic random shocks model (0.5 points)
   - In this model, an individual gets utility from consumption and saving, but her preference for saving is random, or has a "random shock" component. (1 point)
   - In this model, under steady state, top incomes follow a Pareto distribution, with larger variance in taste shocks leading to a lower a, or greater top income inequality. (0.5 points)

Bonus

1. On April 13, we Tweeted an article written by the Washington Center for Equitable Growth's Nick Bunker on determining the optimal U.S. tax rate for higher earners. Bunker reviews Piketty, Saez and Stantcheva's paper on optimal labor income taxation. What are the three elasticities mentioned and why are they important in determining optimal tax rates? (2 points)

   - The elasticities are as follows, along with their estimates:
     (a) labor supply elasticity
     (b) elasticity for income shifting between capital and labor, which is essentially 0
     (c) bargaining elasticity for top earners
   - The authors estimate these elasticities to be 0.2 at most, 0, and 0.3 at the *least*, respectively. Considering only the former elasticity, the typical elasticity we consider when studying optimal taxation, the authors find that the top rate should be 57%. However, when considering all three elasticities, the authors conclude that the top rate should be much higher: 83%. This is because these three different elasticities represent three different behavioral responses to taxation. The third response is the most salient, implying much higher optimal taxes. (1 point)

2. On April 7, we Tweeted an article published in *The Economist* on the relationship between catastrophe and inequality. What is the main argument? (2 points)

   - Catastrophes reduce inequality by destroying large amounts of capital. (1 point)
   - Catastrophes include wars, revolution, collapse of state, and pandemics. (0.5 point)
   - Moreover, catastrophes are the *only* events that have led to significant decreases in inequality. (0.5 points)