Problems for 5700 Fall 2019

**1.** Under 5750 "Problem sets with solutions" do Problem 6 (a), (c); Problem (7)

**2.** Solve the Malthus model discussed in class when the utility function is given by

$$U_t = (1-\beta)\ln(c_t - \widehat{c}) + \beta \ln(n_t), \tag{1}$$

where $\widehat{c}$ is referred to as subsistence consumption. Show in a 45°-diagram that the dynamics for population ($L_t$) can display oscillatory cycles.

**3.** Do Problem 1 under "Practice problems for midterm 2" from the 5110 problems: http://www.nippelagerlof.com/teaching/5110/Problems5110.pdf

**4.** Recall the (discrete-time) Solow model dicussed in class, where we derived

$$k_{t+1} = \frac{sy_t + (1-\delta)k_t}{1+n}. \tag{2}$$

The notation is standard (same as in class), except that we here let $y_t$ denote (non-logged) per-capita GDP [i.e., the same as $f(k_t)$ in class].
(a) Let $y_t = Ak_t^{\alpha}$, where $A > 0$ and $0 < \alpha < 1$. Rewrite (2) as a difference equation expressing $y_{t+1}$ in terms of $y_t$ and exogenous parameters. Your answer should not involve any $k_t$.
(b) Use your answer under (a) to find the steady-state level of $y_t$, denoted $\overline{y}$.
(c) Use your answer under (a) to write the growth rate of $y_t$, here defined as $g_t = \ln(y_{t+1}) - \ln(y_t)$, in terms of $y_t$ and exogenous parameters.
(d) Set $\delta = 1$. Use your answer under (c) to show that we can write $g_t$ as a linear function of the log gap in per-capita GDP, $\ln(y_t) - \ln(\overline{y})$ (and of $\alpha$). How does your answer relate to the Barro-style regression equation discussed in class?
(e) Go back to the case general case, $\delta \leq 1$. Use your answer under (c) to find a linear Taylor approximation of $g_t$ around the steady state level of log per-capita GDP, $\ln(y_t) = \ln(\overline{y})$. Hint: if your answer under (c) is is written $g_t = \Omega(\ln(y_t))$, then the Taylor approximation is given by $g_t \approx \Omega(\ln(\overline{y})) + \Omega'(\ln(\overline{y}))\,[\ln(y_t) - \ln(\overline{y})]$. Again, how does your answer relate to the Barro regressions discussed in class? Can you confirm your answer under (d) when setting $\delta = 1$?

**4.5** Consider the model in Gennaioli et al. (2014).
(a) Explain in words why $\int \widehat{h}_{i,t} di = \int h_{i,t} di$ must hold.
(b) Use (4) in the paper, and $h_{t+1} = \int h_{i,t+1} di$, to show where the expression for the scaling factor $\upsilon_{t+1}$ comes from.
(c) Use (3) and (4) in the paper to derive (5).
(d) Use (5) and $y_{i,t} = A_i h_{i,t}^{\alpha}$ (which you may forward to $t+1$) to derive (6).

**4.6** Consider the paper by Kremer et al. (2021)
(a) Show how they derived eq. (5), i.e., $\beta_t - \beta_t^* = \delta_t \lambda_t$.
(b) Consider two time periods, $t_1$ and $t_2$. Find an expression for the change in the absolute convergence coefficient, $\beta_{t_2} - \beta_{t_1}$, in terms of the four components discussed by Kremer et al. (2021).

**5.** The questions below refer to the sample do file posted on the course website.
(a) Consider the line `gen growth=(F.log_pc_gdp - log_pc_gdp)/10;` This gives an approximation of the growth rate. How would we change the code to compute the exact value? Hint: to generate a new variable `x` that equals the square of per-capita GDP we would write `gen x=pc_gdp^2;`
(b) How would you change the code to drop observations with the 10% smallest population, instead of just the 5% smallest? How many observations (country-years) are dropped in each case?
(c) Consider the scatter plot code `twoway(scatter growth log_pc_gdp if year==1960, msymbol(x) mcolor(red) mlabel(country_name) mlabsize(vsmall) );` Change this code so that the growth rate refers to the period 1970-1980, and initital per-capita GDP to 1970, and such that the plot symbols are circles instead of crosses, and blue instead of red.

**6.** Consider a model of the world where $N > 0$ different countries have different uniform income distributions. In country $i$ the probability density function (pdf) is given by

$$f_i(y) = \begin{cases} a_i & \text{if } y \in [0, \frac{1}{a_i}], \\ 0 & \text{if } y > \frac{1}{a_i}, \end{cases} \tag{3}$$

where $y$ denotes income, and $a_i$ is a parameter that varies across countries.
(a) Illustrate $f_i(y)$ in a diagram. Verify that $f_i(y)$ is a pdf, by showing that it integrates to one, $\int_0^{1/a_i} f_i(y)dy = 1$.
(b) Which country is more equal, one with high $a_i$, or low?
(c) Find the cumulative density function (cdf), defined as $F_i(y) = \int_0^y f_i(\tilde{y})d\tilde{y}$, for all $y \geq 0$. Illustrate $F_i(y)$ in a diagram.
(d) Suppose the $N = 3$ and that $a_1 < a_2 < a_3$. All countries have the same size of population. Find an expression for the world income distribution and draw it in a diagram. (Hint: the cdf's of each country equal one when $y$ reaches above some level that differs across the three countries.)

**7.** Consider Sala-i-Martin's paper. Indicate where in Table I that we can find the four numbers in Figure V referring to the \$570/year poverty line, and show how we can read the same numbers from Figure VI.

**8.** Consider an economy where the poorest 20% earn 10% of total income; the poorest 40% earn 20% of total income; the poorest 60% earn 30% of total income; and the poorest 80% earn 40% of total income. Thus, the richest 20% earn 60% of total income.
(a) Draw the Lorenz curve. Assume it is piece-wise linear with a kink at 80% of the population.
(b) Calculate the Gini coefficient. (Note: the solution has been corrected 7 November, 2017.)

**9.** Create a do file similar to those posted on the course website, starting with the commands `#delimit; drop _all; discard;` Then let the code create a data set of 1000 observations, and a variable called `id` that orders each observation from 1 to 1000, using the commands `set obs 1000; gene id=_n;`
(a) Add a code that creates a variable called `log_y`. For the first 800 observations, `log_y` should equal a random number drawn from a normal distribution with mean 0 and standard

deviation 1, and for the last 200 observations `log_y` should equal a random number drawn from a normal distribution with mean 2 and standard deviation 1. That is, all observations of `log_y` are drawn from a normal distrtibution, but some with a higher mean than others. (Hint: one way to do this is to first create 800 observations using a suitable `if` command, thus leaving some observations empty; then you can replace the missing observations using the command `replace`, together with another `if` command.)
(b) Add a code that creates a histogram of `log_y`, as well as the fitted normal and kernal density plots.
(c) Add a code that creates a plot with `id` on the horizontal azis and `log_y` on the vertical axis.
(d) If `log_y` measures log income, then we can create a variable `y` measuring non-logged income, using `gen y=exp(log_y);` Also, we can create a variable `mean_y` for which all observations equal the mean of `y`, using the command `egen mean_y=mean(y);` Use these insights to write a code that creates a variable `Atkinson`, for which all observations contain the Atkinson inequality index, measured on the variable `y`. Assume an inequality aversion parameter ($\varepsilon$) equal to 0.5.

**10.** Consider an economy with two sectors, indexed A and M, representing agriculture and manufacturing. Total output in each sector is $Y_A$ and $Y_M$, respectively, which are both exogenous. The agricultural output is referred to as food below. The (endogenous) price of food in terms of the manufacturing output is denoted $p$. There is no migration across sectors. A representative agent in each sector earns an income equal to total output of that sector, measured in the good which that sector produces. Thus, the agricultural agent earns $Y_A$ units of food, and the manufacturing agent earns $Y_M$ units of the manufactured good.
The representative agent in sector $i \in \{A, M\}$ has the utility function: $U = (1-\gamma)\ln(C_M^i) + \gamma\ln(C_A^i)$, where $C_M^i$ and $C_A^i$ denotes his/her consumption of the manufactured good, and food, respectively.
(a) Find an expression for how much the agricultural agent spends on food.
(b) Find an expression for how much the manufacturing agent spends on food.
(c) Find an expression for the price of food ($p$) in equilibrium, and illustrate in a supply-and-demand diagram. (Hint: supply of food is the same as total output in the agricultural sector.)
(d) Keeping $Y_A$ constant, how does the well-being of the agricultural agent change when $Y_M$ increases? Explain in words what drives this result.

**11.** Consider a dynamic version of the Murphy-Schleifer-Vishny model, where $n_t$ (the ratio of rent-seekers to market producers in period $t$) evolves according to

$$\begin{array}{lll} n_{t+1} > n_t & \text{if} & R(n_t) > Y(n_t), \\ n_{t+1} = n_t & \text{if} & R(n_t) = Y(n_t), \\ n_{t+1} < n_t & \text{if} & R(n_t) < Y(n_t). \end{array} \tag{4}$$

(a) Explain (4) in words. Phrase your explanation along these lines: "If the payoff to [...] is [...] than the payoff to [...], then over time agents move from [...] to [...]."
(b) Assume $\beta \in (\gamma, \alpha)$. Illustrate all possible steady-state levels of $n_t$ in a diagram with $n_t$ on horizontal axis, and $R(n_t)$ and $Y(n_t)$ on the vertical axis. Show which are stable and unstable, respectively.

**12.** Find and download the dataset for Henderson et al. (2012). In the relevant do file, find the set of commands that replicate Figure 6, Panel B, in the paper.
(a) What is the name of the Stata dataset (ending .dta) used for this figure?
(b) Which four countries are dropped from the plot? Give the names of the countries, not only their codes.
(c) What is the name of the variable on the horizontal axis (x-axis) in the Stata data set?
(d) How would you change the command(s) to generate a best fitted linear regression line in the figure, instead of a non-linear one?

**13.** The questions below refer to Henderson et al. (2012) again.
(a) Consider Figure 2 in the paper. What accounts for the lights in the seas to the east of South Korea? Which Chinese city is most visible in the map?
(b) What do the variables $z_j$ and $\widehat{z}_j$ represent?
(c) From Figure 6, name a country that has had slower growth in GDP/capita than one would expect based on growth in night lights.

Solutions to select problems:

**4.**

(a)

$$y_{t+1} = A\left[\frac{sy_t + (1-\delta)\left(\frac{y_t}{A}\right)^{\frac{1}{\alpha}}}{1+n}\right]^{\alpha}. \tag{5}$$

(b)

$$\overline{y} = A^{\frac{1}{1-\alpha}}\left(\frac{s}{n+\delta}\right)^{\frac{\alpha}{1-\alpha}}. \tag{6}$$

(c) Note from (5) that

$$\frac{y_{t+1}}{y_t} = A\left[\frac{(y_t)^{-\frac{1}{\alpha}}\left[sy_t + (1-\delta)\left(\frac{y_t}{A}\right)^{\frac{1}{\alpha}}\right]}{1+n}\right]^{\alpha} = \frac{A}{(1+n)^{\alpha}}\left[sy_t^{-\left(\frac{1-\alpha}{\alpha}\right)} + (1-\delta)A^{-\frac{1}{\alpha}}\right]^{\alpha}, \tag{7}$$

which gives

$$g_t = \ln(y_{t+1}) - \ln(y_t) = \ln\left(\frac{A}{(1+n)^{\alpha}}\right) + \alpha\ln\left[sy_t^{-\left(\frac{1-\alpha}{\alpha}\right)} + (1-\delta)A^{-\frac{1}{\alpha}}\right]. \tag{8}$$

(d) Setting $\delta = 1$ in (8) gives

$$g_t = \ln\left(\frac{A}{(1+n)^{\alpha}}\right) + \alpha\ln(s) - (1-\alpha)\ln(y_t) = \ln(A) + \alpha\ln\left(\frac{s}{1+n}\right) - (1-\alpha)\ln(y_t). \tag{9}$$

Setting $\delta = 1$ in (6) gives

$$\ln(\overline{y}) = \left(\frac{1}{1-\alpha}\right)\left[\ln(A) + \alpha\ln\left(\frac{s}{1+n}\right)\right]. \tag{10}$$

Now (9) and (10) give

$$g_t = (1-\alpha)[\ln(\overline{y}) - \ln(y_t)] = -(1-\alpha)[\ln(y_t) - \ln(\overline{y})]. \tag{11}$$

(e)

$$g_t = -(1-\alpha)\left(\frac{n+\delta}{1+n}\right)[\ln(y_t) - \ln(\overline{y})]. \tag{12}$$

Note that $\delta = 1$ makes (11) and (12) identical.

**4.5**

(a) We can think of $h_{i,t} - \widehat{h}_{i,t}$ as region $i$'s net export of capital. Since all regions together constitute a closed economy net exports must sum to zero across regions, i.e., $\int\left[h_{i,t} - \widehat{h}_{i,t}\right]di = 0$.

To answer (b)-(d), we first restate eqs. (3), (4), and the expression for $\upsilon_{t+1}$, as given in the paper:

$$\widehat{h}_{i,t+1} = sA_i h_{i,t}^{\alpha}, \tag{13}$$

$$h_{i,t+1} = \upsilon_{t+1}\left(\widehat{h}_{i,t+1}\right)^{\tau}\left(\widehat{A}_i h_{t+1}\right)^{1-\tau}, \tag{14}$$

$$\upsilon_{t+1} = \frac{h_{t+1}^{\tau}}{\int\left(\widehat{h}_{i,t+1}\right)^{\tau}\left(\widehat{A}_i\right)^{1-\tau}di}. \tag{15}$$

(b) Integrate the left- and right-hand sides of (14) to get

$$\begin{aligned} h_{t+1} = \int h_{i,t+1}di = \int \upsilon_{t+1}\left(\widehat{h}_{i,t+1}\right)^{\tau}\left(\widehat{A}_i h_{t+1}\right)^{1-\tau}di \\ = \upsilon_{t+1}h_{t+1}^{1-\tau}\int\left(\widehat{h}_{i,t+1}\right)^{\tau}\left(\widehat{A}_i\right)^{1-\tau}di, \end{aligned} \tag{16}$$

where the last equality moves factors that do not depend on $i$ outside the integral. Solving (16) for $\upsilon_{t+1}$ gives (15).
(c) Using (13) and (14) we get

$$\begin{aligned} h_{i,t+1} &= \upsilon_{t+1}\left(sA_i h_{i,t}^{\alpha}\right)^{\tau}\left(\widehat{A}_i h_{t+1}\right)^{1-\tau} \\ &= \upsilon_{t+1}\left(sA_i\right)^{\tau}h_{i,t}^{\tau\alpha}\left(\widehat{A}_i h_{t+1}\right)^{1-\tau} \\ &= \upsilon_{t+1}\left(sA_i\right)^{\tau}h_{i,t}^{\tau\alpha}\left(\widehat{A}_i \underbrace{s\int A_j h_{j,t}^{\alpha}dj}_{=h_{t+1}}\right)^{1-\tau}, \end{aligned} \tag{17}$$

where we have used (13) and $h_{t+1} = \int \widehat{h}_{j,t+1}dj$. Note that we relabel the variable over which we integrate, since $i$ indexes variables outside the integral. Dividing (17) by $h_{i,t}$ gives (5) in the paper.
(d) Noting that $y_{i,t} = A_i h_{i,t}^{\alpha}$ and $\int y_{i,t}di = y_t$, we can rewrite (17) as

$$\begin{aligned} h_{i,t+1} &= \upsilon_{t+1}sA_i^{\tau}h_{i,t}^{\tau\alpha}\left(\widehat{A}_i\right)^{1-\tau}\left(\int A_j h_{j,t}^{\alpha}dj\right)^{1-\tau} \\ &= \upsilon_{t+1}s\left(A_i h_{i,t}^{\alpha}\right)^{\tau}\left(\widehat{A}_i\right)^{1-\tau}y_t^{1-\tau} \\ &= \upsilon_{t+1}s\left(y_{i,t}\right)^{\tau}\left(\widehat{A}_i\right)^{1-\tau}y_t^{1-\tau}. \end{aligned} \tag{18}$$

Now we see from (18) that

$$y_{i,t+1} = A_i h_{i,t+1}^{\alpha} = A_i \underbrace{\upsilon_{t+1}^{\alpha}s^{\alpha}\left(y_{i,t}\right)^{\alpha\tau}\left(\widehat{A}_i\right)^{\alpha(1-\tau)}y_t^{\alpha(1-\tau)}}_{=h_{i,t+1}^{\alpha}}. \tag{19}$$

Dividing (19) by $y_{i,t}$ and changing the order of the factors gives (6) in the paper.

**5.**
(a) `gen growth=( (F.pc_gdp/pc_gdp)^(1/10) ) - 1;`
(b) `n drop if pct<=2;` Drops 95 observations instead of 48.
(c) `twoway(scatter growth log_pc_gdp if year==1970,`
`msymbol(x) mcolor(red) mlabel(country_name) mlabsize(vsmall) );`

**6.**

(a) The shape of the graph of $f_i(y)$ should be a rectangle with base $1/a_i$, height $a_i$, and area 1.

$$\int_0^{1/a_i} f_i(y)dy = \int_0^{1/a_i} a_i dy = a_i\left(\frac{1}{a_i}\right) - a_i \times 0 = 1 \tag{20}$$

(b) Countries with larger $a_i$ are more equal.

(c)

$$F_i(y) = \min\{1, a_i y\} = \begin{cases} a_i y & \text{if } y \in [0, \frac{1}{a_i}], \\ 1 & \text{if } y > \frac{1}{a_i}. \end{cases} \tag{21}$$

(d) Use your answer to (c) to note that

$$F(y) = \begin{cases} \left(\frac{a_1+a_2+a_3}{3}\right) y & y \in [0, \frac{1}{a_3}] \\ \left(\frac{a_1+a_2}{3}\right) y + \frac{1}{3} & y \in [\frac{1}{a_3}, \frac{1}{a_2}] \\ \left(\frac{a_1}{3}\right) y + \frac{2}{3} & y \in [\frac{1}{a_2}, \frac{1}{a_1}] \\ 1 & y > \frac{1}{a_1} \end{cases} \tag{22}$$

The graph of $F(y)$ should be piece-wise linear, with three kinks, and positive slopes in each segment. The slopes should be flatter for higher levels of $y$.

**8.**

(a) The line should be linear with slope 1/2 from 0 to 0.8 on the horizontal axis, and slope 3 from 0.8 to 1 on the horizontal axis, reaching the coodinate (1,1).

(b) The answer is 0.4. Calculations: The total triangular area under the 45-degree line equals 0.5. The Gini coefficient is the fraction of that area (of size 0.5) which lies between the Lorenz curve and the 45-degree line. The total area below the Lorenz curve is the sum of three sub-areas: one triangle of size $(0.8 \times 0.4)/2 = 0.32/2 = 0.16$; one square of size $0.2 \times 0.4 = 0.08$; and one triangle of size $(0.2 \times 0.6)/2 = 0.12/2 = 0.06$. These sum up to $0.16 + 0.08 + 0.06 = 0.3$. The area between the Lorenz curve and the 45-degree line thus equals $0.5 - 0.3 = 0.2$. Now the Gini coefficient is given by $0.2/0.5 = 0.4$.

**9.**

(a)

```
gen log_y=rnormal(0,1) if id<=800; /* the poor 800 */
replace log_y=rnormal(2,1) if id>800; /* the rich 200 */
```

(b) Following the structure used in Sample do file 2:

```
histogram log_y, bin(100)
normal normopts(lcolor(black) lwidth(thin))
kdensity kdenopts(lcolor(red) lwidth(medium) lpattern(dash) )
legend(on position(2) ring(0))
title(Bimodal distribution);
```

(c)

```
twoway(scatter log_y id);
```

(d)

```
gen y=exp(log_y);
egen mean_y=mean(y);
egen mean_y_root=mean(y^.5);
```

```
gen Atkinson= 1 - ( mean_y_root^2/mean_y );
```

**10.**

(a) The agricultural agent spends $\gamma Y_A$ on food. This follows from maximizing $(1-\gamma)\ln(C_M^A) + \gamma\ln(C_A^A)$ subject to $Y_A = \frac{C_M^A}{p} + C_A^A$; note that the agricultural agent pays $1/p$ units of the agricultural good per unit of the manufactured good.

(b) The manufacturing agent spends $\gamma Y_M/p$ on food. This follows from maximizing $(1-\gamma)\ln(C_M^M) + \gamma\ln(C_A^M)$ subject to $Y_M = C_M^M + pC_A^M$.

(c) Supply is $Y_A$. Demand is $\gamma\left[Y_A + Y_M/p\right]$. Equalizing supply and demand gives is $p = \left(\frac{\gamma}{1-\gamma}\right)\left(\frac{Y_M}{Y_A}\right)$.

(d) The agricultural agent is better off when $Y_M$ increases. To see this, one can write the agricultural agent's utility in terms of $p$. First derive the agricultural agent's optimal spending on the manufactured good from the budget constraint, $C_M^A = p(Y_A - C_A^A) = p(1-\gamma)Y_A$. This gives the maximum utility of the agricultural agent, denoted $U^A$, as follows:

$$\begin{aligned} U^A &= (1-\gamma)\ln(C_M^A) + \gamma\ln(C_A^A) \\ &= (1-\gamma)\ln\left[(1-\gamma)pY_A\right] + \gamma\ln(\gamma Y_A), \end{aligned}$$

which is increasing in $p$, and thus in $Y_M$ (holding constant $\gamma$ and $Y_A$).

[Another way to see the same thing is to note that the maximum amount of food that (s)he can buy is unchanged at $Y_A$, while the maximum amount of the manufacturing good (s)he can buy equals $pY_A = \left(\frac{\gamma}{1-\gamma}\right)Y_M$, which is increasing in $Y_M$. This can be illustrated as an outward shift in the agricultural agent's budget line in an indifference curve diagram.]

Intuitively, when the manufacturing sector grows, the agricultural-sector agent benefits because the price of the good (s)he produces increases.

**11.**

(a)

"If the payoff to rent-seeking is greater than the payoff to production, then over time agents move from production to rent-seeking."

"If the payoff to production is greater than the payoff to rent-seeking, then over time agents move from rent-seeking to production."

(b) See Figure 3 in Murphy, Schleifer and Vishny (1993). Using the notation there, 0 and $n''$ are stable, while $n'''$ is unstable.

**12.**

(a) `global_total_dn_uncal_longdiff9206.dta`

(b) Equatorial Guinea, Bahrain, Singapore, Hong Kong.

(c) `lngdpwdilocallongdiff`

(d) Replace `lowess` by `lfit`

**13.**

(a) The lights in the sea are from fishing boats, which shine lights into the sea to catch, e.g., squid (see p. 2002 in the paper). The Chinese city is Shenyang.

(b) $z_j$ is growth or log difference in GDP/capita from official data, and $\widehat{z}_j$ is the same variable estimated from night lights data. That is, $\widehat{z}_j = \widehat{\psi} x_j$, where $x_j$ is growth or log difference in night lights and $\widehat{\psi}$ is the estimate of $\psi$ from a regression of the type $z_j = \psi x_j + \varepsilon_j$.
(c) For example, the Marshall Islands (MHL) or Zimbabwe (ZWE).