Solutions to review problems for 411-3 midterm

Problem 1: Gini coefficients

Case a: uniform. The average wealth[1] held by people in a uniform distribution $[\underline{x}, \overline{x}]$ is $(\underline{x} + \overline{x})/2$.[2] Let's index people by their percentile $i \in [0, 1]$ in the distribution. The amount of wealth held by someone at percentile i is then $\underline{x} + (\overline{x} - \underline{x})i$, and the share of total wealth held by people of percentile c and below is then

$$\frac{2}{\underline{x} + \overline{x}} \int_0^c (\underline{x} + (\overline{x} - \underline{x})i)\, di = \frac{2}{\underline{x} + \overline{x}} \left(\underline{x}c + (\overline{x} - \underline{x}) \frac{c^2}{2} \right) \tag{1}$$

The relationship (1) between percentile of the distribution and total wealth held under that percentile is the Lorenz curve. Integrating from 0 to 1, the total area under the Lorenz curve is

$$\begin{aligned}
A &= \frac{2}{\underline{x} + \overline{x}} \int_0^1 \underline{x}c + (\overline{x} - \underline{x}) \frac{c^2}{2} dc \\
&= \frac{2}{\underline{x} + \overline{x}} \left(\frac{\underline{x}}{2} + \frac{\overline{x} - \underline{x}}{6} \right) \\
&= \frac{2}{\underline{x} + \overline{x}} \left(\frac{\underline{x}}{3} + \frac{\underline{x}}{6} + \frac{\overline{x} - \underline{x}}{6} \right) \\
&= \frac{2}{\underline{x} + \overline{x}} \left(\frac{\underline{x}}{3} + \frac{\overline{x}}{6} \right) = \frac{2}{3} \frac{\underline{x}}{\underline{x} + \overline{x}} + \frac{1}{3} \frac{\overline{x}}{\underline{x} + \overline{x}}
\end{aligned}$$

Then, the Gini coefficient equals $1 - 2A$, i.e. twice the difference between the maximum possible area under the Lorenz curve ($1/2$) and the actual area A. Implementing this, we get

$$\begin{aligned}
1 - 2A &= 1 - \frac{4}{3} \frac{\underline{x}}{\underline{x} + \overline{x}} - \frac{2}{3} \frac{\overline{x}}{\underline{x} + \overline{x}} \\
&= \frac{\underline{x} + \overline{x} - \frac{4}{3}\underline{x} - \frac{2}{3}\overline{x}}{\underline{x} + \overline{x}} \\
&= \frac{1}{3} \frac{\overline{x} - \underline{x}}{\overline{x} + \underline{x}}
\end{aligned} \tag{2}$$

The expression (2) gives us the Gini coefficient. It's remarkably simple. If $\underline{x}$ is not below zero (i.e. no one has negative wealth), then we see that $\frac{\overline{x} - \underline{x}}{\overline{x} + \underline{x}} \leq \frac{\overline{x} - \underline{x}}{\overline{x}} \leq 1$, so that the maximum possible Gini coefficient when wealth is uniformly distributed on the interval $[\underline{x}, \overline{x}]$ is $\frac{1}{3}$, which is attained whenever $\underline{x} = 0$. When the gap between $\overline{x}$ and $\underline{x}$ is smaller compared to their scale, then the Gini coefficient is lower.

Case b: Pareto. We could do an analysis similar to above, but instead we'll use a trick that we mentioned in class, which is that if wealth is distributed Pareto across people with shape α, the Pareto distribution of "how rich is the person who holds a random dollar of wealth" is Pareto with shape $\alpha - 1$.

If we denote the CDF of the former by F and of the latter by G, and the minimum of the support of the

[1] What the distribution is measuring is unspecified in the question, and it could be income or anything else, but I'll call it "wealth" as a placeholder.

[2] There was some terminological ambiguity here, since people sometimes talk about a "uniform distribution" in the context of wealth as meaning that everyone has the same wealth, in which case the Gini coefficient is zero. Here I literally meant a uniform distribution on the interval $[\underline{x}, \overline{x}]$, with constant density $1/(\overline{x} - \underline{x})$ on the interval and zero elsewhere.

Pareto distribution by $\underline{x}$, then we have

$$\begin{aligned} 1 - F(x) &= \left(\frac{x}{\underline{x}}\right)^{-\alpha} \\ 1 - G(x) &= \left(\frac{x}{\underline{x}}\right)^{-(\alpha-1)} \end{aligned}$$

which combined gives

$$1 - G(x) = (1 - F(x))^{(\alpha-1)/\alpha}$$

The Lorenz curve is the relationship between G and F; letting c be the percentile of the individual wealth distribution, so that $F(x) = c$, then the Lorenz curve is $G(F^{-1}(c))$, or

$$1 - (1 - c)^{(\alpha-1)/\alpha}$$

Integrating with respect to c from 0 to 1, the area under this is

$$\begin{aligned} A &= \int_0^1 1 - (1 - c)^{(\alpha-1)/\alpha} dc \\ &= 1 - \int_0^1 (1 - c)^{(\alpha-1)/\alpha} dc \\ &= 1 - \int_0^1 c^{(\alpha-1)/\alpha} dc \\ &= 1 - \frac{\alpha}{2\alpha - 1} \left[c^{(2\alpha-1)/\alpha}\right]_{c=0}^1 \\ &= 1 - \frac{\alpha}{2\alpha - 1} \end{aligned}$$

Then, the Gini coefficient is

$$\text{Gini} = 1 - 2A = \frac{2\alpha}{2\alpha - 1} - 1 = \frac{1}{2\alpha - 1} \tag{3}$$

We note that this approaches 1 (perfect inequality) as $\alpha \to 1$ from above. This makes sense, because $\alpha \leq 1$ has infinite mean (in which case wealth is infinite and this is undefined), and as we approach it, more and more wealth is held by the extreme richest.

In general, the Gini coefficient is decreasing as we make α higher and the tail thinner, which also makes sense. In the limit $\alpha \to \infty$, all the mass of the Pareto distribution becomes concentrated around $\underline{x}$, and we have perfect equality, and accordingly (3) gives a Gini coefficient of 0.

Problem 2: Pareto

We recall from class that the mean of a Pareto distribution with shape parameter α and minimum $\underline{x}$ is $\frac{\alpha}{\alpha-1}\underline{x}$. Hence $M = \frac{\alpha}{\alpha-1}$, which we can invert to obtain α as a function of M:

$$\begin{aligned} M &= \frac{\alpha}{\alpha-1} \\ M^{-1} &= \frac{\alpha-1}{\alpha} = 1-\alpha^{-1} \\ \alpha^{-1} &= 1-M^{-1} \\ \alpha &= (1-M^{-1})^{-1} \end{aligned} \tag{4}$$

How does this change if instead we define M to be the ratio of the median and the minimum? Recall that the CDF is $1-F(x) = (x/\underline{x})^{-\alpha}$, and the median x^{med} corresponds to $1-F(x^{med}) = 1/2$. In this case, if we write $M \equiv x^{med}/\underline{x}$, we have

$$\begin{aligned} \frac{1}{2} &= \left(\frac{x^{med}}{\underline{x}}\right)^{-\alpha} = M^{-\alpha} \\ 2 &= M^{\alpha} \\ \alpha &= \log_M 2 = \frac{\log 2}{\log M} \end{aligned} \tag{5}$$

So from either the ratio of mean to minimum or the ratio of median to minimum, we can extract the tail parameter α using (4) or (5).

As a side note: if one wants to derive the fact we used above that the mean of a Pareto is $\frac{\alpha}{\alpha-1}\underline{x}$, can just integrate x times the density $f(x) = \alpha x^{-\alpha-1}\underline{x}^{\alpha}$:

$$\begin{aligned} \int_{\underline{x}}^{\infty} x f(x) &= \alpha \underline{x}^{\alpha} \int_{\underline{x}}^{\infty} x^{-\alpha} \\ &= -\frac{\alpha}{\alpha-1}\underline{x}^{\alpha}[x^{-\alpha+1}]_{x=\underline{x}}^{\infty} \\ &= -\frac{\alpha}{\alpha-1}\underline{x}^{\alpha}[0-\underline{x}^{-\alpha+1}] \\ &= \frac{\alpha}{\alpha-1}\underline{x} \end{aligned}$$

Of course, you are free to just remember and cite this fact!

Problem 3: Pareto and taxation

Each income millionaire has exactly \$1 million that is *not* subject to the tax (the first \$1 million that they earn), and all income beyond that is subject to the tax.

The average amount of income subject to the tax among millionaires is then the average income of millionaires, minus \$1 million. The former is $\frac{\alpha}{\alpha-1}$ times \$1 million, where α is the Pareto shape coefficient, so the difference is $\frac{1}{\alpha-1}$ times \$1 million. As a frction of all income among millionaires, which has an average of $\frac{\alpha}{\alpha-1}$ times \$1 million, this is just $\frac{1}{\alpha}$.

So: we conclude that $\frac{1}{\alpha}$ of millionaires' income will be subject to the tax. (Empirically, we found an income α of around 1.64, so this would be $1/1.64 \approx 61\%$, which is quite a bit, but far less than 100%, which is what people might naively assume!)

Problem 4: model of age-dependent wealth distribution

Population dynamics. Let $n_{t,s}$ be the density of people who were born at date s who are still alive at date t, and let $N_t \equiv \int_{-\infty}^{t} n_{t,s} ds$ be the total measure of people alive at date t. Our assumption is that $n_{t,t} = \phi N_t$, i.e. that the flow rate of births equals ϕ times the total population.

We also assumed that there is a constant flow probability η of dying, so that the probability of survival for someone born u periods ago is $e^{-\eta u}$. Combined with the above, we have that

$$n_{t,s} = \phi e^{-\eta(t-s)} N_s \tag{6}$$

Further, the constant rate η of dying implies that the flow rate of deaths in the population is ηN_t. Combining with the gain in population from births ϕN_t, the rate of change in the population is given by the simple ODE

$$\dot{N}_t = (\phi - \eta) N_t$$

implying that

$$N_t = e^{(\phi-\eta)t} N_0 \tag{7}$$

for any t. Combining this with (6) we have the evolution of cohort sizes in the population.

Assets of each cohort. Now, let us denote the assets held at date t by someone who was born at date s by $a_{t,s}$. By assumption, there is a constant rate $ra_{t,s}$ of earnings on wealth, minus consumption $ca_{t,s}$, so that we have an ODE in t

$$\dot{a}_{t,s} = (r - c) a_{t,s}$$

and combined with the assumption that $a_{s,s} = 1$, i.e. everyone is born with 1 dollar, we have

$$a_{t,s} = e^{(r-c)(t-s)}$$

Distribution of assets. Summarizing what we have so far: at any moment t, we have cohorts $s \in (-\infty, t]$ of size $n_{t,s} = \phi e^{-\eta(t-s)} N_s = \phi e^{-\eta(t-s)} e^{(\phi-\eta)s} N_0$, each with assets $a_{t,s} = e^{(r-c)(t-s)}$. Assuming that $r > c$, assets have a minimum of 1 and are increasing in age $t - s$.

Holding t constant and looking across $s \in (-\infty, t]$, the size of cohorts is proportional to $e^{\phi s}$, or equivalently proportional to $e^{-\phi(t-s)}$. In other words, the population is distributed exponentially by age with parameter ϕ. Letting $F(t-s)$ be the fraction of people with age less than $t-s$, it follows that $F(t-s) = 1 - e^{-\phi(t-s)}$.

Now let $G(a)$ be the CDF of the asset distribution. $G(a)$ is the fraction of people with assets less than a, which is equal to $F(t-s)$ for whatever age $t-s$ has assets $a = e^{(r-c)(t-s)}$, which is $t - s = \frac{\log a}{r-c}$. Hence

$$\begin{aligned} G(a) &= F\left(\frac{\log a}{r-c}\right) \\ &= 1 - e^{-\phi \frac{\log a}{r-c}} = 1 - a^{-\frac{\phi}{r-c}} \end{aligned} \tag{8}$$

We conclude that the asset distribution here is Pareto with shape parameter $\frac{\phi}{r-c}$.

The shape parameter decreases—and therefore the Pareto tail becomes fatter and wealth inequality is higher—when the return net of consumption $r - c$ is larger, because then the older, larger fortunes grow more relative to the younger ones.

The shape parameter increases—and therefore the Pareto tail becomes thinner and wealth inequality is lower—when the birth rate ϕ is larger. Intuitively, this is because older cohorts, which in this model hold more wealth, become relatively smaller with a higher birth rate.

Why does only the birth rate ϕ show up here, and not the death rate η? I was a bit puzzled by this at first. The reason, as we can see in the derivation above, is that the distribution of the population by age is exponential with rate ϕ, and does not depend on η. An intuition for this, in turn, is that $\phi = (\phi - \eta) + \eta$, i.e. that is the sum of the growth rate of the population $\phi - \eta$ and the death rate η. The size of older generations, relative to newborn ones, shrinks for two reasons: (1) the overall population is growing at rate $\phi - \eta$, making those older generations smaller by comparison even if they don't die, and (2) they die at rate η.[3] Adding these two forces together, η drops out and we get ϕ.

Finally, we've seen that the asset distribution is Pareto as long as $r > c$. If $r = c$, then the distribution is degenerate, with everyone having $a = 1$. If $r < c$, then assets shrink with age, and no longer follow a Pareto distribution. Instead, in this case, we can go through the same steps above to show that the *inverse* of assets a^{-1} follows a Pareto distribution with shape parameter $\frac{\phi}{c-r}$. But this is very different from a itself being Pareto distributed!

To conclude: in this model, the condition for the asset distribution being Pareto is $r > c$, and in that case the shape parameter is $\frac{\phi}{r-c}$.

Problem 5: hybrid Bewley/Aiyagari model

Assume that we have made assumptions (CRRA preferences, zero borrowing limit) such that the ratio of aggregate assets to labor income, A/wL, is given by some function $a(r)$ that depends only on the real interest rate. (I should have specified this in the question.)

Further, assuming Cobb-Douglas production $Y = K^\alpha L^{1-\alpha}$, we have a constant labor share $wL/Y = 1-\alpha$, so that $A/Y = (A/wL) \cdot (wL/Y) = (1-\alpha)a(r)$. We also have a constant capital share $(r+\delta)K/Y = \alpha$, so that $K/Y = \alpha/(r+\delta)$.

Writing asset market clearing

$$A = K + B$$

and dividing both sides by Y, we get

$$\begin{aligned} \frac{A}{Y} &= \frac{K}{Y} + \frac{B}{Y} \\ (1-\alpha)a(r) &= \frac{\alpha}{r+\delta} + b \end{aligned} \tag{9}$$

where b is the assumed target for B/Y. (9) can be solved to obtain equilibrium r.

[3] Of course, a constant death rate is extremely stylized relative to a realistic lifecycle model!

Taking logs of (9) and totally differentiating with respect to r and $\log b$, we obtain

$$\begin{aligned}
d\log a(r) &= d\log\left(\frac{\alpha}{r+\delta}+b\right) \\
\epsilon_r^d dr &= \frac{K}{A}\cdot d\log\left(\frac{\alpha}{r+\delta}\right)+\frac{B}{A}\cdot d\log b \\
\epsilon_r^d dr &= -\frac{K}{A}\frac{dr}{r+\delta}+\frac{B}{A}d\log b \\
\left(\epsilon_r^d+\frac{K}{A}\frac{1}{r+\delta}\right)dr &= \frac{B}{A}d\log b \\
dr &= \left(\epsilon_r^d+\frac{K}{A}\frac{1}{r+\delta}\right)^{-1}\frac{B}{A}d\log b
\end{aligned} \tag{10}$$

where as in class, $\epsilon_r^d \equiv \partial\log a/\partial r$. We see that a log increase in the debt-to-GDP target b has an effect on real interest rates that is proportional to the share of debt in assets (naturally enough), divided by the sum of the asset demand semielasticity ϵ_r^d and the asset supply semielasticity, which in this case is the semielasticity of capital $\frac{1}{r+\delta}$ times capital's share in assets.

A larger share of bonds in assets $\frac{B}{A}$ therefore implies a larger change dr for two reasons: both the simple fact that $\frac{B}{A}\log b$ is larger, and the more subtle point that since $\frac{K}{A}=1-\frac{B}{A}$, a higher $\frac{B}{A}$ implies a smaller asset supply response because there is less capital to respond.

The change in log capital-output ratio is then

$$d\log\left(\frac{\alpha}{r+\delta}\right) = -\frac{dr}{r+\delta}$$

Problem 6: simple process for income and consumption

The assumptions stated imply that

$$dC_t = 0.5\cdot dY_t + 0.5\cdot dY_{t-1}$$

If dY_t is iid with standard deviation σ (and assuming mean 0), then we have

$$\begin{aligned}
\mathrm{Var}(dC_t) &= \mathbb{E}[dC_t^2] = \mathbb{E}[(0.5\cdot dY_t + 0.5\cdot dY_{t-1})^2] \\
&= 0.25\cdot\mathbb{E}[dY_t^2] + 0.5\cdot\mathbb{E}[dYdY_{t-1}] + 0.25\cdot\mathbb{E}[dY_{t-1}^2] \\
&= 0.5\sigma^2
\end{aligned}$$

since the cross term is 0 because of the iid assumption: $\mathbb{E}[dYdY_{t-1}]=0$. We also have

$$\begin{aligned}
\mathrm{Cov}(dC_t, dC_{t-1}) &= \mathbb{E}[dC_t dC_{t-1}] \\
&= \mathbb{E}[(0.5\cdot dY_t + 0.5\cdot dY_{t-1})(0.5\cdot dY_{t-1} + 0.5\cdot dY_{t-2})] \\
&= 0.25\cdot\mathbb{E}[dY_t dY_{t-1}] + 0.25\cdot\mathbb{E}[dY_{t-1}dY_{t-2}] + 0.25\cdot\mathbb{E}[dY_t dY_{t-2}] + 0.25\cdot\mathbb{E}[dY_{t-1}^2] \\
&= 0.25\sigma^2
\end{aligned}$$

where all terms drop out except $0.25\cdot\mathbb{E}[dY_{t-1}^2]$.

We conclude that the autocorrelation[4] of consumption in successive periods is

$$\text{Corr}(dC_t, dC_{t-1}) = \frac{\text{Cov}(dC_t, dC_{t-1})}{\text{Var}(dC_t)} = \frac{0.25\sigma^2}{0.5\sigma^2} = \frac{1}{2}$$

Problem 7: two-agent New Keynesian model with different taxes

(There was a typo where this was written "two-asset" rather than "two-agent" in the question.)

Recall that we had $\mathbf{M}^{TA} \equiv \mu\mathbf{I} + (1-\mu)\mathbf{M}^{RA}$ in the two-agent model. But now, if all taxes are assessed on the "saver", then the matrix $\mathbf{M}^{RA}$ will govern consumption out of taxes, and the intertemporal Keynesian cross is actually

$$d\mathbf{Y} = d\mathbf{G} - \mathbf{M}^{RA}d\mathbf{T} + \mathbf{M}^{TA}d\mathbf{Y}$$

Now, expanding $\mathbf{M}^{TA}$ and rearranging, this becomes

$$\begin{aligned}
d\mathbf{Y} &= d\mathbf{G} - \mathbf{M}^{RA}d\mathbf{T} + (1-\mu)\mathbf{M}^{RA}d\mathbf{Y} + \mu d\mathbf{Y} \\
(1-\mu)d\mathbf{Y} &= d\mathbf{G} - \mathbf{M}^{RA}d\mathbf{T} + (1-\mu)\mathbf{M}^{RA}d\mathbf{Y} \\
d\mathbf{Y} &= \frac{1}{1-\mu}(d\mathbf{G} - \mathbf{M}^{RA}d\mathbf{T}) + \mathbf{M}^{RA}d\mathbf{Y}
\end{aligned}$$

This is the same as the usual intertemporal Keynesian cross for the representative agent $d\mathbf{Y} = d\mathbf{G} - \mathbf{M}^{RA}d\mathbf{T} + \mathbf{M}^{RA}d\mathbf{Y}$, just with larger shocks $\frac{1}{1-\mu}d\mathbf{G}$ and $\frac{1}{1-\mu}d\mathbf{T}$ to spending and taxes, respectively.

The solution is therefore just the solution we have already identified for the representative-agent case given this larger shock, i.e.

$$d\mathbf{Y} = \frac{1}{1-\mu}d\mathbf{G} \tag{11}$$

This is in contrast to the basic "two-agent" model that we derived in class where the solution was instead $d\mathbf{Y} = \frac{1}{1-\mu}d\mathbf{G} - \frac{\mu}{1-\mu}d\mathbf{T}$. The key difference is that taxes do not appear here in (11), because they are no longer being assessed on any hand-to-mouth households.

An interesting feature of (11) is that it no longer has a balanced-budget multiplier of 1, since even if $d\mathbf{G} = d\mathbf{T}$ we still have a multiplier of $\frac{1}{1-\mu}$ on government spending! This is an interesting case of the balanced-budget multiplier result breaking when the incidence of income and taxes is different. In particular, here, income goes partly to the high-MPC hand-to-mouth households, but taxes are not assessed on those same households; the overall effect is for increases in spending and taxes to be more expansionary.

[4]Normally the correlation of two variables is defined as their covariance divided by the product of their standard deviations, but given stationarity here, the standard deviations are the same and their product is just the variance of dC_t.