

**Econ 230B**  
**Spring 2019**

**FINAL EXAM: Solutions**

The average grade for the final exam is 42.5 (out of 60 points). The final exam is reweighed to 70 points. The average grade (out of 100) including all assignments is XX. The distribution of course grades is:

4 A+, 5 A, 6 A-, 2 B+, 4 B, 3 B-.

**True/False Questions: 30 points**

Answer all 10 questions (3 pts each). Explain your answer fully, since all the credit is based on the explanation.

Only short answers provided here. Full detailed in the class notes and relevant references.

1. Disability insurance has negative effects on labor supply because applicants assigned to tough examiners are more likely to work than applicants assigned to lenient examiners.

True as shown by the Maestas et al. AER'13 study that exploits random assignment to examiners of DI applicants. However, the negative labor supply effect is not very large as getting DI reduces LFP by about 25-30 points, confirming the earlier results by Bound AER'89 that most DI beneficiaries actually cannot work (because DI rejected applicants themselves are not able to work much).

2. The main reason behind the surge in labor force participation of single mothers in the US in the 1990s is the expansion of the Earned Income Tax Credit.

Uncertain: it is true that the surge in labor force participation of single mothers in the US in the 1990s coincided with the expansion of the Earned Income Tax Credit. But welfare reform also happened at the same time. The old literature believed that the EITC was the key element but recent work by Kleven (2019) has cast doubt on this: other EITC expansions did not increase LFP of single mothers. Hence, it is likely that a combination of EITC, welfare reform, and changes in social norms explain the surge in the LFP of single mothers.

3. If the elasticity of taxable income of upper income taxpayers with respect to the net-of-tax

rate is high, it is self-defeating for the government to impose a very high top marginal tax rate on the rich.

True if the elasticity is due to real labor supply responses. (as the revenue maximizing tax rate at the top is given by  $\tau = 1/(1 + a \cdot e)$  (Saez, Restud'01). However, a high taxable income elasticity is in general due to tax avoidance or evasion. By changing the definition of taxable income (broader base, fewer deductions, elimination of tax favored income items), it is possible to reduce the elasticity of taxable income and then increase the level of redistribution from rich to poor. Furthermore, if the elasticity is due to bargaining effects (as in Piketty-Saez-Stantcheva AEJ'14), then a high top tax rate is actually desirable to prevent "over compensation" of top earners at the expense of bottom 99%.

4. The spike in retirement hazard at the Early Retirement Age of the Social Security system is evidence that many individuals do not follow rational model of life-cycle savings.

True: a rational person is not affected by the ERA because somebody who wants to retire earlier than the ERA would save in advance and live off savings before getting the benefits at ERA.

5. Preferential tax systems for highly skilled foreign immigrants have a large positive effect on immigration and hence are desirable even if society cares about redistribution.

True/False: True that preferential tax systems can sometimes have a large effect (study on Denmark by Kleven et al.) so they are desirable for tax revenue reasons from a 1 country perspective. However, from a multi-country perspective, they create harmful tax competition as the gains in tax revenue come at the expense of tax revenue losses in other countries.

6. In the standard Harberger (1962) model, part of the incidence of the corporate tax is on labor.

**Solution:** Wrong: in the Harberger model, 100% of the corporate tax falls on capital. However the corporate tax does not fall only on shareholders, but also on the owners of non-corporate capital (business assets, housing ,etc.).

7. The optimal tax rate on bequest is lower when there is more social mobility.

**Solution:** True. In the Piketty and Saez (2013) model, the optimal tax rate  $\tau^B$  is lower when  $\bar{b}$  (the ratio of the average wealth at death for 0-bequest-receivers to the average wealth at death in the entire population, a measure of wealth mobility) is higher.

8. International tax competition means that a greater fraction of the corporate income tax is shifted to labor in high-tax countries.

**Solution:** Generally speaking this is false: it depends on whether capital moves to low-tax places in response to tax competition. If only paper profits move to low tax places (or if location choices are completely inelastic), tax competition does not reduce wages. However there is some evidence in the data (e.g., the higher capital intensity of U.S. multinationals' affiliates in low-tax places in Wright and Zucman, 2018) that some capital has moved to low-tax places.

9. With high audit rates at the top of the income distribution, tax evasion among the rich can be reduced to zero.

**Solution:** Generally speaking this is false, because not all tax evasion can be detected. What matters more in the existence of third-party information reporting (which makes detecting evasion possible) and the regulation of the supply of tax evasion services.

10. Taxing capital reduces wealth inequality in the long-run.

**Solution:** This is true in a broad class of models (dynamic random shock models) where the long-run concentration of wealth depends on the gap between the net of tax rate of return to wealth and the economy's growth rate.

**PROBLEM (30 pts):**

(a) (3 pts)  $\max_l wl(1 - \tau) + R - l^{1+k}/(1 + k) \Rightarrow w(1 - \tau) = l^k \Rightarrow l = [w(1 - \tau)]^{1/k}$  so  $\varepsilon^u = \varepsilon^c = 1/k$  and  $\eta = 0$ . Let us denote by  $e = 1/k$  the common uncompensated and compensated elasticity.

(b) (3 pts)  $\max_\tau \tau(1 - \tau)^{1/k} \int w^{1+1/k} f(w)dw \Rightarrow \tau^* = k/(k + 1) = 1/(1 + e)$ .

Worst off individual has  $w = 0$  and hence  $l = 0$  and utility  $u = R = \tau Z$  so Rawlsian optimal rate maximizes tax revenue (to maximize  $R$ ) and is set at  $\tau^* = 1/(1 + e)$  from (b). Given that all utilities are linear, there is no concern for redistribution and hence the optimal utilitarian tax rate is zero.

(c) (2 pts) Standard plot.

(d) (4 pts) The budget has a kink generating bunching at  $\bar{z}$ .

Case 1 (first bracket):  $w \leq \underline{w}$ :  $l = [w(1 - \tau_1)]^e$  with  $\underline{w}$  s.t.  $\underline{w}^{1+e}(1 - \tau_1)^e = \bar{z}$

Case 2 (bunching at  $\bar{z}$ ):  $\underline{w} \leq w \leq \bar{w}$ :  $l = \bar{z}/w$

Case 3 (top bracket):  $w \geq \bar{w}$ :  $l = [w(1 - \tau_2)]^e$  with  $\bar{w}$  s.t.  $\bar{w}^{1+e}(1 - \tau_2)^e = \bar{z}$

(e) (3 pts) Amount of bunching is proportional to  $e$  (see Saez AEJ:EP'10 and class notes for details):

$$\frac{\bar{w}^{1+e}}{\underline{w}^{1+e}} = \left( \frac{1 - \tau_1}{1 - \tau_2} \right)^e$$

$$T = \lambda_1 \tau_1 (1 - \tau_1)^e \underline{w}^{1+e} + \lambda_2 \{ \tau_1 \bar{z} + \tau_2 [(1 - \tau_2)^e \bar{w}^{1+e} - \bar{z}] \}$$

(f) (2 pts) take the FOC wrt to  $\tau_2$  to get:

$$\tau_2 / (1 - \tau_2) = (1/e) \cdot [z_2 - \bar{z}] / z_2$$

(g) (4 pts) Take the FOC wrt to  $\tau_1$  to get:

$$\tau_1 / (1 - \tau_1) = (1/e) \cdot [1 + \lambda_2 \bar{z} / (\lambda_1 z_1)]$$

Explaining:  $\tau_2^* < \tau^* < \tau_1^*$ :

1) Increasing the flat tax rate  $\tau$  creates a mechanical increase in revenue proportional to average earnings and creates a negative behavioral response proportional to average earnings as well.

2) Increasing the tax rate  $\tau_2$  in the top bracket creates a mechanical increase in revenue proportional to  $(z_2 - \bar{z})$  but creates a negative behavioral response proportional to  $z_2$ .

3) Increasing the tax rate  $\tau_1$  in the bottom bracket creates a mechanical increase in revenue proportional to  $z_1$  and creates a negative behavioral response proportional to  $z_1$ . However, the tax rate increase also raises more tax from high skilled worker with no negative behavioral response (inframarginal tax).

(h) (5 pts) Disabled workers work iff utility when working is higher than utility when not working. Utility when not working is  $R$ . When working,  $l = w_1^e(1 - \tau_1)^e$  and hence  $u = R + w_1^{1+e}(1 - \tau_1)^{1+e}/(1 + e)$ . Thus, a disabled person will work iff:

$$q \leq w_1^{1+e}(1 - \tau_1)^{1+e}/(1 + e) = \bar{q}.$$

Hence the fraction working is  $P(\bar{q})$ . Note that  $\bar{q}$  is decreasing in  $\tau_1$  and  $dP/d\tau_1 = -P'(\bar{q})z_1$ .

Under this scenario,

$$T = [\lambda_1 + \lambda_0 P(\bar{q})]\tau_1(1 - \tau_1)^e w_1^{1+e} + \lambda_2\{\tau_1 \bar{z} + \tau_2[(1 - \tau_2)^e w_2^{1+e} - \bar{z}]\}$$

Taking the FOC in  $\tau_1$ , one gets:

$$0 = dT/d\tau_1 = -\lambda_0 P'(\bar{q})z_1 \tau_1 z_1 + [\lambda_1 + \lambda_0 P(\bar{q})]z_1[1 - e\tau_1/(1 - \tau_1)] + \lambda_2 \bar{z}$$

$$e \cdot \tau_1/(1 - \tau_1) = 1 + \lambda_2 \bar{z}/(\lambda z_1) - z_1 \lambda_0 P' \cdot \tau_1 z_1/(\lambda \bar{z})$$

where  $\lambda = \lambda_1 + \lambda_0 P(\bar{q})$ .

There are two additional effects relative to (d):

1) There are more low skilled workers (relative to high skilled workers):  $\lambda > \lambda_1$ . This makes the tax  $\tau_1$  less desirable

2) There is another layer of response to  $\tau_1$  through  $\bar{q}$  which makes the behavioral response to  $\tau_1$  larger.

Those two effects will make the new optimal  $\tau_1$  smaller than  $\tau_1^*$ .

(i) (4 pts) Tax reform analysis.

Select the top  $x\%$  share where  $x\%$  is the fractile in the 3rd bracket affected by the reform. Plot the share of income going to the top  $x\%$  in all years. If the graph looks like a step function, that's convincing evidence of a tax effect. Key assumption is no change in income concentration from before to after the reform [absent any tax change]. This can be checked (imperfectly) by

seeing whether the share of income going to a control group not affected by the reform [but fairly close in income level to the affected group] is stable.

An entirely different method would be a bunching method in the years after the method.