Exam in Public Finance – Spring 2015
3-hour closed book exam
(Answers)

Part 1: Questions on various topics

**(1A)** Using only time-variation for a group affected by the reform (i.e., a treatment group) can in principle provide a causal estimate of the reform effect, but is unlikely to do so. In order to provide a causal estimate it is required that there are no other shocks affecting the outcome in question (taxable income) that coincide with the reform. For example, a secular trend in taxable income due to economic growth or business cycles may bias the estimate of the reform effect. In general, it is impossible to rule out that there are no other shocks coinciding with the reform. The student may note (but needs not) that one way to argue that "other shocks" are less of a problem is by showing a longer time series, indicating a constant level before and after the reform with the only change occuring in the reform year. But, of course, this is not proof that there is no bias, it is only consistent with the case of no bias.

**(1B)** The Pareto criterion states that an outcome is Pareto optimal if there is no other outcome that leaves everyone at least as well-off and at least one individual strictly better off.

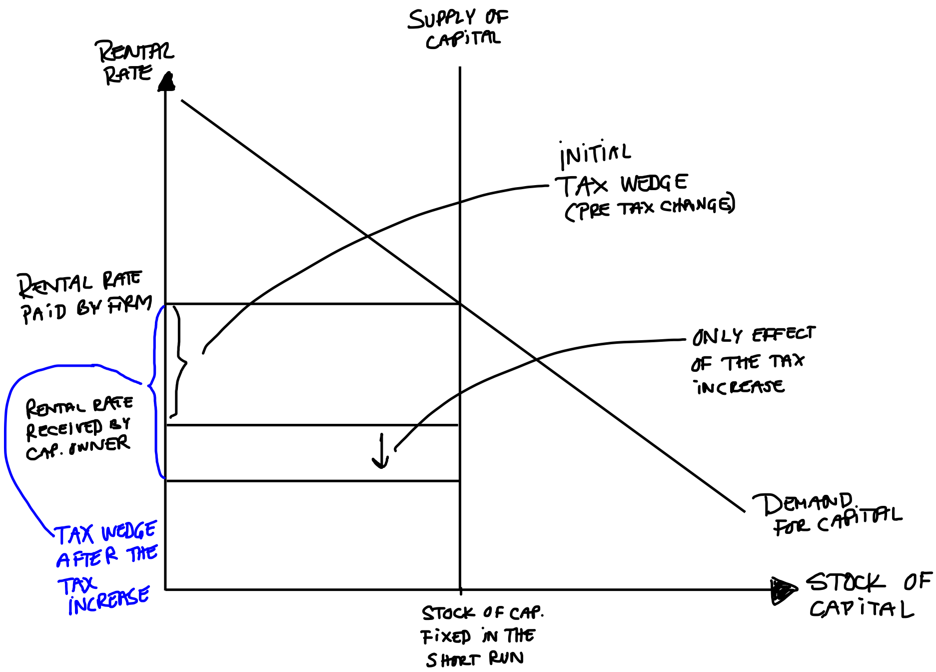
In general, the Pareto criterion will not provide a unique social optimum. The Pareto criterion is based on the ordinal utility concept, which does not allow interpersonal comparison. Therefore, it will not allow the social planner to weight marginal utility gains and losses across individuals. In that sense, the Pareto criterion is more closely related to an efficiency criterion, and it is not suitable for answering distributional questions. For example, *i*) an outcome where one individual consumes everything can be Pareto optimal as long as there is no pure waste of resources and production is efficient, but also *ii*) an outcome where everyone consumes the same amount of all goods can be Pareto optimal (with no pure waste of resources and efficient production). The Pareto criterion does not discriminate between the two cases—it does not entail any particular sort of distributional preference.

Students may also discuss the two welfare theorems and how endowments being private information to the taxpayers lead to a second best outcome in which a Pareto improvement in principle is possible with individualized lump-sum transfers. In other words, searching for a Pareto efficient outcome may lead to infeasible outcomes given the informational restrictions.

**(1C)** The key to answering the question is to note that the capital stock is fixed in the short run, so that supply of capital is perfectly inelastic (vertical in the figure below).

Due to perfect competition in the input markets, firms are willing to pay a rental rate exactly equal to the marginal productivity of capital for renting a unit of capital. With constant returns to scale, the marginal product of capital is decreasing in capital, given a fixed level of labor input in production. This implies that the demand for capital has a decreasing slope in the figure below.

The tax on capital increases, but as capital stock is fixed, there is no change in either the marginal productivity of capital or labor, and therefore the firm is not willing to pay more for renting capital. This means that the demand for capital is unaffected. Since the tax wedge on capital input has increased, the owners of capital will get a smaller after-tax compensation for renting their capital to the firm, whereas the firm is still paying the same rental rate. This is illustrated in the figure below.




The labor market is unaffected as the capital stock, and therefore the marginal product of labor, is unaffected in the short run. Ergo, the capital owners are bearing the full burden of the capital tax in the short run.

Part 2: The revenue-maximizing high income tax rate

**(2A)** The elasticity of taxable income with respect to the net-of-tax rate $(1-m)$ is defined as the percentage change in taxable income due to a one percentage change in the net-of-tax rate. This can be written as

$$\frac{dz/z}{d(1-m)/(1-m)} = \frac{d\log z}{d\log(1-m)}.$$

For calculation purposes, we may rewrite it to

$$\frac{dz}{d(1-m)} \cdot \frac{(1-m)}{z}$$

or use the $d\log$-specification.

To calculate the elasticity for a taxpayer with income $z > \hat{z}$, we use the information given in the text that in optimum such a taxpayer has taxable income $z_i = [a_i(1-m_H)]^{\varepsilon}$. Now we can either write taxable income as $\log z_i = \varepsilon \log a_i + \varepsilon \log(1-m_H)$ and differentiate

$$\frac{d\log z_i}{d\log(1-m_H)} = \varepsilon$$

or note that

$$\frac{dz_i}{d\left(1-m_H\right)} = \varepsilon \cdot \left[a_i\left(1-m_H\right)\right]^{\varepsilon-1} \cdot a_i = \varepsilon \cdot \frac{z_i}{a_i\left(1-m_H\right)} \cdot a_i = \varepsilon \cdot \frac{z_i}{\left(1-m_H\right)}$$

from which it follows that the elasticity is

$$\frac{dz_i}{d(1-m_H)} \cdot \frac{(1-m_H)}{z_i} = \varepsilon.$$

In either case, we get that the elasticity of taxable income with respect to the net-of-tax rate is $\varepsilon$ for taxpayers above the threshold $\hat{z}$.

**(2B)** The "mechanical revenue effect" denotes the effect on government tax revenue keeping taxpayer behavior/the tax base constant. If we are considering a tax reform that raises a tax rate, the mechanical revenue effect will therefore be positive.

Given the definition, the mechanical effect from increasing the marginal tax rate $m_H$ for the $N$ individuals above the threshold is the sum

$$dM = \sum_{i=1}^{N}\left(z_i - \hat{z}\right) \cdot dm_H$$

where the affected tax base for each individual is $(z_i - \hat{z})$ as the tax increase only applies to income above $\hat{z}$. Now use the definition of average income above the threshold to see that

$$\begin{aligned} dM &= \left(\sum_{i=1}^{N} z_i\right) \cdot dm_H - N \cdot \hat{z} \cdot dm_H \\ &= N \cdot \bar{z} \cdot dm_H - N \cdot \hat{z} \cdot dm_H \\ &\Updownarrow \\ dM &= (\bar{z} - \hat{z}) \cdot N \cdot dm_H \end{aligned}$$

**(2C)** The revenue-maximizing tax rate, $m_H^*$, is the tax rate that has the property that a small change in the tax rate ($dm_H$) induces no change in government tax revenue:

$$dR = dM + dB = 0,$$

given that second-order conditions for a global maximum are fulfilled ($d^2R/dm_H^2 < 0$). Inserting for $dM$ and $dB$ we get

$$\begin{aligned}
\Rightarrow 0 &= (\bar{z} - \hat{z}) \cdot N \cdot dm_H - \varepsilon \cdot \frac{m_H^*}{1 - m_H^*} \cdot \bar{z} \cdot N \cdot dm_H \qquad (1) \\
&= \bar{z} - \hat{z} - \varepsilon \cdot \frac{m_H^*}{1 - m_H^*} \cdot \bar{z} \\
&= 1 - \varepsilon \cdot \frac{m_H^*}{1 - m_H^*} \cdot \underbrace{\frac{\bar{z}}{\bar{z} - \hat{z}}}_{\equiv \alpha} \\
\Rightarrow \varepsilon \cdot \alpha \cdot m_H^* &= 1 - m_H^* \\
\Rightarrow (1 + \varepsilon \cdot \alpha) \cdot m_H^* &= 1 \\
\Rightarrow m_H^* &= \frac{1}{1 + \varepsilon \cdot \alpha}.
\end{aligned}$$

As the tax rate increases, the behavioral revenue effect (second term in (1)) becomes increasingly larger. A tax rate above $m_H^*$ will reduce tax revenue because the negative behavioral effect on government revenue ($dB$) becomes large relative to the positive mechanical effect on government revenue ($dM$).

The revenue-maximizing high-income tax rate depends negatively on the two parameters $\varepsilon$ and $\alpha$.

$\varepsilon$ affects the strength of the behavioral response from a tax increase. With a positive marginal tax rate, if $\varepsilon$ is larger, the negative behavioral response to a tax increase will be larger, which lowers the revenue-maximizing rate $m_H^*$.

$\alpha$ measures the relative strength of the mechanical revenue effect. $\alpha = \bar{z}/(\bar{z} - \hat{z})$ is larger when the income distribution is relatively compressed above the threshold (i.e., $\bar{z} - \hat{z}$ is small), so that the amount of tax base located above the threshold is smaller. Recall that the tax rate $m_H$ only applies to income above $\hat{z}$. This means that the positive mechanical revenue effect from an increase in the tax rate is smaller, given the size of the behavioral response. The implication is that a higher $\alpha$ implies a lower revenue-maximizing rate $m_H^*$.

## Part 3: Extensive labor supply responses

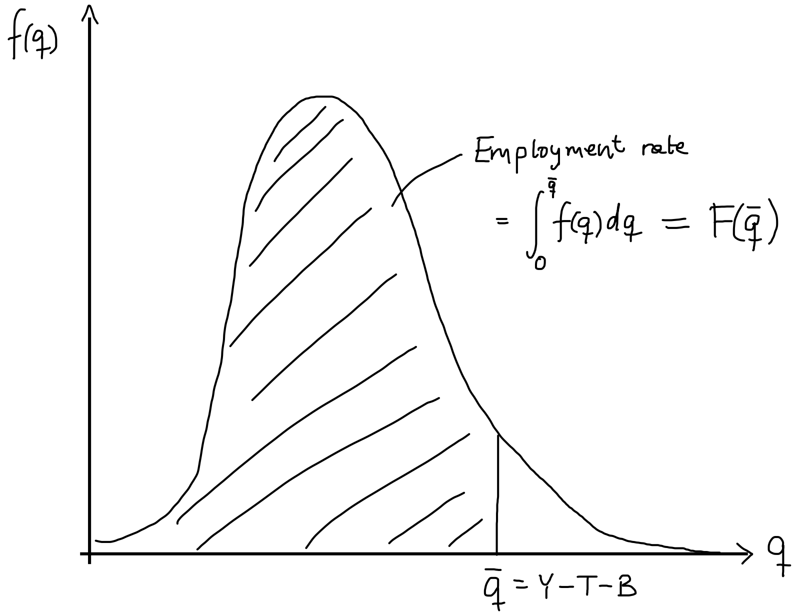
**(3A)** An individual is just willing to work (i.e., indifferent between working or not) if and only if

$$u_1 = u_0,$$

where $u_1 = Y - T - q$ denotes utility if working, and $u_0 = B$ denotes utility if not working. The marginal individual therefore has a fixed cost, $\bar{q}$, that satisfies

$$\begin{aligned}
Y - T - \bar{q} &= B \Leftrightarrow \\
\bar{q} &= Y - T - B.
\end{aligned}$$

The diagram below shows an example of a density $f(q)$ with a choice of $\bar{q}$. Those willing to work are individuals with $u_1 \geq u_0 \Leftrightarrow \bar{q} \leq Y - T - B$, i.e., those with a fixed cost lower than the net benefit from going from not working to working. These individuals are shown as the hatched area under the graph to the left of $q = \bar{q}$. The total mass/number of individuals is 1 (the density integrates to 1) so the hatched area, $F(\bar{q})$, is exactly the employment rate.




**(3B)** The threshold fixed cost $\bar{q}$ is decreasing in both the tax $T$ and the benefit $B$ according to

$$\frac{d\bar{q}}{dT} = \frac{d\bar{q}}{dB} = -1.$$

That is, $\bar{q}$ shifts left in the diagram when either $T$ or $B$ increases. This implies a drop in the employment rate as

$$\frac{dF(\bar{q})}{dT} = \frac{\partial F(\bar{q})}{\partial \bar{q}} \cdot \frac{\partial \bar{q}}{\partial T} = f(\bar{q}) \cdot (-1) = -f(\bar{q}) < 0$$

and using the same argumentation

$$\frac{dF(\bar{q})}{dB} = -f(\bar{q}) < 0.$$

The economic interpretation: $T + B$ is the effective tax on labor market participation, which is the difference in disposable income in the working state versus the unemployed state. When the participation tax increases, more individuals (those with progressively lower and lower fixed cost of working) will no longer find it worthwhile to work. The size of the drop in employment depends on the mass/number of taxpayers on the margin of working/not working, $f(\bar{q})$.

**(3C)** There are two questions here. First, an EITC lowers the tax burden on low-income earners on the labor market without increasing disposable income for those outside the labor market. That is, $T$ decreases and $B$ is unchanged. This amounts to a drop in the participation tax, and according to our model, this should increase labor force participation relative to a situation without an EITC.

Second, the TRA86 expansion of the EITC amounts to a further decrease in the size of the participation tax. In the model, this leads to an increase in the employment rate for the affected group, single women with a qualifying child.

It may be noted that the actual reform affects marginal tax rates differently in different income regions as well as the participation tax rate, which complicates matters. But given our simple model without intensive margin optimization, we would expect labor force participation of the targeted group to increase.

**(3D)** Eissa and Liebman (1996) estimate the impact of the EITC expansion using a difference-in-differences (DiD) estimation framework. In order for this to work, they need two similar groups, where one is affected by the reform and one is not. The affected group (treatment group) is single women with a qualifying child, and the unaffected group (control group) is single women without children.

Estimating the impact of the reform using the time-difference in labor force participation of the treatment group alone is likely biased by shocks to labor force participation coinciding with the reform. Even a secular trend in labor force participation would hamper identification. By subtracting the time-difference for the control group, the authors hope to cancel out any of such shocks/trends. This explains why the identifying assumptions are the following: (*i*) time-effects are identical across groups, and (*ii*) there is no change in composition within groups. These two assumptions amounts to assuming that the time trend in the treatment and control groups would be identical in the absense of treatment, the so-called "common trend" assumption.

If (*i*) does not hold, we cannot hope to get rid of time trends by subtracting the time difference for the control group from that of the treatment group. If (*ii*) does not hold, the same problem arises. A compositional change could for example be a change in the share of unionized workers in the treatment group coinciding with the reform, which could affect their labor force participation and not the control group's participation. As long as compositional changes occur in observable characteristics, we can control for it in a regression framework, but if occurs in unobservable characteristics, we have a problem.

In Table II we see the simple DiD estimation table without controlling for observable characteristics of the two groups (as can be done in a regression framework). Denote by $E_t^i$ the labor force participation of group $i$ (= **T**reatment or **C**ontrol) at time $t$ (**pre** or **post** reform). Column 1 shows pre-reform labor force participation of the treatment group (top row) and the control group (bottom row). Column 3 shows the time differences of each group

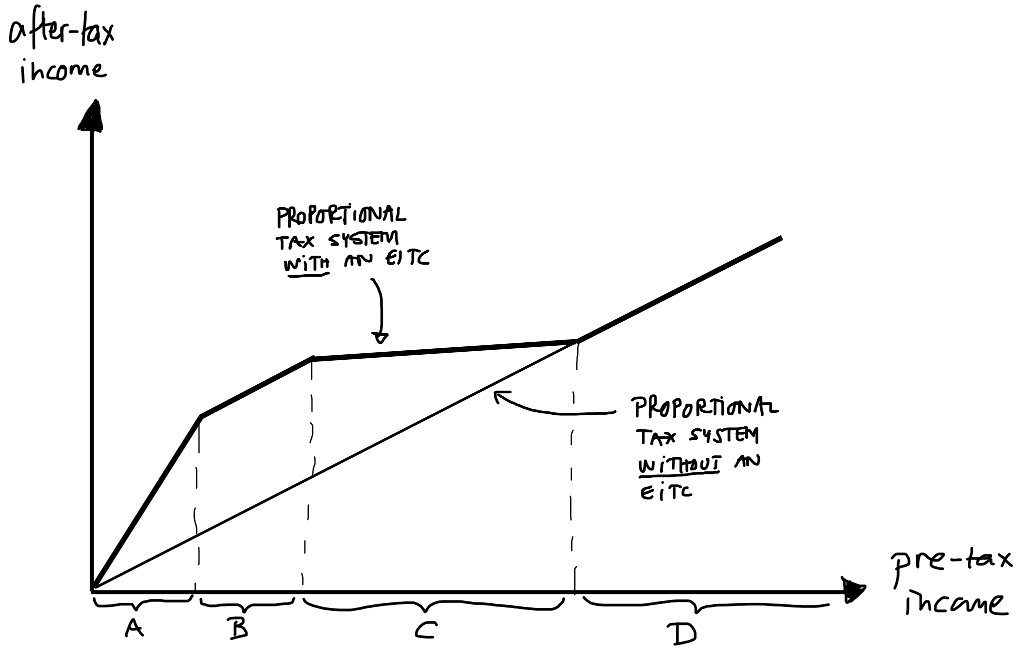
$$\begin{aligned} E^T_{\text{post}} - E^T_{\text{pre}} &= 0.753 - 0.729 = 0.024 \\ E^C_{\text{post}} - E^C_{\text{pre}} &= 0.952 - 0.952 = 0.000. \end{aligned}$$

Column 4 shows the DiD estimate as

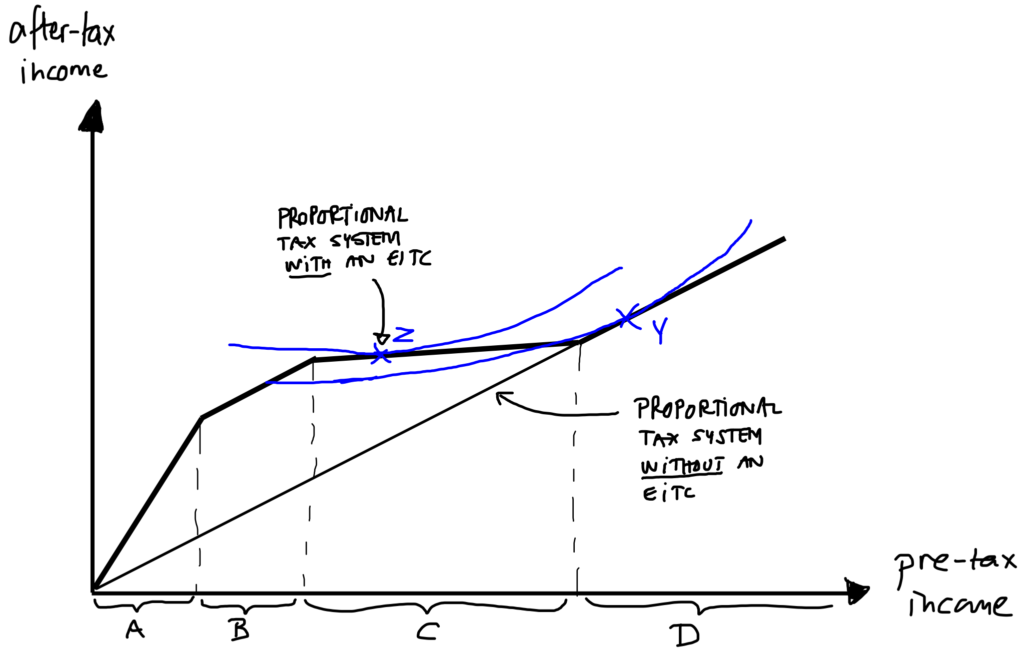
$$\left(E^T_{\text{post}} - E^T_{\text{pre}}\right) - \left(E^C_{\text{post}} - E^C_{\text{pre}}\right) = 0.024 - 0.000 = 0.024.$$

Their estimate of the impact of the reform has the expected sign as the expansion seems to have led to an increase in labor force participation of the target group. However, we may be concerned about the comparability of the two groups given that their pre-reform level of labor force participation are so different (0.729 vs. 0.952).

**(3E)** The graph below shows the relationship between pre- and after-tax income in a proportional tax system without an EITC and the same tax system with an EITC of the US-type in place. In either case, the slope of the curve equals $1 - m_j$, where $m_j$ is the effective marginal tax rate in income region $j = A, B, C, D$. Pre-tax income is split into four income regions: a phase-in (A), a plateau (B), a phase-out (C), and a region where the marginal tax rate is unchanged.




In region A the marginal tax decreases following the EITC introduction. In that case the substitution effect is working towards more hours worked, whereas the income effect is working in the opposite direction. Theoretically, the net effect is ambiguous, but empirically we tend to find that substitution effects dominate income effects. In region B, the marginal tax rate is unchanged but the average tax is lower. In that case, there is no substitution effect and the income effect pushes towards fewer hours worked. In region C, the marginal tax rate has increased and the average tax rate has decreased. This implies that both income and substitution effects push towards fewer hours worked. Finally, in region D, the marginal and average tax rates are unaffected and no behavioral response should come about. However, it is possible that some taxpayers initially in region D will get a discrete increase in utility by moving down to region C (from point Y to Z in the graph below).




The table below summarizes the effects discussed here.

| | Income region | | | |
|---|---|---|---|---|
| | A | B | C | D |
| Marginal tax rate | – | 0 | + | 0 |
| Average tax rate | – | – | – | 0 |
| | Effect on hours worked | | | |
| Substitution effect | + | 0 | – | 0 |
| Income effect | – | – | – | 0 |
| Total effect | ? (+) | – | – | (–) |

As the discussion revealed, there are effects working in opposite directions. Therefore it is not possible from theory alone to say whether intensive margin responses will increase or decrease aggregate number of working hours. It depends among other things on the relative size of income and substitution elasticities and on the income distribution (cf. the discussion of the $\alpha$-parameter in question 2C).