Exam - Tax Policy - Fall 2016 - Answers

Part 1: Firm taxation

(1A) **Q:** The value of the firm is the net present value of the after-tax cash-flows accruing to the shareholders. The value of the firm can generally be expressed as:

$$V = (1-t_d)D - E + \frac{(1-t_d)[(1-t_c)f(I) + X - D] + E}{1+r}$$

where the first two terms reflect the cash-flow at the beginning of period 0 and the last term reflects the cash-flow Under the assumption that there is an interior solution for D, we have:

$$\frac{\partial V}{\partial D} = \frac{(1-t_d)r - (1-t_c)(1-t_d)f'(I)}{1+r} = 0$$

This implies that the optimal investment level is determined by:

$$f'(I^*) = \frac{r}{(1-t_c)}$$

Hence, the investment level of the firm is unaffected by the dividend tax. Under the assumption that E=0, it follows trivially that optimal dividend payments are given by:

$$D^* = X - I^*$$

Since X is exogenous and I^* is unaffected by the dividend tax, also the firm's dividend payments are unaffected by the dividend tax.

Intuitively, the marginal source of finance is retained earnings. The relevant choice for the firm is therefore whether to distribute dividends now or invest and distribute later. This choice is unaffected by the dividend tax rate, because this tax will apply regardless of whether dividends are paid out now or later. Hence, dividend taxes have no bearing on firms' dividend and investment policies.

Q: The maximized value of the firm is given by

$$V = (1-t_d)\left\{D^* + \frac{[(1-t_c)f(I^*) + X - D^*]}{1+r}\right\}$$

Since V is proportional to $(1-t_d)$ and all variables inside the curly brackets are independent of t_d, a percentage increase in $(1-t_d)$ causes a percentage increase in V. Intuitively, the cash-flows received by shareholders will correspond to the sum of retained and future earnings, all of which will be subject to dividend taxation.

(1B) **Q:** In the U.S., corporations can choose between two fundamentally different tax treatments. They can elect that current profits are taxed at the corporate level at the rate t_c and that distributed profits are taxed at the shareholder level at the rate t_d. Corporations choosing this tax treatment are

known as C-corporations. Alternatively, they can elect that current profits are taxed at the shareholder level at the personal income tax rate t_p and that no corporate or dividend taxes apply. Corporations choosing this tax treatment are known as S-corporations. The choice of tax status does not influence the corporation's legal status: both S- and C-corporations have limited liability for instance. However, there are certain restrictions on the choice of S-corporation status, which makes it unattractive for very large companies: the number of shareholders cannot exceed 100 and all shareholders need to be physical persons residing in the U.S.

The empirical design of Yagan (2015) exploits that C-corporations are subject to the dividend tax whereas S-corporations are not. This allows for a difference-in-differences type of estimator where S-corporations work as a "control group" for the C-corporations that are "treated" with a large dividend tax reduction.

Q: Empirically, C-corporations are overrepresented among large firms whereas S-corporations are overrepresented among small firms. This raises the concern that unobserved shocks to investment policies that correlate with size will cause divergence between the "treatment group" and "control group". This concern is addressed in two ways. First, the estimation sample is restricted to firms of intermediate sizes. In this range, the size (and industry) distribution is shown to be reasonably balanced across C-corporations and S-corporations. Second, the regression weights are adjusted such that S-corporations have the precise same weight as C-corporations within each narrow size-industry bin. This ensures that size-industry specific shocks to investment decisions do not affect the difference-in-differences estimator.

(1C) **Q:** The finding of a zero effect of dividend taxes on investment is clearly consistent with the new view under which the dividend tax does not the firm's cost of capital. **Q:** Strictly speaking, the old view only states that dividend taxes affect the firm's cost of capital: marginal investment is financed with new equity implying that corporate investment is subject to a double tax wedge - corporate tax and dividend tax - relative to the alternative investment in bonds. One possible interpretation of the empirical results is therefore that the old view is flawed. An alternative interpretation is that the dividend tax, consistent with the old view, does affect the firm's cost of capital, but that investment choices are insensitive to this cost of capital.

Part 2: Income taxation

(2A) **Q:** The *mechanical revenue effect*: Holding behavior constant, the tax revenue increases by:

$$\Delta M = \Delta\tau \Delta z[1 - H(z)]$$

The mass of people with income above z, that is $1-H(z)$, each pay an increaed tax bill of $\Delta\tau\Delta z$ because the marginal tax is increased by $\Delta\tau$ over an income interval of length Δz.

The *behavioral revenue effect*: Behavioral responses to the tax reduce tax payments by:

$$\Delta B = -h(z)\frac{\partial z}{dT'(z)}\Delta\tau T'(z)$$

The mass of people who become subject to a higher marginal tax, that is $h(z)$, each reduce their taxable income by $-(\partial z/dT'(z)) \cdot \Delta\tau$ and each dollar of reduction in taxable income generates a revenue loss of $T'(z)$. This equation can be rewritten in the following way using the elasticity of taxable income $e(z)$:

$$\Delta B = h(z)e(z)z\frac{T'(z)}{1-T'(z)}\Delta\tau\Delta z$$

The *social welfare cost*: The increased tax payment is associated with utility costs with a social value of:

$$\Delta W = \Delta\tau\Delta z[1-H(z)]G(z)$$

This is simply the increased tax payment (absent behavioral responses) for persons with income exceeding z multiplied by the average social welfare weight for persons with income exceeding z, that is $G(z)$. As usual, behavioral responses to a small tax change have no first-order effect on individual utilities.

(2B) **Q:** In an optimum, it must hold that the social gain of a small tax increase equals the social cost of a small tax increase:

$$\Delta M = \Delta W + \Delta B$$

Insert the expressions derived under the previous question:

$$\Delta\tau\Delta z[1-H(z)] = \Delta\tau\Delta z[1-H(z)]G(z) + h(z)e(z)z\frac{T'(z)}{1-T'(z)}\Delta\tau\Delta z$$

Reduce and rearrange to obtain:

$$\frac{T'(z)}{1-T'(z)} = \frac{1-G(z)}{e(z)} \cdot \frac{1-H(z)}{zh(z)}$$

Q: A larger value of $e(z)$ implies that those who face increased marginal tax rates respond more strongly to taxation. This in turn implies a larger marginal efficiency loss and a lower optimal marginal tax rate. A larger value of $1-G(z)$ implies lower welfare weight on those facing increased tax bills. This in turn implies a higher optimal marginal tax rate. A larger value of $1-H(z)$ implies larger mass of people who pay higher tax bills after reform but face no change in marginal tax rates and hence do not change behavior. This in turn implies a larger mechanical revenue gain relative to the behavioral revenue loss and therefore a higher optimal marginal tax rate. A larger value of $h(z)$ implies a larger mass of people who face an increase in the marginal tax rate. This in turn implies a larger marginal efficiency loss and a lower optimal marginal tax rate.

(2C) **Q:** At a constant marginal tax rate, tax payers optimally choose their labor supply such that the marginal rate of substitution between consumption and income equals $(1-t)$ where t is the marginal tax rate. Given heterogeneity across workers, a smooth distribution of skills and preferences translates into a smooth distribution of labor incomes. When a kink is introduced in the tax schedule by increasing the marginal tax rate to $t+dt$ at some income level z^*, there will be bunching at the kink because those choosing the income level z^* are no longer just individuals with $MRS(z^*) = (1-t)$ but all individuals with $MRS(z^*)$ between $(1-t-dt)$ and $(1-t)$.

Saez (2010) shows that the amount of excess mass at the kink point is determined by three factors:

(i) the size of the kink: the larger the jump in the marginal tax rate, the larger the range of individuals who will find it optimal to choose the income level z^*.

(ii) the counter-factual density around z^*: the more individuals who have incomes in a neighborhood around z^* without the kink, the more will choose the income level z^* in the presence of the kink

(iii) the elasticity of taxable income: the larger the elasticity of taxable income, the larger the range of individuals who will find it optimal to choose the income level z^*

While the size of the kink (i) can be measured directly, the excess mass at z^* as well as the counterfactual density (ii) can be easily approximated. With knowledge of three out of four variables in an equation, it is possible to back out the unknown fourth variable, the elasticity of taxable income.

Part 3: Shorter questions

(3A) **Q:** In the equilibrium, it must hold that labor demand at the effective labor cost of $w + t$ equals labor supply at the effective labor remueration of $w + \alpha t$:

$$D(w + t) = S(w + \alpha t)$$

Differentiate with respect to t to obtain:

$$\begin{aligned} D'(\cdot)(\frac{dw}{dt} + 1) &= S'(\cdot)(\frac{dw}{dt} + \alpha) \\ &\Leftrightarrow \\ \frac{dw}{dt} &= \frac{D'(\cdot) - \alpha S'(\cdot)}{S'(\cdot) - D'(\cdot)} \end{aligned}$$

For a small change in t and starting from $t = 0$, this can be rearranged to yield:

$$\frac{dw}{dt} = -1 + (1 - \alpha)\frac{\varepsilon^S}{\varepsilon^S - \varepsilon^D}$$

When the mandated benefit is valued at its full cost ($\alpha = 1$), the entire incidence is on the workers; the wage rate falls by the full amount of the cost. When the mandated benefit is not valued at all ($\alpha = 0$), the mandated benefit is equivalent to a tax on employers and the standard incidence formula applies. In intermediate cases, the part of the cost that is valued by the workers is born entirely by the workers and the part that is not valued by the workers is split between workers and employees in proportions given by the standard incidence formula.

Q: To the extent that the mandated benefit is valued by workers (and thus shifted into the wage), it does not affect labor cost from the perspective of employers nor the remuneration of labor from the perspective of workers, hence the equilibrium employment level is unchanged. This implies that mandated benefits are a more efficient tool for providing goods and services to workers than universal provision through taxation. The key is the direct link between work and benefits, which makes workers internalize the mandated benefit in the labor supply decision.

(3B) **Q:** A firm reporting zero profits is a crude indicator for aggressive tax avoidance. The figure shows that firms' propensity to report zero profits correlate strongly with their incentives to shift profits to foreign affiliates: when the tax rate of a firm's foreign affiliates is lower than its own tax rate, it is much more likely to report zero profits than when the tax rate of its foreign affiliates is higher than its own tax rate.

Q: First, the simple comparison of corporations with low-tax and high-tax parents only uses part of the variation in the tax incentive to shift profits; it ignores that the tax saving from profit shifting is not the same for all corporations with low-tax parents, but proportional to the tax differential. Second, it is effectively making comparisons across corporations operating in different countries, comparing, for instance, a corporation in Poland with a low-tax parent to a corporation in Georgia with a high-tax parent. This is problematic if there are cross-country differences in the propensity to report zero profits for other reasons than profit shifting. Both of these issues are addressed in a regression framework that explains the propensity to report zero profit with the tax saving associated with a dollar of profits shifted and country fixed effects.

(3C) **Q:** Sales taxes in the U.S. are generally not salient: only the pre-tax price is reported on the price tag and the tax is added at the register after the consumption decision has been made. Chetty, Looney and Croft (2009) study how tax salience affects consumer choices in a controlled experiment: they choose a number of "treatment" products in "treatment stores" where information about the after-tax price is included on the price tag. The price tags of "control" products are untouched just like all price tags in "control" stores are untouched.

Q: This experimental set-up allows for a difference-in-difference strategy where the change in the sales of treated products in treatment stores (**-1.3**) is compared to the change in the sales of control products in treatment stores (**0.84**). This yields a DiD estimate (**-2.14**) which is robust to store-specific shocks, but not product-specific shocks. It also allows for a difference-in-difference-in-difference strategy where the general product-specific shock is estimated in the control stores (**0.06**) and subtracted from the DiD treatment effect. This yields an DiDiD estimate of the treatment effect (**-2.2**), which is robust to both product- and store specific shocks, but not to product-store specific shocks.