Answers to exam in Tax Policy, January 2014.

Part 1: Firm taxation

(1A) The value of the firm is the net present value of the cash-flow to the shareholders. The first-period cash-flow is the net-of-tax dividend payment $(1-t_d)D$ minus the injection of new equity E. The second-period cash-flow is the corporate profits net of corporate taxes and dividend taxes $(1-t_d)(1-t_c)f(I)$, retained earnings from previous periods net of dividend taxes $(1-t_d)(X-D)$ as well the reimbursement of equity capital E, which is tax free.

Differentiate V with respect to D and E to obtain:

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

$$\frac{\partial V}{\partial E} = \frac{(1-t_c)(1-t_d)f'(I) - r}{1+r} \tag{2}$$

It follows that:

$$\frac{\partial V}{\partial D} + \frac{\partial V}{\partial E} = -\frac{rt_d}{1+r}$$

Hence, if $E > 0$ and $D > 0$, it is profitable to reduce dividend payment and equity issues by the same amount to keep the investment level I constant. This reduces dividend tax payments in period 1 by t_d and increases them by the same amount in period 2, which yields a "liquidity gain" of rt_d in period 2, or equivalently a present value gain of $rt_d/(1+r)$.

(1B) Assume that the firm is sufficiently **cash-rich** so as to ensure that $(1-t_c)f'(X) < r$. We have shown above that either $E = 0$ or $D = 0$. First assume that $D = 0$ and evaluate (2) at this point:

$$\left(\frac{\partial V}{\partial E}|D=0\right) = \frac{(1-t_c)(1-t_d)f'(X+E) - r}{1+r} \tag{3}$$

This must be negative for any $E \geq 0$ given the initial assumption that $(1-t_c)f'(X) < r$ and the concavity of $f()$, hence there is no interior solution for E. Then assume that $E = 0$ and evaluate (1) at this point:

$$\left(\frac{\partial V}{\partial D}|E=0\right) = \frac{(1-t_d)r - (1-t_c)(1-t_d)f'(X-D)}{1+r} \tag{4}$$

It follows that there is an interior solution $E = 0$ and $D > 0$ where D satisfies:

$$(1-t_c)f'(X-D) = r \tag{5}$$

Under the new view, the firm finances the marginal investment with retained earnings so the relevant choice on the margin is between dividend payments and investment. It follows from (5) that the dividend tax does not affect this choice and hence has no bearing on dividend payments and investment.. This is because retained earnings are hit by the dividend tax regardless of whether they are distributed now or reinvested and distributed later (trapped cash) and, hence, the dividend tax does not affect the choice between these two options. The prediction is contradicted by Chetty and Saez (2005) who find that the

US dividend tax cut in 2003 had an immediate and significant positive effect on dividend payments. The good answer may elaborate on the method and findings by Chetty and Saez (2005).

(1C) Agency problems may occur in large firms with separation of ownership and management. The management ("agent") makes the day-to-day decisions but may not aim to maximize the value of the firm, which would be in the interest of the owners ("principal"). This is captured in the model by the assumption that the allocation of resources is made to maximize the pay-off to the manager V^M. This pay-off has two components. First, it is assumed that the manager holds a fraction α of the shares in the firm, hence the first term of V^M represents the value of the shares owned by the management. Second, it is assumed that the manager is able to divert J of the firm's resources for unproductive investment, which creates benefits $g(J)$ for the manager himself but no benefits for the other owners. Monitoring by the owners captured by γ, however, reduces the benefits the manager derives from his pet project. The second term thus represents the present value of the benefit derived from J given the monitoring level.

Rewrite the equation as

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

where $\omega \equiv \alpha(1-t_d)(1+\gamma)$. The first-order conditions for optimal levels of I and D are:

$$\omega(1-t_c)f'(I) = g'(X - I - D) \tag{6}$$

$$\omega r \leq g'(X - I - D) \tag{7}$$

If α and γ are sufficiently large, (7) holds with equality, hence

$$\omega(1-t_c)f'(I) = \omega r \Leftrightarrow (1-t_c)f'(I) = r$$

so that $I = I^*$. The equality of (7) implies that the optimum is interior with $D > 0$.

The tax on dividends does not distort the investment level, which is at $I = I^*$. Intuitively, the manager puts sufficient weight on firm value relative to his pet project to pay positive dividends in the optimum. Hence, he faces a trade-off on the margin between dividend payments and productive investment and this behavioral margin is not affected by the agency problems. Hence, funds are allocated to productive investment until the marginal product $(1-t_c)f'(I)$ is the same as the return to the alternative investment r regardless of the dividend tax rate. The manager also faces a trade-off on the margin between dividend payments and unproductive investment. This trade-off is affected by the dividend tax because the the tax reduces the value of dividend payments relative to the private return to unproductive investment. In sum, a dividend tax cut should increase the level of dividend payments immediately by making such payments more attractive relative to unproductive investment by self-interested managers. This is consistent with the findings by Chetty and Saez (2005) described above.

Part 2: Commodity taxation

(2A) The langrangian to the government problem writes

$$\mathcal{L}_G = V(q, Z) + \lambda[\sum_j t_j X_j(q, Z) - T]$$

The first-order condition for q_k equals

$$\frac{\partial \mathcal{L}_G}{\partial q_k} = \frac{\partial V}{\partial q_k} + \lambda[X_k + \sum_j t_j \partial X_j / \partial q_k] = 0$$

Use $\alpha = \partial V / \partial Z$ and Roy's identity to rewrite as:

$$(\lambda - \alpha) X_k + \lambda \sum_j t_j \partial X_j / \partial q_k = 0$$

Now insert the Slutsky equation to obtain:

$$(\lambda - \alpha) X_k + \lambda \sum_j t_j (S_{jk} - X_k \partial X_j / \partial Z) = 0$$

Insert $\mu \equiv \alpha + \lambda(\sum_j t_j \partial X_j / \partial Z)$ to obtain:

$$(\lambda - \mu) X_k + \lambda \sum_j t_j S_{jk} = 0$$

Rearrange as:

$$\frac{\lambda - \mu}{\lambda} = -\frac{\sum_j t_j S_{jk}}{X_k}$$

The numerator on the right-hand side is the revenue effect of the compensated behavioral responses to a small increase in the tax on good k. This can be interpreted as the marginal excess burden of the tax increase. The denominator is the mechanical revenue effect of a small increase in the tax on good k. The right hand side thus expresses the share of the potential revenue gain from a small tax increase that is lost. The result shows that this share should be equalized across all instruments. This ensures that the total excess burden is minimized given the revenue constraint. The optimal commodity tax system thus maximizes economic efficiency.

(2B) When the economy comprises more individuals with different levels of unearned income and the government has a preference for redistribution, the optimal commodity tax system redistributes from high-income to low-income individuals by raising the tax on goods consumed disproprotionately by high-income individuals and lowering the tax on goods consumed disproprotionately by low-income individuals. The optimal commodity tax system thus sacrifices some efficiency in order to enhance equity.

An individual who has self-control problems in the way modeled by O'Donoghue and Rabin (2003) overconsumes goods that generate immediate utility and disutility in the future ("sin goods") according to his own long-term preferences. The optimal commodity tax system corrects this inefficiency by raising the tax on sin goods and lowering the tax on other goods. This is akin to a Pigouvian tax on goods that generate externalities. Here, the tax corrects for internalities, that is choices, which cause future harm on the individual itself and which is not fully taken into account because of the self-control problems. Recent examples are the recent Danish taxes on fatty and sugary foods.

(2C) Doyle and Samphanthrak (2008) exploit that two U.S. states, Illinois and Indiana, first repealed and later reinstated, their gasoline taxes whereas the neighboring states left the gasoline taxes unchanged. The timing of these tax changes give rise to three natural experiments: (i) the simultaneous repeal of the gasoline tax by Illinois and Indiana; (ii) the reinstatement of the gasoline tax by Indiana; (iii) the reinstatement of the gasoline tax by Illinois. The causal impact on the consumer price is estimated by estimating the price change in the "treatment" state over and above the price change in the "control states" in a short time window around the tax change. The Figure plots for each day the log of the average retail price on gasoline in the treatment states minus the log of the average retail price in the control states. This is in effect the percentage difference in average retail prices between treatment and control states. Moreover, it fits a line through these data points by local linear estimation (may be explained in more detail) while allowing for a jump at the reform date. The estimated treatment effect is the size of the jump at the reform date. This is the change in the price difference between treatment and control states induced by the tax change and thus a difference-in-difference estimator. The estimator relies on the assumption of parallel trends, i.e. that price difference between treatment and control states would have constant (the line flat) in the absence of a tax change. The paper also accounts for changing observables in a regression framework.

The model of optimal commodity taxation developed in (2A) assumes that producer prices are fixed, so that a dollar increase in taxes causes a dollar increase in consumer prices. This assumption implies that the full incidence is on the consumers.

Part 3: Shorter questions

(3A) The optimal income tax balances three effects. Increasing the marginal tax at income level z, (i) mechanically increases government revenue; (ii) mechanically reduces disposable income of those with income above z; (iii) causes behavioral responses that lowers taxable income and reduces the government revenue. The elasticity of taxable income e(z) captures the strength of the behavioral responses for a given individual who becomes subject to a higher marginal tax. The larger the responses, that is the larger e(z), the lower the optimal marginal income tax. The density of the income distribution h(z) captures the mass of people whose behavior is distorted by the marginal tax at income level z. The larger this mass, the more important the negative behavioral revenue effect and the lower the optimal income tax. The term $1 - H(z)$ represents the fraction of the population with incomes above z. When $T'(z)$ is increased, these individuals pay higher taxes but their marginal tax is unchanged. Under the assumption of zero income effects, their behavior is therefore not distorted. Hence, the larger.$1 - H(z)$, the larger the size of the mechanical revenue effect and theregore the higher the optimal marginal tax at income level z. Finally, the term $1 - G(z)$ captures how much government revenue is valued socially relative to disposable income in the hands of those who become subject to a higher tax burden. The larger the term $1 - G(z)$, the more valuable is government revenue and the higher the marginal tax rate.

(3B) In the theoretical model by Chetty, Kroft and Looney (2009), consumer prices are the sum of producer prices and sales taxes, however, sales taxes are less salient - and therefore induce smaller responses by consumers - than producer prices. The latter assumption is motivated by the fact that the sales tax in many U.S. states is not included in the posted price, but is added at the register at purchase. The figure shows the supply and demand as a function of the producer price p. When a tax is imposed, the demand curve shifts to the south-west because consumers are willing to buy a smaller quantity at a given producer price, however, the demand shift is smaller than for an equivalent increase in the producer price. The demand shift implies that there is over-supply, which pushes down the producer price, which, in turn, reduces the over-supply by increasing demand and reducing supply. Clearly, salience is a determinant of tax incidence. When salience is low (high) the shift of the demand curve is small (large) and a relatively small (large) drop in the producer price is required to equilibrate the market. Hence, the less salient the tax, the more of the burden is borne by consumers. Moreover, also the relative size of supply and demand elasticities matter like in the standard model with full salience. For instance, when the supply elasticity is large relative to the demand elasticity, it takes a relatively small drop in the producer price to equilibrate the market given the initial shift in the demand curve and more of the burden is borne by consumers.