Answers to exam in Tax Policy, Summer 2017.

Part 1: Paternalism

(1A) **Q**: The parameter δ expresses standard time discounting: the tendency to attach more weight to utility gains that are closer in time. Under this form of discounting, the consumer is indifferent between one unit of consumption in period t and $1/\delta$ units of consumption in period $t+1$ for any t. This form of discounting implies that consumers make *time consistent* consumption plans because the discount factor between two time periods is constant over time. Both individuals with and without self-control problems should have $\delta > 0$ (implying that an increase in future consumption increases life-time utility) and $\delta < 1$ (implying some degree of impatience).

The parameter β expresses hyperbolic discounting: an additional discounting of future consumption relative to current consumption or, in other words, a "present-bias" in the intertemporal preferences. This form of discounting implies that consumers make *time inconsistent* consumption plans because the discount factor between two time periods changes over time. For instance, when making consumption plans at $t-1$, consumption at $t+1$ is discounted at the rate δ relative to consumption at t. When making consumption plans at t, consumption at $t+1$ is discounted at the rate $\beta\delta$ relative to consumption at t, hence the plan made at $t-1$ is no longer perceived as optimal.

Q: Combine the intertemporal and instantaneous utility functions to obtain:

$$\begin{aligned} U^0 &= \rho \ln(x_t) + \sigma \ln(y_t) + z_t - \gamma \ln(x_{t-1}) + \\ &\quad \beta\delta \left\{\rho \ln(x_{t+1}) + \sigma \ln(y_{t+1}) + z_{t+1} - \gamma \ln(x_t)\right\} + \\ &\quad \beta\delta^2 \left\{\rho \ln(x_{t+2}) + \sigma \ln(y_{t+2}) + z_{t+2} - \gamma \ln(x_{t+1})\right\} + \\ &\quad ... \end{aligned} \tag{1}$$

Construct the Lagrangian $\mathcal{L}$ and differentiate with respect to x_t, y_t and z_t to obtain the following first-order conditions for optimal consumption in period t:

$$\frac{\partial \mathcal{L}}{\partial x_t} = \frac{\rho}{x_t} - \frac{\beta\delta\gamma}{x_t} - \lambda p_x = 0 \tag{2}$$

$$\frac{\partial \mathcal{L}}{\partial y_t} = \frac{\sigma}{y_t} - \lambda p_y = 0 \tag{3}$$

$$\frac{\partial \mathcal{L}}{\partial z_t} = 1 - \lambda = 0 \tag{4}$$

where λ_t denotes the Lagrangian multiplier in period t. It follows from (4) that $\lambda_t = 1$. Insert into (2) and (3) to obtain:

$$x^* = \frac{\rho - \beta\delta\gamma}{p_x}$$

$$y^* = \frac{\sigma}{p_y}$$

Insert x^* and y^* into the budget constraint and rearrange to obtain:

$$z^* = B - (\rho - \beta\delta\gamma + \sigma)$$

(1B) **Q:** At period t, long-run utility is maximized by choosing consumption for periods t and onwards that maximizes utility evaluated in any period before period t, for instance $t-1$. Utility in period $t-1$ is given by:

$$U^{t-1} = u_{t-1} + \beta(\delta u_t + \delta^2 u_{t+1} + \delta^3 u_{t+2} + ...)$$

Maximizing U^{t-1} with respect to consumption in period t and onwards is equivalent to maximing the terms in U^{t-1} that concern period t or later (as consumption choices in period t and onwards cannot affect instantaneous utility in periods before period t). Maximizing those terms is equivalent to maximizing

$$\widetilde{U}^t = u_t + \delta u_{t+1} + \delta^2 u_{t+2} + ... \tag{5}$$

This expression can be interpreted as life-time utility at period t evaluated at the long-run preferences that ignore the consumer's present-bias

Q: The consumer's choices suffer from a present-bias: in any period t he eats more potato chips than he would have if he could have made consumption choices for period t in any period before period t. Optimal paternalism prescribes that governments should influence people's choices in ways that make them better off as judged by themselves. By

maximizing $\widetilde{U}^t$, the government induces the consumer to make consumption choices that make him as well off as possible as evaluated by his own long-term preferences.

(1C) **Q**: Inserting instantaneous utility functions into (5) yields:

$$\begin{aligned} &\rho \ln(x_t) + \sigma \ln(y_t) + z_t - \gamma \ln(x_{t-1}) + \\ &\delta \left\{ \rho \ln(x_{t+1}) + \sigma \ln(y_{t+1}) + z_{t+1} - \gamma \ln(x_t) \right\} + \\ &\delta^2 \left\{ \rho \ln(x_{t+2}) + \sigma \ln(y_{t+2}) + z_{t+2} - \gamma \ln(x_{t+1}) \right\} \\ &... \end{aligned} \tag{6}$$

Construct the Lagrangian $\mathcal{L}$ and differentiate with respect to the tax rate on potato chips in period t to obtain the following first-order condition for optimal taxes in period t:

$$\frac{d\mathcal{L}}{dt_x} = \frac{(\rho - \delta\gamma)}{x_t} \frac{\partial x_t}{\partial t_x} + \frac{\sigma}{y_t} \frac{\partial y_t}{\partial t_x} + \frac{\partial z_t}{\partial t_x} + \mu_t \left\{ t_x \frac{\partial x_t}{\partial t_x} + x_t + t_y \frac{\partial y_t}{\partial t_x} \right\} = 0 \tag{7}$$

where μ_t denotes the Lagrangian multiplier concerning period t.

It follows from the demand functions that:

$$\frac{\partial x^*}{\partial t_x} = -\frac{\rho - \beta\delta\gamma}{(p_x)^2} \tag{8}$$

$$\frac{\partial y^*}{\partial t_x} = 0 \tag{9}$$

$$\frac{\partial z^*}{\partial t_x} = 0 \tag{10}$$

Insert (8)-(10) into (7) to obtain:

$$\frac{(\rho - \delta\gamma)}{\frac{\rho - \beta\delta\gamma}{p_x}} \left(-\frac{\rho - \beta\delta\gamma}{(p_x)^2} \right) + \mu \left\{ t_x \left(-\frac{\rho - \beta\delta\gamma}{(p_x)^2} \right) + \left(\frac{\rho - \beta\delta\gamma}{p_x} \right) \right\} = 0$$

Reduce to obtain:

$$\frac{t_x}{1 + t_x} = \frac{\mu - \frac{\rho - \delta\gamma}{\rho - \beta\delta\gamma}}{\mu}$$

Q: The optimal tax formula ressembles the inverse elasticity rule where the elasticity of demand is one and the marginal social value of private income is one due to the quasi-linear form of the instantaneous utility function. The formula thus differs from the inverse elasticity rule only by having $-(\rho - \delta\gamma)/(\rho - \beta\delta\gamma)$ as the second term in the

numerator rather than -1. In the special case where $\beta = 1$ and the consumer does not suffer from hyperbolic discounting, the term $-(\rho - \delta\gamma)/(\rho - \beta\delta\gamma)$ collapses to -1 and the standard inverse elasticity rule prevails. In cases where $\beta < 1$ and the consumer suffers from hyperbolic discounting, the tax on potato chips is higher, which reduces demand for this good. The optimal tax formula thus corrects for the optimization error made by the consumer. This is akin to a Pigouvian tax correcting for an externality. In this case, the tax corrects for an "internality", that is the harm inflicted by the consumer on himself due to the hyperbolic discounting.

Part 2: Firm taxation

(2A) **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 at the beginning of period 1.

Q: Under the assumption that there is an interior solution for E, we have:

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

This implies that the optimal investment level is determined by:

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

Hence, the user cost of investment is increasing in both the dividend tax and corporate tax suggesting that the investment level of the firm should be decreasing in both tax rates. Intuitively, the marginal source of finance is new equity so the relevant choice for the owners is whether to invest in the firm, in which case profits will be subject to the corporate tax and dividend tax, or to invest in bonds, in which case these taxes do not apply. The owners therefore inject equity into the firm up until the point where the marginal return on investment net of corporate and dividend taxes equals the return on bonds.

(2B) **Q:** The empirical design of Yagan (2015) exploits that U.S. corporations can choose between two fundamentally different tax treatments. If they opt for a status as "C-corporations", current profits are taxed at the corporate level at the rate t_c and distributed profits are taxed at the shareholder level at the rate t_d. If they opt for a status as "S-corporations", current profits are taxed at the shareholder level at the personal income tax rate t_p and no corporate or dividend taxes apply. While large firms tend to be C-corporations and small firms tend to be S-corporations, the two types of firms co-exist at intermediate sizes within all industries.

The tax reform in 2003 lowered the dividend tax rate considerably and therefore lends itself to a difference-in-differences empirical strategy whereby S-corporations serve as a "control group" for the "treatment group" of C-corporations. Intuitively, the actual change in the investment level of S-corporations from the pre-reform to the post-reform period is assumed to be the counterfactural change in the investment level of C-corporations absent the reform. Hence, the *differential* change in the investment level of C-corporations is the difference-in-difference estimate of the effect of the tax reform. The strategy is implemented in a regression framework where the difference-in-difference estimator is captured by the interaction between an indicator for C-corporations and an indicator for post-reform periods.

Column 2 suggests that the decrease in the dividend tax rate reduced the ratio of investment to lagged capital by around 0.0002. This treatment effect should be compared to an average investment ratio of 0.2428 across the treated firms before the reform; hence the treatment effect is less than 0.1% evaluated at this sample mean. Combining the size of the change in the tax rate with the estimated treatment effects provides an implied elasticity of the investment ratio with respect to the net-of-dividend-tax of 0.00. Importantly, the standard errors of the estimate are relatively small. At the 95% confidence level, it can be rejected that the true elasticity exceeds 0.08.

Q: The theoretical model finds that dividend taxes increase the firm's cost of capital. The empirical results suggest that dividend taxes have no bearing on firms' investment choices. There are two possible interpretations. Either the old view is flawed, for instance because the marginal source of finance is retained earnings, as stipulated by the new view,

such that the cost of capital is independent of the dividend tax rate. Or, the old view is accurate in assuming that the dividend tax rate does affect the firm's cost of capital, but investment choices are insensitive to this cost of capital.

(2C) **Q:** Chetty and Saez (2005) estimate the effect of the reduction of the U.S. dividend tax rate in 2003 on dividend payouts. The main analysis does not employ a control group and therefore bases inference on a simple difference-estimator: the change in dividend pay-outs from the pre-reform to the post-reform period This identification strategy is vulnerable to unobserved shocks to dividend pay-outs that coincide with the reform.

The main results are suggestive that the reform had a large and immediate effect on the extensive margin of dividend payouts: the first figure shows that dividend initiations soared immediately after the reform whereas dividend terminations dropped. Both effects suggest that the reform increased the number of firms paying dividends. Moreover, there was a considerable effect on the intensive of dividend payouts: the second figure shows that the number of firms raising their dividend payments increased immediately after the reform whereas the number of firms lowering their dividend payments decreased. These results are robust to controls for observable shocks.

Q: The theoretical model implies that dividend taxes have no bearing on dividend payments in period 0. This follows directly from the assumption that firms need to raise new equity to finance investment and pay no dividends in this period. The model also implies that lower dividend taxes raise dividend pay-outs in period 1 by increasing new equity in period 0 and thus increasing the profits to be distributed when the firm shuts down. These theoretical results are suggestive that dividend taxes have no short-run but may have a negative long-run effect on dividend pay-outs, which is inconsistent with the immediate payout response identified by Chetty and Saez (2005).

Part 3

(3A) **Q:** The key assumption underlying the Ramsey model is that governments are unable to tax leisure. The Corlett-Hague rule states that commodities, which are more complementary to leisure, i.e. where the compensated cross-elasticity with respect to

the wage rate is lower, should be taxed at a higher rate. Intuitively, the government attempts to tax leisure indirectly by imposing higher rates on commodities that are consumed together with leisure

Q: The results can rationalize the reduced rates on passenger transport, restaurant meals and repairs of private dwellings observed in many countries because these are goods and services that are not complementary but rather substitutable with leisure. Indeed, individuals in work need to transport themselves to their place of work (thereby purchasing more transport services) and additional hours of work may induce individuals to substitute home-production of meals and repairs with professional services provided by restaurants and craftsmen.

(3B) **Q**: In terms of methodology, Johannesen, Tørsløv and Wier (2017) differs from previous papers in two ways. First, they consider not only the intensive margin of profit reporting but also the extensive margin: whether firms report any profits at all. This is important because the reporting of zero profits in high-tax countries is the most salient outcome of aggressive tax planning and because such zero-profit outcomes are effectively omitted from the analysis when the outcome variable is the log of profits as in most of the existing literature. Second, they identify profit shifting exclusively from variation in foreign tax rates by always including dummies that fully absorb the effect of domestic tax rates and other domestic policies (country dummies in cross-sectional regressions and country-time dummies in panel regressions). This implies that profit shifting is effectively identified from comparisons of corporations in the same country whose incentives for profit shifting differs due to the location of their foreign affiliiates.

Q: The main result of the paper is that reported profits are more sensitive to tax incentives for profit shifting the lower the level of economic and institutional development. The table shows the results for the cross-sectional specification where the outcome is an indicator for reporting zero profits. The results shows that corporations in low- and middle income countries are around 3 percentage points more likely to report zero profits when the tax rate facing their parent is 10 percentage points lower whereas the corresponding effect for corporations in high-income countries is only 1.5 percentage points. Similar results obtain when the average tax rate across foreign affiliates is used to measure

profit shifting opportunites and when the level of economic and institutional development is treated as a continuous variable.