Answers

Re-examination in Public Finance - Summer 2017

3-hour closed book exam

By
Jakob Egholt Søgaard

Part 1: Questions on various topics

(1A) *In a market for a good, where supply is perfectly elastic, a higher demand elasticity implies that the incidence of a tax on the consumer is lower.*

False. The share of the burden of a tax (the incidence) falls most heavily on the side of the market with the relative lowest elasticity. Mathematically we have the share of the incidence on the buyers I_B can be approximated by

$$I_B = \frac{\varepsilon_S}{\varepsilon_S + \varepsilon_B} = \frac{1}{1 + \frac{\varepsilon_B}{\varepsilon_S}},$$

where ε_S is the supply elasticity and ε_D is the demand elasticity. For any finite demand elasticity this expression goes to 1 when the supply elasticity goes to infinity (perfect elastic supply). Hence a higher (finite) demand elasitcity does not affect the incidence on the buyers in this case.

The answer might also be assisted by a graphical illustration.

(1B) *Starting from a situation with no initial taxation, the deadweight loss of introducing a tax is larger under perfect competition than when the market is supplied by a monopoly.*

False. In a competetive market without initial taxation, buyers buy goods until the marginal benefit equals the price and sellers sell their goods until the marginal costs equal the price. There is therefore no (first order) loss in welfare, when a marginal tax increase reduces the number of goods sold marginally. Matematically we see this from the fact that the marginal deadweight loss is linear in t ($MEB \approx t \cdot \varepsilon \cdot x_0$).

In constrast, when the market is supplied by a monopoly, the monopoly sets the price with a markup above the marginal costs. Hence the monopoly earns a profit on the marginal good sold, and this profit (also called the producer surpluss) is lost, when a marginal tax increase that reduces the number of goods sold marginally. Therefore the deadweight is larger in this situation.

The answer can benefit from a graphical illustration of the two situations.

(1C) *In the paper "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark" by Kleven, Knudsen, Kreiner, Pedersen and Saez (published in Econometrica in 2011), the empirical results primarily support that tax evasion in Denmark is low because the Danes are unable to evade.*

True. The empirical results in Kleven et al. are based on a large audit experiment conducted together with the Danish tax agency (SKAT), where a (stratified) random sample of taxpayers are audited thoroughly and homogeneously. The results most relevant for the question are: The overall tax gap is reasonably low (2-3%). On average, the tax agency knows a very high share of net-income of a taxpayer (95%) from third-party information. Underreporting of self-reported income is very high (40%), while underreporting of third-party reported income is very low (0.3%). These results indicate that third-party information is very effective in reducing tax evasion, and making it very difficult to evade large amounts.

The authors find that the differences in tax evasion across individuals to a relative large extend can be explained by their share of self-reported income. In constrast, demographic variables such as gender or being member of the church only explain the differences in tax evade to a much small extend. Hence the results support that it is large amount of third-party information available to the tax agency that prevent the Danes from evading, rather than Danes being more honest than others.

Part 2: Social Insurance: Adverse Selection

Consider an economy where individuals face a risk of becoming unemployed. If they become unemployed they incur a loss of income $d = 1$ assumed to be the same for all individuals. The risk of becoming unemployed θ is exogenous and heterogenous across individuals. Assume that θ is uniformly distributed between $[0, 1]$ in the population. The individuals' willingness to pay *for an insurance that fully compensates them in the case of unemployment is given by:*

$$w(\theta) = (1+\alpha)\,\theta, \tag{1}$$

where α is a measure of risk aversion.

(2A) *Give the intuition for why the individuals' willingness to pay for insurance is consistent with the individuals being risk averse. Describe the first best allocation of insurance. What share of the population should be covered by insurance?*

A potential income loss of $d = 1$ that incur with probability θ implies an expected income loss of $\theta \cdot d = \theta$. A risk neutral person will value insurance at the expected benefits equal the expected loss they would have without insurance ($w(\theta) = \theta$). A risk averse person will value the insurance higher than the expected benefits consistent with $w(\theta) = (1+\alpha)\,\theta$, where $\alpha > 0$.

In a situation with risk averse individuals, the first best allocation of insurance is full coverage (100 percent of the population), because - for very individual with $\theta > 0$ - the willingness to pay for the insurance is larger than the expected costs.

In a private insurance market with the market price (π) only individuals with a willingness to pay above π buy insurance. Assuming that the market is characterized by perfect competition the market equilibrium price (π) is equal to the expected costs of providing insurance, i.e:

$$\pi = E\left[\theta \cdot d \,|w(\theta) > \pi\right] = E\left[\theta\,|(1+\alpha)\,\theta > \pi\right] \tag{2}$$

(2B) *Show that the market equilibrium price equals $\pi^* = \frac{1+\alpha}{1+2\alpha}$. Compare the share of individuals who buy insurance at this price with the first best allocation in (2A).*

Given a market price of π the marginal individual just willing to buy insurance has a risk level $(1+\alpha)\,\hat{\theta} \Leftrightarrow \hat{\theta} = \frac{\pi}{1+\alpha}$. Hence all individuals with $\theta > \hat{\theta} = \frac{\pi}{1+\alpha}$ will buy insurance. This implies that the average risk (given a uniform distribution) is $E\left[\theta\,|(1+\alpha)\,\theta > \pi\right] = \frac{\frac{\pi}{1+\alpha}+1}{2}$ (the midpoint between $\frac{\pi}{1+\alpha}$ and 1).

This gives an equilibrium price of

$$\pi^* = E\left[\theta\,|(1+\alpha)\,\theta > \pi^*\right] \Leftrightarrow \pi^* = \frac{\frac{\pi^*}{1+\alpha}+1}{2} \Leftrightarrow \pi^* = \frac{1+\alpha}{1+2\alpha}.$$

With this price only individuals with a willingness to pay above $(1+\alpha)\,\theta > \frac{1+\alpha}{1+2\alpha}$ will by insurance and hence a fraction $\frac{1}{1+2\alpha}$ will be left uninsured contrary to the first best allocation. I.e. only high risk individuals will buy the insurance (adverse selection).

The answer may benefit from a illustration of the share of individuals who buy insurance.

(2C) *Discuss how government intervention could achieve the first best allocation in (2A). Would this intervention be a Pareto improvement if the government financed the insurance scheme with a lump sum tax of* $1/2$ *per individual?*

Given that the first best allocation of insurance covers the entire population, the government can implement this by public provision of the insurance with the costs covered by taxation. If the government finance the provision by a lump sum tax of 1/2 per individual (the expected costs of the insurance when the entire population is covered) most of the population is made better off, but the low risk individuals are made worse off, because the tax is higher than their valuation of the insurance.

Hence this government intervention achieves the first best allocantion (with is Pareto effecient), but the intervention is not a Pareto improvement, because it also redistributes from low to high risk individuals.

Part 3: Social Insurance: Moral Hazard

Consider an unemployed individual, who has to decide how hard to search for a new job. If the individual chooses a search level of s, he finds a job with probability $p(s) = s$, however searching for a new job has the disutility cost of $v(s)$ with $v'(s) > 0$ and $v''(s) > 0$. Once employed, the individual earns an income of y and pays taxes τ. If the individual remains unemployed, he receives the benefits b. The individual's expected utility is given by:

$$U = s \cdot u(y - \tau) + (1 - s) \cdot u(b) - v(s), \tag{3}$$

where $u(\cdot)$ is the utility of consumption with $u'(\cdot) < 0$ and $u''(\cdot) < 0$. The government's budget constraint is given by $s \cdot \tau = (1 - s)b$.

(3A) *Show that the first best insurance scheme (where the government can control s directly) implies that individuals have full insurance ($y - \tau = b$).*

To show this we need to maximize (3) wrt. s, τ and b subject to the government's budget constraint.

Using the government's budget constraint we can first elimate τ in the expected utility function by rewriting and inserting:

$$s \cdot \tau = (1-s)b \Leftrightarrow \tau = \frac{1-s}{s}b \Rightarrow U = s \cdot u\left(y - \frac{1-s}{s}b\right) + (1-s) \cdot u(b) - v(s).$$

Maximizing wrt. b (while holding s fixed) yields:

$$\frac{\partial U}{\partial b} = -s \cdot u'\left(y - \frac{1-s}{s}b\right) \cdot \frac{1-s}{s} + (1-s) \cdot u'(b) = 0$$

$$\Leftrightarrow u'\left(y - \frac{1-s}{s}b\right) = u'(b) \Leftrightarrow u'(y-\tau) = u'(b).$$

Given the properties of the utility function $u'(y-\tau) = u'(b)$ implies that also $y - \tau = b$. That is, if the government can fully constrol the individuals search effort (s), it is optimal to fully smooth consumption over the two states (employed and unemployed).

The answer may also show the optimal s, but this is not necessary.

(3B) *Show that the individual optimization, when b and τ are taken as given, implies $v'(s) = u(y-t) - u(b)$. What would be the consequence if the individual had full unemployment insurance in this case? Can government intervention overcome the problem?*

In order to find the individual optimum we maximize 3 wrt. s while taking τ and b as given.

$$\frac{\partial U}{\partial s} = u(y-\tau) - u(b) - v'(s) = 0$$

$$\Leftrightarrow u(y-\tau) - u(b) = v'(s).$$

If the government provided full unemployment insurance in this case the left hand side of the equation above would be zero, and hence search effort would likely be zero too, which is too low compared to the first best (moral hazard).

Given that the government cannot control s, it is very difficult to overcome this problem with (further) government intervention, and the government will have to lower b in order to secure sufficient search effort from individuals. Hence the government faces a trade off between efficiency (sufficient search effort) and social insurance (coverage of income loss).

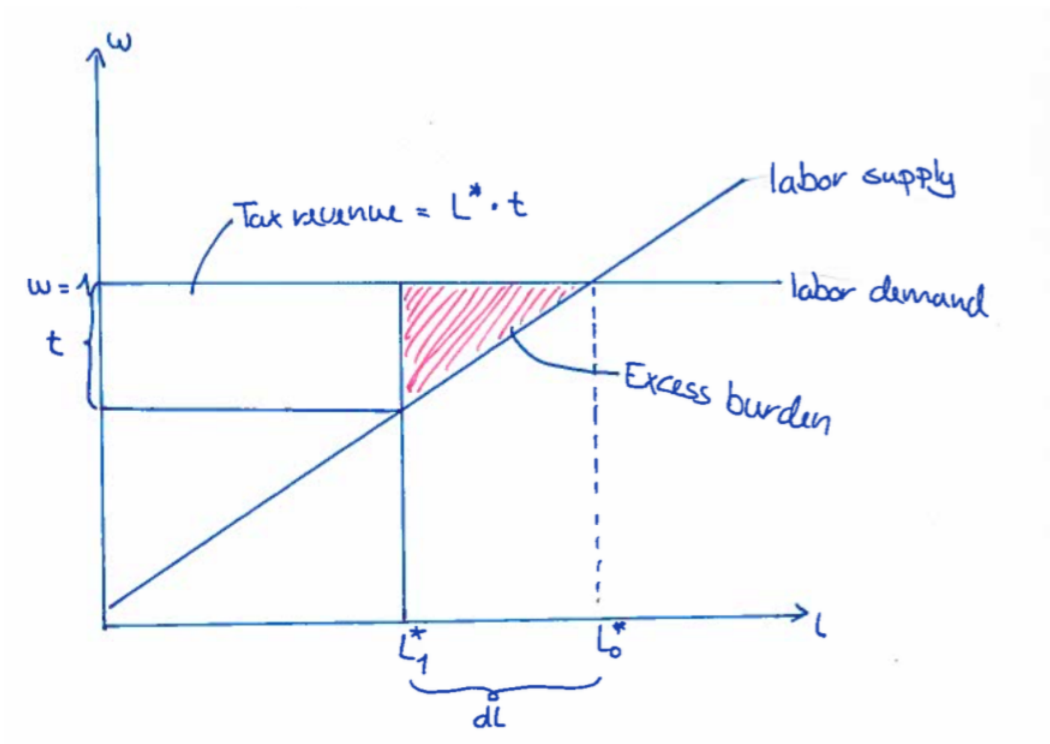
It may be mentioned that one way to (partially) overcome the problem is with active labor market policies that can help to counter the moral hazard problem.

Part 4: The elasticity of taxable income

Consider an economy where the public sector taxes labor income with a constant tax rate t. Assume that labor supply is increasing in the after-tax wage rate and that labor demand is perfectly elastic at a pre-tax wage rate of 1.

(4A) *Illustrate the equilibrium in the labor market with a given tax rate t in a diagram with the number of hours worked l on the primary axis and the wage rate (w) on the secondary axis.*

The market situation is illustrated below.



The tax reduces the labor supply from L_0 (the pre-tax equilibrium) to L_1 and therefore creates the excess burden (or deadweight loss) equal to the red triangle. The government revenue is given by $L_1 \cdot t$.

(4B) *Consider a reform that increases t marginally. Illustrate the effect on the market equilibrium and government revenue. Provide the intuition for why the marginal deadweight loss* *of a small increase in t is equal to the behavioral effect on the government revenue. What are the key parameters that determine the size of the marginal deadweight loss?*

A marginal increase in $t$increases the height of the triangle in the figure above, and thus pushing the equilibrium labor supply further to the left (this should be illustrated in the answer). The resulting change in government revenue can be broken up in to parts: A mechanical effect ($dM = dt \cdot L_0$) and a behavioral effect ($dB = t \cdot dL$).

The resulting increase in the deadweight loss can be broken up in to two pieces: A rectangle equal $dL \cdot t$ (the first order effect) and a small triangle $dt \cdot dL$ (the second order effect). When we are considering a marginal change the second order effect is approximately zero, and the marginal deadweight loss is equal the behavioral effect on the government budget. The intuition for this result is the following:

In equilibrium individuals have optimized so that the disutility from the last unit of labor supplied is equal to gain in terms of extra consumption. This implies that individuals are indifferent about supplying one unit of labor more or less. However with an pre-exiting tax t, the wage that individuals require to work an extra hour is lower than that the employers are willing to pay and this "social surplus" is exactly equal to the tax rate (which is captured by the government). When a small increase in t decreases l by one unit, the loss in term of social surplus (the deadweight loss) is therefore equal to t, which is also the loss of revenue for the government, when l is decreased by one.

Besides t, the size of the marginal deadweight is also affected by the size of dL, which is determined from the labor supply elasticity. Here a higher labor supply elasticity implies that a given tax increase, reduces the labor supplied more. And a larger reduction in labor supply implies a larger marginal deadeight loss.

(4C) *Give the intuition for why it might be more correct to look at a broader concept of labor supply responses than just hours worked. That is, why would it be more correct to look at the change in taxable income with computing the marginal deadweight loss?*

Above we only considered the effect of the tax on the number of hours worked, but the tax might also affect behavior in other dimensions. E.g. a higher tax might reduce the willingness to accepted a higher paying job further away or give a higher incentive to transform earns to fringe benefits (better coffee machines etc.). All of these dimensions cause distortions, and are captured by the change in taxable income. The elasticity of taxable income (ETI) is therefore often called a suffecient statistics.

The answer may contain more elaborate explanations.

In the paper "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform" by Martin Feldstein (published in the Journal of Political Economy in 1995), the author investigates the effect of the 1986 tax reform on the taxable income reported by different income groups. The reform significantly reduces marginal tax rates on taxable income while broadening the tax base. Below is a copy of Table II from the article showing the main estimate from the paper.

TABLE 2

ESTIMATED ELASTICITIES OF TAXABLE INCOME WITH RESPECT TO NET-OF-TAX RATES

Taxpayer Groups Classified by 1985 Marginal Rate	Net of Tax Rate (1)	Adjusted Taxable Income (2)	Adjusted Taxable Income Plus Gross Loss (3)
	Percentage Changes, 1985–88		
1. Medium (22–38)	12.2	6.2	6.4
2. High (42–45)	25.6	21.0	20.3
3. Highest (49–50)	42.2	71.6	44.8
	Differences of Differences		
4. High minus medium	13.4	14.8	13.9
5. Highest minus high	16.6	50.6	24.5
6. Highest minus medium	30.0	65.4	38.4
	Implied Elasticity Estimates		
7. High minus medium		1.10	1.04
8. Highest minus high		3.05	1.48
9. Highest minus medium		2.14	1.25

NOTE.—The calculations in this table are based on observations for married taxpayers under age 65 who filed joint tax returns for 1985 and 1988 with no age exemption in 1988. Taxpayers who created a subchapter S corporation between 1985 and 1988 are eliminated from the sample.

(4D) *Describe the empirical analysis and explain, using Table 2 above, how the author arrives at the estimates of the implied elasticity of taxable income (ETI). What are the main identifying assumptions needed for the estimates to be the causal effect of the marginal tax rates on taxable income?*

Feldstein (1995) estimate the effect of changes in the marginal tax rate using a difference-in-differences (DiD) estimation framework. However contrary to "normal" DiD estimates, he does not have an untreated control group but rather different groups with different treatment intensity. Here the treatment intensity rises with income. I.e. high income groups experience the largest changes in their marginal tax rate.

In Table II we see the simple DiD estimation table. Denote the precentage change in taxable income from before to after the reform of group i as $\Delta log(E^i)$ where i can be Medium (M), High (H1) or Highest (H2). Similar denote the precentage change in the net-of-tax income $(1-t)$ as $\Delta log(1-t^i)$.

Using this, we can compute the implied elasticity in row 7 (High minus medium) by

$$ETI = \frac{\Delta log(E^{H1}) - \Delta log(E^M)}{\Delta log(1-t^{H1}) - \Delta log(1-t^M)} = \frac{21,0-6.2}{25,6-12.2} = 1.10,$$

which is the DiD estimate. The estimate has the expected sign as the reduction in marginal tax rates seem to have increased the taxable income most among the groups that experience the largest fall in marginal tax rate. However, the size of the implied elasticities are much larger than modern studies show today.

For the estimates to be causal we need two assumptions:

- The common/parallel trend assumption, which states that, in the absence of the reform, the change in the taxable income for the three groups should be the same.
- Same underlying elasticity for all groups. I.e. if the groups had gotten the same change in the net of tax rates, we would have see the same change in taxable income for all groups.

If one of these assumptions are invalid the estimates will be biased.

(4E) *Describe how you could validate the main identifying assumptions needed in (4D) and what kind of data you would need to do so.*

One way to validate if the common trend assumption seems plausible is to consider the evolution of the taxable income rates of the diffenret group before and/or after the reform. If the taxable incomes move in parallel in these years it speaks to the validity of the common trend assumption. More formally this can be tested by running the same DiD estimators in non-reform years. The estimates from these "Placebo tests" should be insignificant.

The same underlying elasticity for all groups assumptions is more difficult to validate. Here you would need other reforms, where all groups were treated to the same extend and see if taxable incomes move in parallel in these years.