Answers

Re-examination exam in Public Finance - Spring 2018

3-hour closed book exam

By

Jakob Egholt Søgaard

## Part 1: Tax Incidence

*Consider a worker with the following utility function*

$$u(c, h) = c - v(h), \tag{1}$$

*where $c$ is consumption and $h$ is labor supply. The worker maximizes utility subject to the following budget constraint*

$$c = w_S h, \tag{2}$$

*where $w_S$ is the after-tax wage rate that the worker receive. The consumer takes $w_S$ as given.*

***(1A)*** *Show that the effect of a marginal increase in $w_S$ on the worker's utility is given by*

$$\frac{\partial u(c,h)}{\partial w_S} = h. \tag{3}$$

*Provide intuition for the result. In particular, for why the effect is independent of the worker's behavioral response to the higher after-tax wage rate.*

#

**First we find** the worker's first order condition

$$\frac{du(c,h)}{dh} = \frac{dc}{dh} - v'(h) = w_S - v'(h) = 0.$$

This equation implicitly defines $h$ as a function of $w_s$. Next we differentiate (1) wrt. $w_S$ and uses the first order condition above

$$\frac{du(c,h)}{dw_S} = \frac{dc}{dw_S} - v'(h)\frac{\partial h}{\partial w_S},$$

where $\frac{dc}{dw_S} = \frac{\partial c}{\partial w_S} + \frac{\partial c}{\partial h}\frac{\partial h}{\partial w_S} = h + w_S\frac{\partial h}{\partial w_S}$. Use this to get

$$\frac{du(c,h)}{dw_S} = h + w_S\frac{\partial h}{\partial w_S} - v'(h)\frac{\partial h}{\partial w_S} = h + \underbrace{(w_S - v'(h))}_{=0}\frac{dh}{dw_S} = h,$$

which is what we were asked to show.

The result implies that when the worker experiences a marginal increase in the wage rate ($w_S$) his utility is simply affected by the mechanical effect: with initial work hours of $h$, a 1\$ increase in the wage rate increases his disposable income by $h$\$. The result also implies that even thought the worker is likely to increase his labor supply in response to the change in $w_S$, this behavioral response has no first order effect on his utility. This is due to the fact that the individual initially has maximized his utility, and therefore he is different between working one hour more or less (The envelope theorem).

*Assume that the after-tax wage is given by $w_S = w_D - t$, where $w_D$ is the wage rate that firms pay out to workers, and that the initial equilibrium on the labor market is given by $S(w_S) = D(w_D)$, where $S(w_S)$ is aggregate labor supply and $D(w_D)$ is aggregate labor demand.*

***(1B)*** *Show that the effect of an increase in $t$ on the after-tax wage ($w_S$) is given by*

$$\frac{dw_S}{dt} = -\frac{\varepsilon_D}{\varepsilon_D + \varepsilon_S\frac{w_D}{w_S}} \approx -\frac{\varepsilon_D}{\varepsilon_D + \varepsilon_S}, \tag{4}$$

*where $\varepsilon_D = -\frac{dD(w_D)}{dw_D}\frac{w_D}{D(w_D)}$ is the (numerical) elasticity of labor demand and $\varepsilon_S = -\frac{dS(w_S)}{dw_S}\frac{w_S}{D(w_S)}$ is the elasticity of labor supply. The last approximation holds when $t$ is small. Describe the economic intuition behind the formula.*
#

**Note that there** is a typo in this question. The elasticity of labor supply should be defined as $\varepsilon_S = \frac{dS(w_S)}{dw_S}\frac{w_S}{D(w_S)}$.

In order to derive (4) we start from the equilibrium condition on the labor market, differentiate wrt. the tax, rewrite the result to isolate the two elasticities:

$$\frac{dS(w_S)}{dw_S}\frac{dw_S}{dt} = \frac{dD(w_D)}{dw_D}\frac{dw_D}{dt} \Leftrightarrow \underbrace{\frac{w_S}{S(w_S)}\frac{dS(w_S)}{dw_S}}_{=\varepsilon_S}\frac{dw_S}{dt}\frac{S(w_S)}{w_S} = \underbrace{\frac{w_D}{D(w_D)}\frac{dD(w_D)}{dw_D}}_{=-\varepsilon_D}\frac{dw_D}{dt}\frac{D(w_D)}{w_D}.$$

Next, using that $S(w_S) = D(w_D)$ and that $w_S = w_D - t \Rightarrow \frac{dw_S}{dt} = \frac{dw_D}{dt} - 1$, we obtain

$$\varepsilon_S \frac{dw_S}{dt}\frac{w_D}{w_S} = -\varepsilon_D\left(\frac{dw_S}{dt} + 1\right) \Leftrightarrow \frac{dw_S}{dt} = -\frac{\varepsilon_D}{\varepsilon_D + \varepsilon_S\frac{w_D}{w_S}} \approx -\frac{\varepsilon_D}{\varepsilon_D + \varepsilon_S}.$$

The formula shows that the economic incidence on the worker is determined by the relative size of the elasticities. Consider, for example, the case where labor supply is completely inelastic ($\varepsilon_S = 0$). In this case, workers supply a fixed number of work hours, completely independent of the after tax wage rate. A higher tax will reduce labor demand, but, because the labor supply curve is vertical, this will go directly into a lower wage to workers. In the other extreme case, labor supply is perfectly elastic ($\varepsilon_S \to \infty$) implying that workers are will to work any number of hours as long as the wage rate is above some fixed level. In this case, firms will reduce demand until the point where they are willing to pay workers the original wage plus the tax, and firms will therefore bear the full burden of the tax. These points may be illustrated graphically.

## Part 2: Labor taxation in the short run and in the long run

*Consider a model economy where firms hire labor ($L$) and rent capital ($K$) to produce output ($Y$) according to the following Cobb-Douglas production function*

$$Y = K^{\alpha}L^{1-\alpha}. \tag{5}$$

*All markets are perfectly competitive and the wage rate ($w_D$) and the rental rate ($r_D$) that firms pay therefore equal the marginal product of labor and capital, respectively. Both labor and capital income are taxed so that the wage rate that workers receive ($w_S$) and the rental rate that capital capital owners receive ($r_S$) are given by*

$$w_S = (1 - t_L)w_D \tag{6}$$

$$r_S = (1 - t_K)r_D. \tag{7}$$

*Finally, assume that workers supply labor according to an aggregate labor supply function*

*$L(w_S)$ with a constant elasticity $\varepsilon$. Log transforming and total differentiating the model above yields the following five model equations*

$$\begin{aligned}
\hat{w}_D &= \alpha\left(\hat{K}-\hat{L}\right), && (8)\\
\hat{r}_D &= -(1-\alpha)\left(\hat{K}-\hat{L}\right), && (9)\\
\hat{w}_S &= -\frac{dt_L}{1-t_L}+\hat{w}_D, && (10)\\
\hat{r}_S &= -\frac{dt_K}{1-t_K}+\hat{r}_D, && (11)\\
\hat{L} &= \varepsilon\hat{w}_S, && (12)
\end{aligned}$$

*where $\hat{x} = dx/x$, that is, the percentage/relative change in $x$. The government revenue is given by*

$$R = t_L w_D L + t_K r_D K. \tag{13}$$

***(2A)*** *Show that in the short run, where the capital stock is assumed fixed ($K = \bar{K}$), the effect of an increase in the tax on labor ($t_L$) on labor supply is given by*

$$\hat{L}_{Short} = -\frac{\frac{1}{\alpha}\varepsilon}{\frac{1}{\alpha}+\varepsilon}\frac{dt_L}{1-t_L}. \tag{14}$$

*Comment on the expression and on the importance of $\alpha$ and $\varepsilon$, respectively.*
#

**We find the** effect on labor supply in the short run by inserting the equations (8) and (10) into (12):

$$\hat{L} = \varepsilon\left(-\frac{dt_L}{1-t_L}-\alpha\hat{L}\right) \Leftrightarrow \hat{L} = -\frac{dt_L}{1-t_L}\frac{\varepsilon}{1+\alpha\varepsilon} = -\frac{\frac{1}{\alpha}\varepsilon}{\frac{1}{\alpha}+\varepsilon}\frac{dt_L}{1-t_L}.$$

From this equation we see that the fall in equilibrium labor following an increase in $t_L$ depends positively on $\frac{1}{\alpha}$ and $\varepsilon$. $\varepsilon$ comes from the labor supply function and determines how much a given change in $\hat{w}_S$ affects labor supply. $\frac{1}{\alpha}$ comes from the labor demand function, where $\alpha$ determines how much $\hat{w}_D$ increases for a given fall in $L$. When $\alpha$ is large ($\frac{1}{\alpha}$ is small), labor demand is inelastic in which case an increase in $t_L$ push up $\hat{w}_D$. In this case $\hat{w}_S$ – and therefore $\hat{L}$ – fall less. It could be noted that this formula corresponds to the standard formula in a market with both supply and demand effects.

***(2B)*** *Show that in the long run, where the capital stock is assumed to be perfectly elastic* *at the world interest rate level ($r_S = \bar{r}$), the effect of an increase in the tax on labor ($t_L$) on labor supply is given by*

$$\hat{L}_{Long} = -\frac{dt_L}{1-t_L}\varepsilon \tag{15}$$

*Comment on the expression and compare it to the effect of an increase in the tax on labor in the short run. Provide intuition for the difference.*

#

**Under the assumption** that $r_S = \bar{r} \Rightarrow \hat{r}_S = 0$ we see from equation (11) that $\hat{r}_D$ is independent of $t_L$. From equation (9) this further implies that so is $\hat{K} - \hat{L}$ (the capital-labor ratio). A fixed capital-labor ratio implies that $\hat{w}_D = 0$ from equation (8) and the effect of $t_L$ on labor supply is therefore simply given by inserting equation (10) into (12)

$$\hat{L} = \varepsilon\left(-\frac{dt_L}{1-t_L}\right) = -\varepsilon\frac{dt_L}{1-t_L}.$$

From this equation we see that labor supply in the long run only depends on the labor supply elasticity. This is due to the fact that capital is assumed to be perfect elastic in the long run (the result can also be obtain by letting $\frac{1}{\alpha} \to \infty$ in equation (14)). This also implies that the fall in equilibrium labor following an increase in $t_L$ is larger in the long run than in the short run. The intuition for this result is that in the short run (when capital is fixed), an increase $t_L$ is mitigated by an increase $w_D$, and thereby a small decrease in $w_S$ and $L$.

***(2C)*** *Is the behavioral effect on the government revenue of an increase in the labor tax largest in the short or in the long run? Does the answer depend on whether capital is taxed ($t_K > 0$) or not ($t_K = 0$)?*

#

**The behavioral effect** on the govenrment's revenue of an increase in $t_L$ is proportional to $\hat{L}$, and the effect is largest in the long run following the argument in the answer to (2B). The answer is independent of whether the tax on capital is zero or positive. In fact a positive tax on capital will reinforce the behavioral effect in the long run (but not in the short run). In the long run $\hat{K} - \hat{L} = 0 \Leftrightarrow \hat{K} = \hat{L}$ and a fall in $L$ therefore results in a proportional fall in $K$. When capital is taxed, this fall also reduces the government's revenue (a larger behavioral effect). This is not the case in the short run, where capital is fixed.

## Part 3: Inequality and intergenerational mobility

***(3A)*** *Define the Pigou-Dalton principle and discuss when the principle is sufficient to rank two (income) distributions according to their degree of inequality.*

#

**The Pigou-Dalton principle** states that an inequality measure must decrease if there is a transfer of income from a richer household to a poorer household conditional on the transfer being a so-called Pigou-Dalton transfer, that is:

- The transfer preserves the rank of the two households (the transfer must not be so large that the poorer household becomes the richer).

- The transfer leaves total income unchanged (pure redistribution).

Pigou-Dalton principle is sufficient to rank distributions according to their degree of inequality if you can transform one of the distribution to the other using only Pigou-Dalton transfer. This point can be illustated by comparing two Lorenz curves. If the two Lorenz curves do not cross, you can transform the lowest Lorenz curve to the highest Lorenz curve by a number of Pigou-Dalton transfer; and as each Pigou-Dalton transfer reduces inequality, the lowest Lorenz curve is unambigiously associated with a lower level of inequality.

If instead the two Lorenz curves cross, then you will need Pigou-Dalton transfers to transform one of the curves to the other in one part of the income distribution and opposite Pigou-Dalton transfers in other parts of the income distribution. In this case it is therefore not possible to unambigiously rank the two distributions according to their degree of inequality using only the Pigou-Dalton principle.

***(3B)*** *Describe how "inequality" and "intergenerational mobility" are different concepts (although they may be related).*

#

**Inequality measures the** variation across individuals in some outcome, for example variation in income or wealth at a given point in time or differences in lifetime income across individuals. Intergenerational mobility measures to what extent outcomes are related across generations. A high degree of intergenerational persistence (low degree of mobility) implies that a high degree of inequality is transmitted to the next generation. The concepts are therefore related to each other, although they are not the same.

To see the difference between the two concepts, consider as an example two countries that have the same income inequality throughout an extended period of time. One country has no intergenerational mobility, implying that a child gets the same position in the distribution as the parents, while the other country has perfect intergenerational mobility, implying that the position of a child in the distribution is completely unrelated to that of its parents. Thus, the two countries have the same distribution, but very different intergenerational mobility, with parents being crucial for outcomes of children in one country, but not in the other country.

*The article "Intergenerational Wealth Formation over the Life Cycle: Evidence from Danish Wealth Records 1984–2013" in the American Economic Review (2016) by Boserup, Kopczuk and Kreiner studies the impact of bequests following parental death on the wealth distribution of the next generation. Below (next page) is a copy of Figure 2 and Figure 3 from the article.*

***(3C)*** *Describe the results in each of the graphs and how the graphs lead to different conclusions about the effect of bequests on inequality. How can this difference be explained/reconciled?* #

**Boserup et al. (2016)** divide individuals into treatment and control groups depending on whether a parent dies at a given point in time (denoted by 0 in the graphs) or do not die. The aim is to measure the impact on the wealth distribution of the next generation receiving bequests, which is not directly observable in the data. The left panel in Figure 2 displays the percentage difference in the variance of the wealth distribution between the treatment group and the control group over time. The difference in the seven years before death of a parent is close to zero and then increases by around 30 percent after death of the parent. Thus, bequests increase the variance of the wealth distribution of the next generation. The right panel in Figure 2 displays the difference in 75th wealth percentile between the treatment group and the control group over time. Similar to the left panel, the difference is close to zero before time 0. After the death of the parent the difference increases to around 150,000 DKK.

Figure 3 shows the share of wealth owned by the one percent most wealthy in the treatment group and the control group, respectively. In the seven years prior to death of the 3 parent, the top 1% share of the control group lies somewhat higher than the treatment group, and the wealth share varies over time with the business cycle, but the wealth shares of the two groups co-vary reasonably well with a fixed difference over time. After death of the parent, the gap between the two curves increases (actually, it occurs somewhat already the year before, which may be due to wealth transfers in order to avoid inheritance taxation). This implies that the wealth share of the treatment group falls compared to the control group (by 6 percentage points), implying that bequests reduce the share of wealth owned by the top 1% wealthiest (and equivalently that the wealth share of those not in the top 1% group increases).

As described above the conclusion from Figure 2 is that bequests increase wealth inequality, while the conclusion from Figure 3 is that bequests reduce wealth inequality. The reason for this difference in conclusions is the use of different inequality concepts. The variance of the distribution is a measure of absolute inequality, focusing on absolute differences across people. Inheritances stretch the distribution because the rich inherits larger amounts, which increases absolute inequality. The top 1% wealth share measures one groups wealth out of total wealth, which is a measure of relative inequality. If the rich inherit a smaller percentage of their initial wealth than the less rich (which is the case here) then relative wealth inequality decreases. To conclude, whether bequests increase or decrease wealth inequality depends on how we measure wealth inequality.

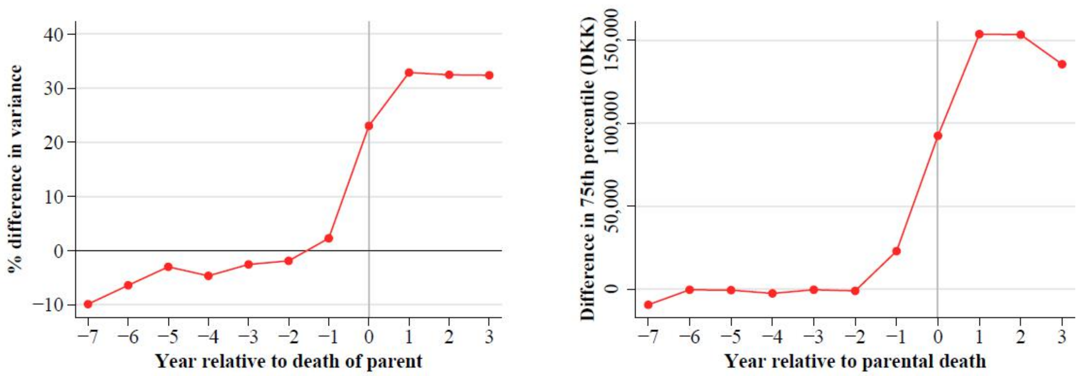
**(3D)** Provide an argument for whether or not the results in Boserup et al. (2016) are likely to be causal estimates of the effect of bequests on the wealth distribution of the next generation. Is there anything in the graphs that validates or invalidates a causal interpretation?

#

**The estimation method** in Boserup et al. (2016) is essentially a difference-in-difference estimation that compares a treatment and control group over time. The key assumption in this estimation method is the common/parallel trend. The parallel trend assumption states that the treatment and control groups should have evolve in parallel absent the treatment (in this case the death of a parent).

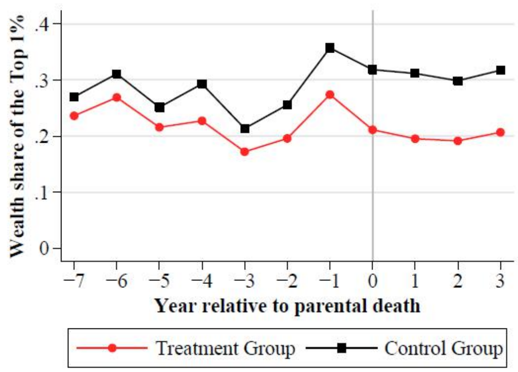
The parallel trend assumption cannot be proven as we do not observe the evolution of the treatment group absent the treatment, but the assumption can be validated by comparing the treatment and control group on outcomes that should not be effected by the treatment, such as e.g. outcomes prior to the treatment. The figures from Boserup et al. (2016) enable exactly this type of validation. In particular in Figure 2 (the right panel) and in Figure 3 we see that the treatment and controls groups evolve in parallel prior to the treatment, which speaks to the validity of the parallel trend assumption, and thereby that the estimates are the causal effect of bequests on the wealth distribution of the next generation.

Figure 2. Effects of Bequests on the Variance and the 75th Percentile of the Wealth Distribution




*Note:* Percentage difference in variance of treatment group relative to control group (left panel), and difference between treatment group and control group in the value of the 75th percentile (right panel). Variance based on the distributions censored at the 1st and 99th percentiles. Weighting as in Figure 1. $1 = DKK5.6 in 2010.

Figure 3. Effect of Bequests on Top 1% Share of Wealth




*Note:* Top 1% share in the treatment and control groups. Weighting as in Figure 1.