Public Economics

Problem Set 2: Welfare Analysis, Consumption Taxes and Labor Income Taxation

Claire LEROY

claire.leroy@ensae.fr

Fall Semester 2022 (PSE)

Due date: November 17th at 10pm. Send your homework to me by email.
NB: This solution is deliberately more detailed than what is expected from you. The additional details are provided in order to help you understand the key concepts mentioned in this problme set.

1 Empirical Discussion (10pt)

Read the following paper: "Labor Supply Responses to Learning the Tax and Benefit Schedule" (AER, 2021) by A. Kostøl and A. Myhre[1]. Answer the following questions using both the article and the concepts seen in the lecture. Your answers should be concise and not exceed a small paragraph.

1. This paper studies labor supply responses and the labor earnings elasticity with respect to tax incentives. Why is this elasticity an interesting parameter to estimate in public economics? (1 pt)

   Answer: There is an extensive empirical literature in economics trying to estimate the elasticity of labor supply (or labor earnings) with respect to the net of tax rate. When designing and setting the level of taxes and transfers, policy makers need to balance redistributive objectives with the potential efficiency loss introduced. The elasticity of labor earnings can help assess the magnitude of the variation in deadweight loss. The larger the labor earnings response the larger taxes and transfers will create distortions. While this elasticity is informative, it might not be a sufficient statistic to assess the overall deadweight loss.

2. What is the research question addressed by the paper? Why is it a relevant one? (1 pt)

   Answer: The paper addresses two questions: how responsive is the labor supply to tax incentives? and how much do different frictions attenuate labor supply responses? The authors are able to estimate a frictionless elasticity and

[1] The paper is available here and you should also have received it by email.

to document the relative importance of information frictions versus other types of frictions in reducing this elasticity. While the frictionless elasticity is often thought as a structural parameter coming from individual preferences, policy makers can affect frictions and thus shape earnings responses. Understanding the channel through which frictions arise is key because (i) it implies different policy instruments and (ii) it has different implications in terms of optimal policy design.

3. Describe briefly how the disability insurance (DI) benefit worked in Norway before the 2015 reform. In your opinion, what motivated the elimination of the notch? (1 pt)

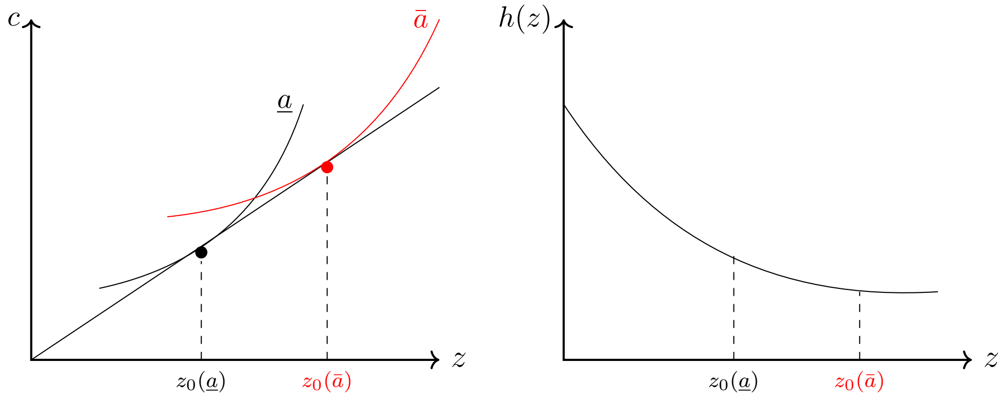
   Answer: The disability insurance benefit in Norway is a cash transfer targeting individuals with disabilities that prevent them from working (either completely or partially). The benefit can be fully cumulated with labor earnings up until a threshold called SGA. Above this threshold, benefit can only be partially cumulated with earnings. Two systems of means-testing existed.

   - Individuals who entered the program before 2004 saw their benefit phased out at a rate τ when their earnings exceeded the SGA threshold. This means that for each additional euro of labor earnings they get, their benefit is decreased by $100 \times \tau$ cents.
   - Individuals who entered the program after 2004 face a discontinuous decrease of T in their benefit (i.e. a notch) if their earnings exceed the SGA threshold. After this, their benefit is also phased out at a certain rate τ.

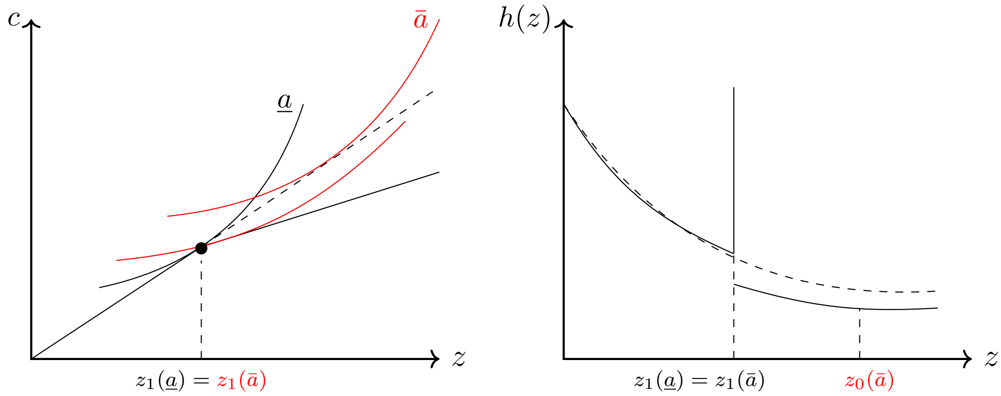
   The elimination of the notch system after the 2015 reform could be motivated by both a redistributive and an efficiency concern. First, the notch makes individuals just above the SGA experience a large decrease in their disposable income which might conflict with the redistributive motives of the government. Second, eliminating the notch should encourage work among DI recipients as a notch strongly discourage workers who would like to earn just above the SGA threshold and generates the idea that it does not always pay to work more.

4. One of the strategies used by the authors to estimate earnings elasticities is to use bunching at the kink of the DI system. Represent graphically a kinked budget set as in Figure 2 Panel A and illustrate with indifference curves why we should expect bunching. Represent in another figure the effect of the introduction of a kink on the distribution of labor earnings z. (1 pt)

   Answer: For the purpose of clarity, we start by representing these figures in the case of a linear tax system. We also assume that the preferences are convex and the distribution of abilities a is continuous and smooth.



Now we consider what happens if a kink is introduced in the budget constraint at income $z_0(\underline{a})$. We also assume here that there are no adjustment costs ψ and that elasticities are homogeneous. The figures become:



Suppose that $\bar{a}$ is the type of the marginal buncher. The figure illustrates that when there is a kink in the budget constraint, the distribution of labor earnings z is no longer continuous because a share of individuals (with abilities $a \in [\underline{a}; \bar{a}]$ here) bunch at the kink point $z_0(\underline{a})$. This is why we have a discontinuous increase in the earning density at that point. Above that point, since individuals who would earn between $z_0(\underline{a})$ and $z_0(\bar{a})$ have now decreased their labor earnings to bunch, we have a missing mass in the distribution in this range. Note that even if individuals with abilities higher than $\bar{a}$ do not bunch at the kink, they might still reduce their labor earnings due to the increase in the marginal tax rate. In that case, there is not a zero mass in the $[z_0(\underline{a}); z_0(\bar{a})]$ range because people with abilities higher than $\bar{a}$ might now locate in this income range.

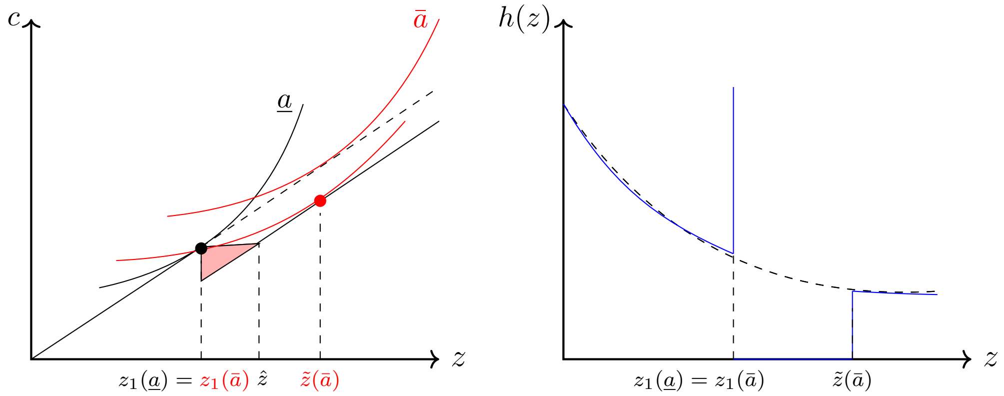
Highly recommended reading about bunching: Kleven (2016).

(a) *(Bonus question)* Consider the model described in part III.A. With a linear tax system, there is a smooth and continuous density of ability a_i leading to a smooth and continuous density of z_i. Using equation (1), show that with a kink at $z = K$, there will be bunching i.e that there will be a range of individuals with skills a_i locating at $z = K$. [Hint: For what range of ability level a_i is $z_{F,i} < K$ and for what range is $z_{F,i} > K$?]

(b) *(Bonus question)* Using the answer to the previous question, derive the range of pre-reform earnings z for which we expect a behavioral response to the reform. Express the excess mass of bunchers at the kink z and show that it is related to the structural elasticity e and the tax parameters.

5. The paper also relies on bunching at notches to identify earnings elasticities as well as the share of nonoptimizers. Answer the same question as in question 4 above but in the case of a notch. Do you expect bunching to be larger at a notch or at a kink and why? (1 pt)

   Answer: Start with the same initial situation as depicted in the answer for question 4. We introduce a notch at $z = z_0(\underline{a})$, i.e. a discontinuous increase in the tax amount. Again, we suppose that individuals with ability $\bar{a}$ are the marginal bunchers. We can see that the marginal buncher is indifferent between locating at the notch or at $\tilde{z}(\bar{a})$. As for kinks, there will be bunching at the notch for all individuals with ability $[\underline{a}; \bar{a}]$. Contrary to the case of kinks, a notch should create an empty hole in the earning distribution above the notch, i.e. for income in the range $[z_0(\underline{a}); \tilde{z}(\bar{a})]$. Note that there is a dominated region $[z_0(\bar{a}); \hat{z}]$ but that the hole in the distribution correspond to a larger region than just the dominated region.



   (a) *(Bonus question)* Re-rederive equation (4) from the paper. [Hint: use the fact that the marginal buncher at the notch, who has earnings $z + \Delta z^*$ pre-reform, must be indifferent between locating at the notch or at a point after the notch.]

6. A key challenge with the bunching technique is the estimation of the counterfactual distribution of earnings. What are the two methods used in the paper to estimate it? In which case can we say that bunching is a non-parametric estimation approach? (1 pt)

   Answer: The bunching technique[2] is a method to estimate the labor earnings elasticity $e = \frac{\Delta z^*}{\Delta \tau} \frac{\tau}{z}$ using the variation of marginal tax rate τ that arises around a kink at income $z = K$. Earnings z are observed and the magnitude of the tax change $\Delta\tau/\tau$ around the kink is known. We thus only need to estimate the behavioral response to the tax change Δz^*.

[2] We will present here the method in the case of a kink. The general idea is quite similar in the case of a notch.

The method relies on the relationship that exists between (i) the excess mass of bunchers at a kink ($B/h_0(K)$) and (ii) the structural labor earning response Δz^* to this kink:

$$B/h_0(K) = \int_K^{K+\Delta z^*} h_0(z)dz, \quad \text{with } h_0(.) \text{ the pre-kink pdf of } z$$

We are interested in estimating Δz^* from this equation. The total mass of bunchers B is observed, but to estimate Δz^* from this equation, we need to know the counterfactual density h_0 (i.e. the density that would arise in a world without a kink). This is challenging because, by definition, h_0 is unobserved.

In this paper, the authors use two methods to estimate the counterfactual earning distribution in 2015.

(a) A non-parametrical approach that relies on variation of the DI schedule over time

The idea is to infer the distribution of earnings in 2015, absent any notches or kinks at the SGA_{15} (=\$ 8,000) threshold, by using the observed distribution of earnings in 2014. The 2014 distribution is indeed not "polluted" by bunching behaviors around the SGA_{15} threshold because the threshold was higher ($SGA_{14} = \$12,000$). Therefore, while we should expect bunching around SGA_{14} in 2014, the 2014 distribution should look like h_0 around the SGA_{15} threshold.

(b) A structural approach that relies on a polynomial fit

The other method consists in estimating h_0 using the observed 2015 earnings distribution, except around the kink. Indeed, with optimization frictions, we should expect the presence of a kink to affect the number of individuals with income around the kink (not just exactly at the kink). However, for income range further from the kink, there should be no effect of the kink and the observed density should be close to h_0. See Equation (3) in the paper for how they fit a polynomial to estimate the density.

7. What is the identification strategy for the information treatment? On which identification assumptions does it rely on? Based on Table 4, what is the magnitude of the treatment effect? Do you see any limitation(s) to this strategy? (1pt)

Answer: The effect of the information treatment on the elasticity of labor earnings is identified by a methodology mixing bunching and simple difference[3]. The bunching elasticity of those receiving the information letter (treated) is compared to the bunching elasticity of those eligible to the letter but who did not receive it in the end (control). The simple difference approach relies on the assumption that the treatment assignment is quasi-random among the sample of

[3]The paper also relies on an event-study approach to identify the effect of the treatment on other outcomes in Section 6.

eligible individuals. The authors explain how the administration administered the information treatment: "*To implement the targeted intervention, the SSA used monthly earnings records from January to May 2015 to forecast annual earnings. Our research design uses forecast errors due to fluctuations in monthly payments and electronic reporting by employers generating quasi-random variation in information letters*" (p.3734). In Table 4, the authors estimate an elasticity of 0.15 for the treated and 0.058 for the control. Providing information about the kink location thus increased by 0.092 p.p.t the earning elasticity.

One limitation of this approach arises from the fact that the bunching technique requires large sample size while policy intervention like information letters are usually conducted on a small number of individuals. In this paper, the main sample of interest is made of about 40,000 individuals (cf. Table 2) but the number of treated and controls is relatively small (3,642 and 1,521 respectively). In result, the histograms are very bumpy (cf. Figure 9) which might lead to a bias when estimating the counterfactual density h_0.

8. Table 5 presents the main results of the paper. Explain with words how do the authors estimate or compute the following parameters: (2 pt)

   (a) The share of non-optimizers

   Answer: The share of non-optimizers (α) is obtained by exploiting the theoretical prediction that there should be a density of zero in the dominated range of the notch if there were no frictions (cf. question 5 above). If an individual locates in this region (a "non-optimizer" hereafter), it means that she is facing strong enough frictions to prevent her from bunching at the notch. In the case of Table 5, they estimate this share among the sample of individuals eligible for the information letter, that faced a notch in 2014.[4] The share of non-optimizer is simply the observed number of individuals locating in the dominated region of the notch in 2014 divided by the counterfactual number of individuals locating in this region obtained by using the estimated counterfactual density.

   (b) The observed elasticity

   Answer: The observed elasticity for Panel A is obtained by using the bunching technique applied to the notch created by the SGA_{14} threshold. This technique allows to estimate the magnitude of the labor earning response Δz by using an estimated counterfactual density h_0 and the observed amount of people bunching at the notch. Equation (4) then relates the elasticity with this estimated response and the size of the notch (Δt). Note that Equation (4) does not provide directly a closed-form expression of the elasticity in terms of Δz and Δt but it can still be solved for numerically. For Panel B, the observed elasticity is obtained using the bunching technique applies to a kink (cf. details in question 6).

[4] Remember that only those who entered the DI program after 2004 faced a notch in 2014. Column 1 of Table 5 shows that only 2,381 individuals among the 5,163 eligible did face a notch. By definition, it is not possible to estimate the share of the non-optimizer among those facing a kink instead of a notch.

(c) The structural elasticity

Answer: The structural elasticity is obtained by combining the estimated observed elasticity (i.e. the elasticity in a world with frictions) and the share of non-optimizers (i.e. the share of individuals that were prevented from responding to the notch because of frictions). The structural (frictionless) earning response is $\Delta z^* = \Delta z/(1-\alpha)$, it is a function of the share of non-optimizer α and the observed earning response Δz. Using Equation (4), the observed earning response Δz and the estimated share of non-optimizer α, the authors are able to estimate the structural elasticity.

(d) The total attenuation in the elasticity

Answer: The total attenuation in the elasticity is simply given by the estimated structural elasticity (i.e. the frictionless elasticity) minus the observed elasticity (i.e. the attenuated elasticity. Note that the structural elasticity was only estimated for year 2014 and for the sub-sample of individuals facing a notch. The authors still use it to derive the attenuation for other years and other samples (as in Panel B of Table 5). This relies on the key assumption that the structural elasticity is indeed a structural parameter, i.e. that is policy-invariant (same e for a notch or for a kink) and time-invariant (same e for 2014 or 2015).

(e) The attenuation due to information frictions

Answer: The authors are able to decompose the total attenuation in labor earnings elasticity by decomposing between information frictions vs other types of friction. Their estimated share of the attenuation arising from information frictions is simply the treatment effect of the information letter, discussed in question 7. This estimate is a lower bound of the true share as the information letter might not have fully corrected the misperceptions of individuals regarding the DI system.

9. What are the main takeaways of the paper? (1 pt)

Answer: Using bunching at a notch and under-optimal choices by individuals locating in the dominated region of this notch, the paper estimates a frictionless elasticity of labor earnings with respect to the net after-tax earnings of 0.286. In reality, the observed elasticity is about 70% lower due to the existence of frictions. Using a randomized information letter experiment, the paper shows that at least 30% of this attenuation is due to imperfect information about the DI system and especially about the location of kinks and notches.

2 Exercise: Consumption Subsidies (10pt)

In its Q3-2022 Gas Market Report, the International Energy Agency (IEA) notes that "Russia's invasion of Ukraine has exacerbated the tightening supply of natural gas underway since mid-2021, further pushing up prices for consumers". Many fear that the current energy crisis will increase significantly the cost-of-living this winter, especially for low-income households. This exercise studies the effects of a gas subsidy to address this public policy problem.

Suppose that there is a representative consumer in the economy, with an income of R euros that she can spend on two goods: (i) natural gas g (measured in kWh) with a price of p euros per kWh and (ii) a composit numeraire good x. The utility function of the consumer is defined by $U(x,g) = x^{\alpha} g^{1-\alpha}$ with $\alpha \in \,]0;1[$. There is also a representative producer that can produce an amount of g kWh of gas by using $C(g) = \frac{1}{3}g^3$ units of the composit good.

1. Write and solve the consumer's utility maximization problem.

   Answer: The consumer faces the following problem:

$$\begin{cases} \max_{\{x,g\}} \; U(x,g) = x^{\alpha} g^{1-\alpha} \\ s.t. \;\; x + pg \leq R \end{cases}$$

   Solving for this problem, the demand for gas is $g(p,R) = (1-\alpha)\frac{R}{p}$ and the demand for the composite good is $x(p,R) = \alpha R$.

2. Following a negative supply shock, the price of gas has increased from p to p'. The government would like to compensate the consumer for the welfare loss it will cause her.

   (a) Derive the following three measures of the consumer's welfare loss in terms of the parameters of the problem: (i) the change in the consumer surplus, (ii) the equivalent variation (EV) and (iii) the compensating variation (CV).

   Answer: The variation in consumer surplus is given by

$$\begin{aligned} \int_{p'}^{p} g(p,R)dp &= \int_{p'}^{p} (1-\alpha)\frac{R}{p}dp \\ &= [(1-\alpha)R \; ln(p)]_{p'}^{p} \\ &= (1-\alpha)R \; ln(p/p') \end{aligned}$$

   Note that this variation is negative as $p \leq p'$.
   To derive the equivalent variation and the compensating variation, we need to express the indirect utility first. The indirect utility $V(p,R)$ is

$$V(p,R) = (\alpha R)^{\alpha} \left((1-\alpha)\frac{R}{p}\right)^{(1-\alpha)}$$

   Since indirect utility is an ordinal concept, we can apply without loss of generality a log transformation to $V()$ and deduce some constant c to define a new indirect utility as

$$\tilde{V}(p,R) = log(V(p,R)) - c = ln(R) - (1-\alpha)ln(p)$$

   The compensating variation (CV) is the amount of money the individual is willing to pay to experience the price increase. It is defined by:

$$\begin{aligned}
&\tilde{V}(p,R) = \tilde{V}(p', R - CV) \\
\Leftrightarrow\ & ln(R) - (1-\alpha)ln(p) = ln(R - CV) - (1-\alpha)ln(p') \\
\Leftrightarrow\ & ln(R - CV) = ln(R) + (1-\alpha)(ln(p') - ln(p)) \\
\Leftrightarrow\ & CV = R\left(1 - \left(\frac{p'}{p}\right)^{1-\alpha}\right)
\end{aligned}$$

Note that the CV is negative because $p' > p$. The equivalent variation (EV) is the minimal amount of money the consumer would need to be paid in order to accept that the price increase is not implemented. It is defined by:

$$\begin{aligned}
&\tilde{V}(p, R + EV) = \tilde{V}(p', R) \\
\Leftrightarrow\ & ln(R + EV) - (1-\alpha)ln(p) = ln(R) - (1-\alpha)ln(p') \\
\Leftrightarrow\ & EV = R\left(\left(\frac{p}{p'}\right)^{1-\alpha} - 1\right)
\end{aligned}$$

Again, note that since $p' > p$, we have that $EV \leq 0$. We can also use this inequality to show that $|CV| \geq |EV|$. The welfare loss measured by CV is higher than the one measured by EV in the case of a price increase.

Proof: We are going to compare $ln\left(|CV|/R\right)$ and $ln\left(|EV|/R\right)$ which is equivalent to comparing $|CV|$ and $|EV|$.

$$\begin{aligned}
ln\left(\frac{|CV|}{R}\right) - ln\left(\frac{|EV|}{R}\right) &= ln\left(\left(\frac{p'}{p}\right)^{1-\alpha} - 1\right) - ln\left(1 - \left(\frac{p}{p'}\right)^{1-\alpha}\right) \\
&= ln\left(\left(\frac{p'}{p}\right)^{1-\alpha}\right) = (1-\alpha)ln\left(\frac{p'}{p'}\right) \geq 0
\end{aligned}$$

(b) Why do these concepts lead to different measures of the welfare effect of the subsidy? Which one do you consider the most relevant in this context and why?

Answer: The discrepancy between the three measures arise due to the presence of income effects. As the variation of consumer surplus is measuring the effect of the price decrease using the uncompensated demand, it is capturing a mix of the effect of the pure price change and of the associated income effect. The EV and CV are only capturing the substitution effect and are thus an exact measure of the welfare loss. However, the EV and the CV differ because they use difference reference points. The EV is measured in the ex-ante world where prices are lower, while the CV is measured in the ex-post world where prices are higher. The purchasing power of any money transfer is higher in the first case compared to the second, therefore the EV will be lower than the CV here.

Since the government is interested in compensating the consumer from the welfare loss associated with the price increase that already happened, the compensating variation seem to be the most appropriate measure. In particular, if the government transfers CV euros to the consumer it will completely cancel the welfare loss she experienced.

(c) What would be the first best solution to compensate the consumer?

Answer: The first best solution would be to set up a lump-sum transfer to compensate exactly the consumer for her welfare loss (e.g. using the amount of the CV). This would allow the government to reach its goal without creating any deadweight loss (contrary to a gas subsidy that would distort consumption behaviors).

(d) Suppose that in reality, the government faces a heterogeneous population and has redistributive preferences such that it wishes to protect mostly the low income households from the price increase. Why is the first best solution not feasible in practice?

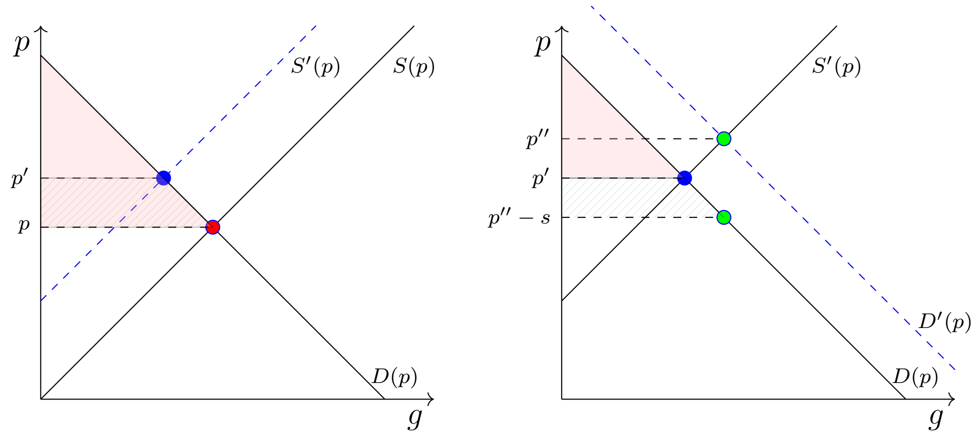
Answer: In reality, lump-sum transfers are impossible to set up because of asymmetric information between consumers and the government. The government does not know the welfare loss and thus the right CV to set for each consumer. The government could set up a means-tested transfer instead but this would distort the incentives to work and thus not be a first best policy anymore.

3. The government introduces instead a subsidy of s euros per kWh on gas consumption, directly remitted to the consumer. Represent the effects of the negative supply shock and then of the introduction of the subsidy on he equilibrium price and quantity of gas in two separate supply and demand diagrams. Illustrate the changes in consumer surplus in both graphs. [NB: You do not need to derive the supply and demand functions but can represent generic demand and supply curves for this question.] (1 pt)

   Answer: In Figure 1.(a), there is an initial equilibrium represented by the red dot with an initial consumer surplus represented by the red area. Following a negative supply shock, the gas price increases from p to p' and the new equilibrium is represented by the blue dot. The consumer faces a surplus loss represented by the shaded area. In Figure 1.(b), the equilibrium now moves to the point represented by the upper green dot following the introduction of the subsidy. The market price for gas increases further to p'' but the consumer only pays $p'' - s$. The consumer surplus increases by the shaded area.

4. Opponents to the subsidy argue that this policy will mostly benefit gas-producing countries in the end instead of local consumers because it will push prices even above p'. To address this question, the government would like to know the incidence of the subsidy. (1.5pt)

   (a) Derive the demand function D in terms of the after-subsidy price $p - s$ and the supply function S in terms of the before-subsidy price p. Compute the price elasticities of demand and supply. (0.5pt)



(a) Effect of the negative supply shock (b) Effect of the subsidy introduction

Answer: The demand function is found by solving the new consumer's problem:

$$\begin{cases} \max_{\{x,g\}} \; U(x,g) = x^{\alpha} g^{1-\alpha} \\ s.t. \;\; x + (p-s)g \leq R \end{cases}$$

The demand is $D(p-s) = (1-\alpha)\frac{R}{p-s}$. The supply function is given by the profit maximization problem of the producer and is shown to be $S(p) = p^{1/2}$. The demand elasticity is:

$$\varepsilon^D = \frac{\partial D(p-s)}{\partial(p-s)} \frac{(p-s)}{D(p-s)} = -1$$

The supply elasticity is:

$$\varepsilon^S = \frac{\partial S(p)}{\partial p} \frac{p}{S(p)} = 0.5$$

(b) Re-derive the formula of the incidence of a subsidy s. What are the two main assumptions needed for this formula to hold? (0.5pt)

Answer: The competitive equilibrium price p is defined by the equality between supply and demand[5]. We differentiate this equation with respect

[5]Note that to start from $D(p-s) = S(p)$, we are assuming perfect competition.

to s in order to derive the effect of the introduction of a subsidy s on price p:

$$\frac{dD(p-s)}{ds} = \frac{dS(p)}{ds} \tag{1}$$

$$\Leftrightarrow \frac{\partial D(p-s)}{\partial p}\frac{dp}{ds} + \frac{\partial D(p-s)}{\partial s} = \frac{\partial S(p)}{dp}\frac{dp}{ds} \tag{2}$$

$$\Leftrightarrow \frac{\partial D(p-s)}{\partial (p-s)}\left(\frac{dp}{ds} - 1\right) = \frac{\partial S(p)}{dp}\frac{dp}{ds} \tag{3}$$

$$\Leftrightarrow \frac{dp}{ds} = \frac{\frac{\partial D(p-s)}{\partial (p-s)}}{\frac{\partial D(p-s)}{\partial (p-s)} - \frac{\partial S(p)}{\partial p}} \tag{4}$$

$$\Leftrightarrow \frac{dp}{ds} = \frac{\frac{\partial D(p-s)}{\partial (p-s)}\frac{(p-s)}{\partial Q}}{\frac{\partial D(p-s)}{\partial (p-s)}\frac{(p-s)}{\partial Q} - \frac{\partial S(p)}{\partial p}\frac{(p-s)}{\partial Q}} \tag{5}$$

$$\Leftrightarrow \frac{dp}{ds} \simeq \frac{\frac{\partial D(p-s)}{\partial (p-s)}\frac{(p-s)}{\partial Q}}{\frac{\partial D(p-s)}{\partial (p-s)}\frac{(p-s)}{\partial Q} - \frac{\partial S(p)}{\partial p}\frac{p}{\partial Q}} \tag{6}$$

$$\Leftrightarrow \frac{dp}{ds} \simeq \frac{\varepsilon^D}{\varepsilon^D - \varepsilon^S} \tag{7}$$

To go from (2) to (3) we need to assume that consumers react the same to a change in pre-subsidy price than to a change in post-subsidy price so that $\partial D/\partial p = \partial D/\partial(p-s) = \partial D/\partial s$. This means that subsidies are perfectly salient to the consumer. The other assumption we need is to assume that the subsidy introduced is small to go from (5) to (6). If s is small, then $p - s \simeq p$ and we find back the classical incidence formula. Note that $Q = S(p^*) = D(p^* - s)$ is the equilibrium gas quantity.

(c) What share of the subsidy will go to the consumer? In the context of the European energy crisis, where there are multiple countries demanding gas, why could the government of one country still argue that an energy subsidy will benefit its citizen? (0.5pt)

Answer: Plugging the value of the supply and demand elasticities in the incidence formula, we find that following the introduction of a small subsidy of 1 euro per kWh, the market price of gas would go up by 2/3 euro per kWh. We can also show that $\partial(p-s)/\partial s = \partial p/\partial s - 1$ so the price faced by the consumer would only go down by 1/3 euro per kWh. The incidence falls mostly on the gas producers which are less elastic. Here, we have been reasoning using a representative consumer. In reality, there are many consumers in different countries. A country in which the population represents a sufficiently small portion of the overall gas demand might be tempted to implement a subsidy because it would not affect the market too much, i.e. the market price should not go up much. However, each country has an incentive to do the same and the incidence would then still fall on the producers.

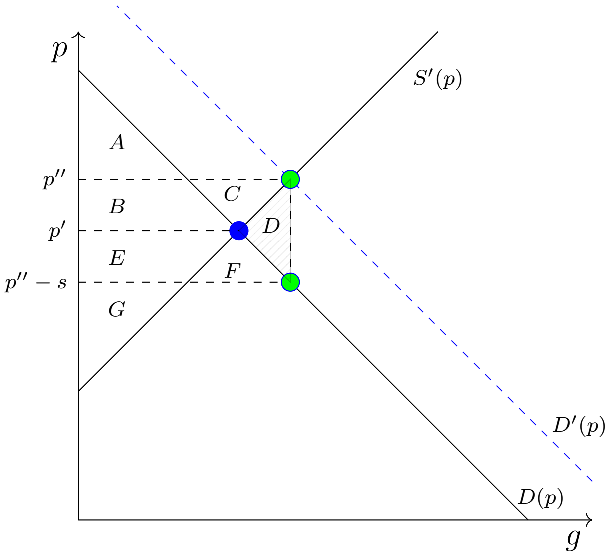
5. The government is also worried about the inefficiency that this subsidy could introduce in the economy. Represent the deadweight loss associated with the

introduction of a subsidy using a supply and demand diagram as in question 3. (1pt)

Answer: The shaded area in the figure below represents the deadweight loss associated with the introduction of the subsidy. The subsidy distorts behaviors and generates over-consumption of gas. The variation in total social surplus is represented in the figure by the following areas:

$$\begin{aligned}\Delta \text{ Social surplus} &= \Delta \text{ Consumer Surplus} + \Delta \text{ Producer Surplus} + \Delta \text{ Government Revenue}\\ &= ((A+B+E+F)-(A+B)) + ((G+E+B+C)-(E+G))\\ &\quad + (-(B+C+D+E+F)-0)\\ &= D\end{aligned}$$

We can see graphically that the deadweight loss is measured by $s\frac{dg}{2}$.



(a) *(Bonus question)* Derive the deadweight loss in terms of s and p using Harberger's formula.

Answer: Following the introduction of a small subsidy s, the deadweight loss is given by:

$$\begin{aligned} DWL &= -\frac{1}{2}dQ \times s && (8)\\ &= -\frac{1}{2}\frac{\partial S(p)}{\partial p}dp \times s && (9)\\ &= -\frac{1}{2}\frac{\partial S(p)}{\partial p}\frac{\varepsilon^D}{\varepsilon^D-\varepsilon^S} \times s^2 && (10)\\ &= -\frac{1}{2}\varepsilon^S\frac{S(p)}{p}\frac{\varepsilon^D}{\varepsilon^D-\varepsilon^S} \times s^2 && (11)\\ &= -\frac{1}{2}pQ\frac{\varepsilon^D\varepsilon^S}{\varepsilon^D-\varepsilon^S} \times \left(\frac{s}{p}\right)^2 && (12)\end{aligned}$$

In the gas subsidy case studied here, the deadweight loss is $-\frac{1}{6}p^{3/2}\left(\frac{s}{p}\right)^2$.

(b) *(Bonus question)* Suppose that the government does not have access to reliable estimates of the supply and demand elasticities for gas. What alternative formula of the deadweight loss could the government use? How you would estimate empirically the sufficient statistics needed to assess the deadweight loss using this formula?

Answer: An alternative expression for the deadweight loss is:

$$DWL = \frac{1}{2}\varepsilon^{Q} pQ \times \left(\frac{s}{p}\right)^{2}$$

where $\varepsilon^{Q} = -\frac{\partial Q}{\partial s}\frac{p}{Q}$. As seen in the lecture, this elasticity could be estimated using data on gas quantities traded, market price and gas subsidy. However, it would require an exogenous change in the subsidy amount. Here for example, as the subsidy is introduced at a time where other unobserved factors were affecting the gas (e.g. war), using simple time series data and OLS would lead to a bias. A possible solution would be to design a randomized experiment and estimate this elasticity by a difference in difference method.

6. In real life, consumption taxes and subsidies are not always fully salient to consumers, especially when they do not bear the statutory incidence. Discuss the possible effects of a reform that would shift the statutory incidence of the energy subsidy from the consumer to the producer (i.e. remit the subsidy to the producer). (1pt)

   Answer: If remitting the subsidy to the producer makes it less salient to the consumer, it will reduce the elasticity of demand with respect to the subsidy. This seems like a good policy here as it would both reduce the deadweight loss associated with the subsidy and shift the incidence a bit more towards the consumer.

7. Using Californian data, Auffhammer and Rubin (NBER WP, 2018) estimate a price elasticity of gas consumption of -0.21. They show that there is substantial heterogeneity across seasons, with demand being much more sensitive in the winter than in the summer. Moreover, low-income households (who are eligible to subsidized gas prices) are more elastic than other households in the winter and less elastic in the summer. Based on this observation, how could such a subsidy program be re-designed in a way that is both more efficient and more redistributive? (1pt)

   Answer: The government faces a classical equity-efficiency trade-off. It wishes to redistribute to low-income households through a subsidy but such a redistribution will create a deadweight loss. A gas subsidy that would only be provided in the summer (instead of a permanent subsidy over the year) could improve both equity and efficiency. Indeed, if gas consumption is almost inelastic in the summer, the deadweight loss would be much smaller. Moreover, since the low-income households are even less elastic than the others, the incidence of the subsidy would become more progressive by shifting towards these households.