Redistribution

Course note to accompany Lecture 11 of Elements of Economics II

E. Glen Weyl
University of Chicago

Fall 2011-2

1 Framework

The general problem of optimal income taxation is quite complex. Individuals earn income over the course of a lifetime, have income and substitution effects in their choice of working time, may evade taxes, etc. In order to focus on the core issues and understand the basic trade-offs clearly, we will abstract from many of these features. In particular,

1. We will assume that all income earning occurs in one go.

2. We will assume that there are no income effects. In particular, we will ignore the fact that, given this specification, when consumers are given money by the government, they will reduce the amount they work typically. While ignoring income effects may seem strange (and inconsistent), as they are clearly an important part of these decisions, Saez (2001) shows that the effect these have only a small effect on the analysis.

3. We will make some stark, but realistic, assumptions about evaluating social welfare. We will assume that we care about the sum of all consumers utilities; this is justified primarily by Harsanyi's "Veil of Ignorance" argument. We will assume that all consumers have logarithmic utility of income. This is mostly justified by Sacks et al. (2011)'s data on the relationship between the logarithm of income and reported subjective well-being.

4. Finally, we will explore a limited set of policy tools. Effectively we will assume that taxes are flat over most of the range and then that there is a (linear) surcharge on incomes at the top end of the spectrum. All revenues earned in this manner are then given back as a lump sum on a per-capita basis to all consumers. This is a proxy for a range of government services as well as direct transfers.

While income tax systems in the real world are more complex (with several brackets that gradually become more progressive) the basic intuition of how these more complex systems should work can be seen more easily in this simplified setting. In particular, the linear part of the tax is a good proxy for the average marginal tax rate most people should face, while the top tax rate is good proxy for what the super-rich should pay in marginal rates.

5. We will make some (fairly well-justified, but especially useful) parametric assumptions about the distribution of income.

This simplified model brings out the fundamental trade-off in optimal tax theory. On the one hand, because logarithmic utility is concave (because people are risk-averse) we would like to redistribute income from the rich to the poor, who can better use that income. On the other hand, raising taxes to do this can be self-defeating as the more heavily income is taxed, the less incentive individuals have to work for the income that can then be redistributed. That is, you don't want to kill the goose that lays the golden egg. This trade-off is typically known as the "equity-efficiency" trade-off or the "Laffer curve".

The trade-off is very similar to the monopoly problem: raising prices raises profits, but eventually reduces them as consumers reduce their purchases. The only differences here are the government should take into account the utility of those taxed, at least to the extent that they are not super-rich, and when charging the super-rich we must take into account the fact that the surcharge only hits their income above some level, but has the full impact on their marginal incentives. This makes charging them extra somewhat less attractive as it effectively raises the elasticity of their taxed income. The rest of this note formally explores these issues.

2 Linear taxation

If a linear income tax is imposed, individuals after tax income $i = (1-t)I + t\overline{I}$ where $\overline{I}$ is the average income in the population, $\overline{I} = \int_{I=0}^{\infty} If(I)dI$ where f is the density of pre-tax incomes. The reason is that each individual pays a fraction t of their income, but receives back a fraction t of the average income.

However, pre-tax incomes also fall when tax rates rise. This happens for two reasons:

1. *Substitution effect*: individuals now get less money for each hour they work, encouraging them to work less. Suppose we know the elasticity of individuals' labor supply with respect to their wage, ϵ. Let's call the average value of this elasticity for individuals with pre-tax income I, $\overline{\epsilon}(I)$. How much of a reduction in wage is an increase in t by one unit equivalent to? The individual's after-tax wage is effective $w(1-t)$ so we must calculate

$$\frac{dw(1-t)}{dt}\frac{1}{w(1-t)} = -\frac{w}{w(1-t)} = -\frac{1}{1-t}.$$

Thus when t increases by one percentage point, the average income of individuals currently earning pre-tax income I will fall by $-\frac{100\overline{\epsilon}(I)}{1-t}$ percent. The total reduction percentage reduction in average income this causes is the average value in the population of this fall, weighted by income:

$$\frac{d\overline{I}}{dt}\frac{1}{\overline{I}} = \frac{\int_{I=0}^{\infty} I \frac{\overline{\epsilon}(I)}{1-t} f(I) dI}{\int_{I=0}^{\infty} I f(I) dI} \equiv \frac{\overline{\epsilon}}{1-t}.$$

Thus a one percentage point increase in taxes cases a $\frac{100\overline{\epsilon}}{1-t}$ percent fall in average pre-tax income.

2. *Income effect*: individuals now receive a larger lump-sum transfer. This causes them to work less hard. As mentioned before, we ignore this effect from now on, but note that it is being left out.

Social welfare is the average utility of income; we can ignore welfare from leisure by the envelope theorem (if that doesn't mean anything to you, just ignore this comment and go with the flow). This welfare is

$$\int_{I=0}^{\infty} \log\left((1-t)I + t\overline{I}\right) f(I) dI.$$

Taking the derivative with respect to t yields

$$\int_{I=0}^{\infty} \frac{\overbrace{\overline{I}\left(1-\frac{t\overline{\epsilon}}{1-t}\right)}^{\text{value of marginal tax revenue}} - \overbrace{I}^{\text{welfare cost to taxpayers}}}{(1-t)I + t\overline{I}} f(I) dI = \left(1-\frac{t\overline{\epsilon}}{1-t}\right)\overline{I}\overline{w} - \overline{Iw},$$

where w denotes the *social welfare weight* on an individual (their marginal utility of income) and $\overline{x}$ denotes the average value of x. Rearranging this expression and noting that $\overline{Iw} - \overline{I}\overline{w} = \text{Cov}[I, w]$ yields

$$-\frac{t\overline{\epsilon}}{1-t}\overline{I}\overline{w} - \text{Cov}[I, w].$$

Setting this to 0 and solving out gives the optimal tax rate as

$$t^\star = \frac{\sigma(t)}{\overline{\epsilon} + \sigma(t)}, \tag{1}$$

where $\sigma(t) \equiv -Cov\left[\frac{I}{\overline{I}}, \frac{w}{\overline{w}}\right]$. The covariance is always negative, as wealthier citizens receive lower welfare weights as utility is concave. As it becomes large, the optimal tax rate goes to 1; as it becomes small (as when there is no inequality), the optimal tax rate goes to zero. This is quite intuitive: the more unequal is society, the more negative covariance there is between the marginal utility of money and how much money an individual has. Inequality makes redistribution desirable. The tax also falls with the average elasticity of labor supply: the more elastic is labor supply, the less it makes sense to tax labor, as doing so kills the goose that lays the golden egg. Thus we see an elegant expression of the equity-efficiency trade-off. We will go further in calibrating this expression in Section 4.

3 Top tax rates

This formula gives us a good sense of what determines tax rates for most citizens. But President Obama has recently been pushing for an extra tax on the super-rich, those earning more than a million dollars a year. The useful thing, from an analytic stand point, about such individuals is that we can largely disregard the welfare costs of redistributing money from them. The marginal utility of someone earning a million dollars a year is about one thirtieth of that of a typical American. Thus it is only the reduction in revenue we need to take into account. This significantly simplifies the analysis.

Let's consider a similar procedure to the one we followed in the previous section, except now a) the tax considered hits only incomes above \$1 million and b) we disregard any reduction in the welfare to those earners. Let's call the tax rate on those earning above \$1 million t_{10^6}. If we increase this rate by one percentage point again we have two effects:

1. Mechanical revenue effect: the revenue earned from these individuals will, mechanically, increase by the amount of income they earn above this level $I - 10^6$.

2. Labor supply reduction: the total income of an average individual earning $I > 10^6$ will fall by $\frac{I\overline{\epsilon}(I)}{1-t}$. Given that the marginal tax rate faced by this individual is t, this will reduce tax collected by from them by $\frac{tI\overline{\epsilon}(I)}{1-t}$.

Because we do not care about the marginal welfare reduction from taking money away from these super-rich, we only need consider these two effects. Again, the optimal value of t_{10^6} balances these two effects averaged over all individuals earning:

$$\int_{10^6}^{\infty} \left(I \left[1 - \frac{t^\star_{10^6}\overline{\epsilon}(I)}{1 - t^\star_{10^6}} \right] - 10^6 \right) f(I)dI = 0 \iff \int_{10^6}^{\infty} I \left[1 - \frac{t^\star_{10^6}\overline{\epsilon}(I)}{1 - t^\star_{10^6}} \right] f(I)dI = 10^6 \int_{10^6}^{\infty} f(I)dI.$$

If we divide both sides by $\int_{10^6}^{\infty} f(I)dI$ and let $\overline{I}_{10^6} \equiv \frac{\int_{10^6}^{\infty} If(I)dI}{\int_{10^6}^{\infty} f(I)dI}$ and $\overline{\epsilon}_{10^6} \equiv \frac{\int_{10^6}^{\infty} I\overline{\epsilon}(I)f(I)dI}{\int_{10^6}^{\infty} If(I)dI}$ we obtain

$$\overline{I}_{10^6} - 10^6 = \overline{I}_{10^6}\overline{\epsilon}_{10^6}\frac{t^\star_{10^6}}{1 - t^\star_{10^6}}$$

or

$$t^\star_{10^6} = \frac{\frac{\overline{I}_{10^6} - 10^6}{\overline{I}_{10^6}}}{\overline{\epsilon}_{10^6} + \frac{\overline{I}_{10^6} - 10^6}{\overline{I}_{10^6}}}. \tag{2}$$

Note how similar this is to equation 1: $\overline{\epsilon}_{10^6}$ replaces $\overline{\epsilon}$ and $\frac{\overline{I}_{10^6}-10^6}{\overline{I}_{10^6}}$ replaces $\sigma(t)$. Thus it will be optimal to have a surcharge on the rich to the extent that their (income-weighted) average elasticity of labor supply is lower and to the extent that $\frac{\overline{I}_{10^6}-10^6}{\overline{I}_{10^6}} > \sigma(t)$ and it will be optimal to charge them a lower marginal rate in the reverse case. We now turn to the empirical calculation of these quantities to figure out which case the real world lies in.

4 Empirical calibration

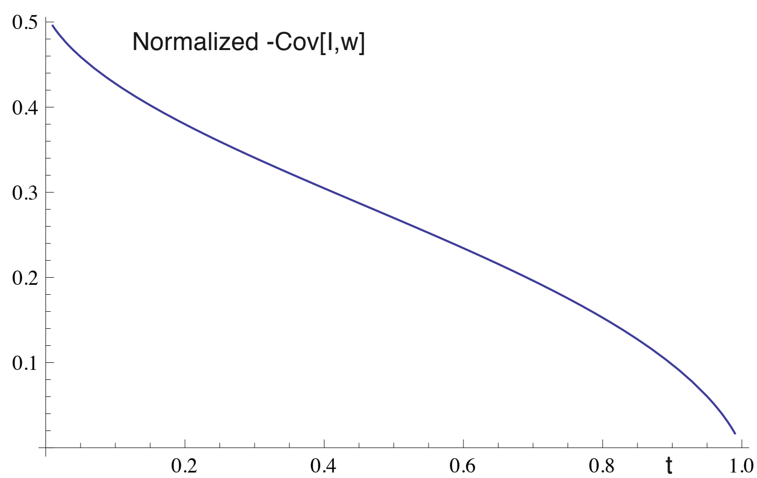

Figure 1: Negative covariance between social welfare weights and income, as a function of the tax rate.

Both $\sigma(t)$ and $\frac{\overline{I}_{10^6}-10^6}{\overline{I}_{10^6}}$ are relatively straight-forward to calculate empirically, given available data on the distribution of income. First consider

$$\sigma(t) = 1 - \frac{\int_{I=0}^{\infty} \frac{I}{(1-t)I+t\overline{I}} f(I)dI}{\overline{I}\int_{I=0}^{\infty} \frac{1}{(1-t)I+t\overline{I}} f(I)dI}.$$

We can calibrate $\overline{I}$ immediately from US GDP per capita which is approximately \$46,000. To calculate the other quantities, it is useful to a log-logistic distribution of income; this is very similar to a log-normal distribution, but even simpler to work with in this context. The log-logistic distribution has density function:

$$f(I) = \frac{\beta\left(\frac{I}{\alpha}\right)^{\beta-1}}{\alpha\left[1+\left(\frac{I}{\alpha}\right)^{\beta}\right]^2},$$

where $\beta = \frac{1}{G}$, G is the country's Gini coefficient and mean incomes is $\frac{\alpha\pi}{\beta\sin\left(\frac{\pi}{\beta}\right)}$. For the United

States, as we found in Problem Set 3, the Gini is .45. US mean incomes is $46,000$ so we can solve out for α:

$$\frac{.45\alpha\pi}{\sin\left(.45\pi\right)} = 46,000 \implies \alpha \approx 32,168.$$

Plugging into Mathematica allows us to numerically compute the above integrals. Figure 1 shows a graph of $\sigma(t)$ as a function of t. As suspected, it declines as t increases, as eventually there is no inequality and thus no variance or covariance. We can then determine optimal tax rates by graphing t against $\frac{\sigma(t)}{\overline{\epsilon}+\sigma(t)}$ for various assumptions about $\overline{\epsilon}$. Figure 2 shows that with $\overline{\epsilon} = .1, .3, .5, 1$, $t^\star \approx .67, .48, .38$ and $.26$ respectively. Thus the elasticity makes a big difference in the optimal tax rate. Kotlikoff and Rapson (2008) find that average marginal tax rates in the US population, taking into account taxation at all levels of government, are approximately .4. Thus tax rates should be, on average, higher if the average elasticity is much below .5 and should be lower if the elasticity is much below this level.

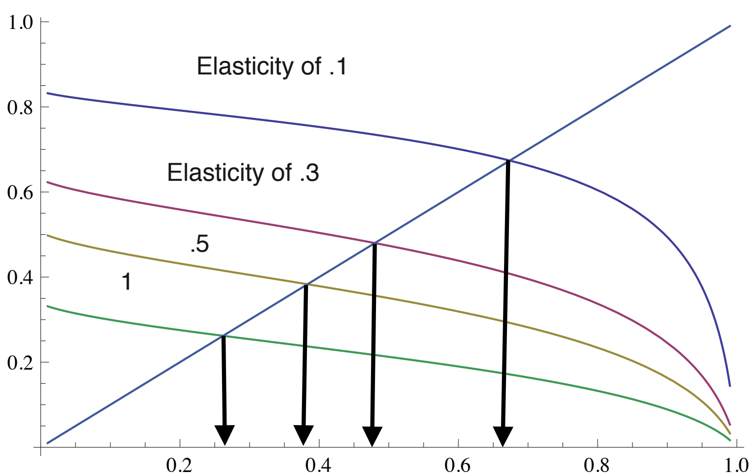

Figure 2: Optimal linear tax rates for various assumed levels of the average elasticity.

Calibrating $\frac{\overline{I}_{10^6}-10^6}{\overline{I}_{10^6}}$ is even easier. The upper part of the income distribution is well-known to follow a Pareto distribution with a coefficient $\alpha \approx 1.5$ (Diamond and Saez, Forthcoming). We can thus calculate

$$\frac{\int_{10^6}^{\infty} \frac{\alpha}{\underline{I}} I \left(\frac{I}{\underline{I}}\right)^{-\alpha-1} dI}{\int_{10^6}^{\infty} \frac{\alpha}{\underline{I}} \left(\frac{I}{\underline{I}}\right)^{-\alpha-1} dI} = \frac{\alpha \underline{I}^{\alpha+1}\int_{10^6}^{\infty} I^{-\alpha} dI}{\alpha\underline{I}^{\alpha+1}\int_{10^6}^{\infty} I^{-\alpha-1}dI} = \frac{-\frac{1}{\alpha-1}\, I^{-(\alpha-1)}\Big|_{10^6}^{\infty}}{-\frac{1}{\alpha}\, I^{-\alpha}\Big|_{10^6}^{\infty}} = \frac{\frac{10^{-6(\alpha-1)}}{\alpha-1}}{\frac{10^{-6\alpha}}{\alpha}} = 10^6\frac{\alpha}{\alpha-1}$$

Thus $\frac{\bar{I}_{10^6}-10^6}{\bar{I}_{10^6}} = \frac{\frac{\alpha}{\alpha-1}-1}{\frac{\alpha}{\alpha-1}} = 1 - \frac{\alpha-1}{\alpha} = \frac{1}{\alpha}$ and, assuming $\alpha \approx 1.5$, $\frac{\bar{I}_{10^6}-10^6}{\bar{I}_{10^6}} \approx \frac{2}{3}$. Thus the optimal tax rate for earners above $1,000,000 is approximately

$$t^{\star}_{10^6} = \frac{\frac{2}{3}}{\bar{\epsilon}_{10^6} + \frac{2}{3}} = \frac{1}{1.5\bar{\epsilon}_{10^6} + 1}.$$

If we again plug $\bar{\epsilon}_{10^6} = .1, .3, .5, 1$ into this we obtain $t^{\star}_{10^6} = .87, .69, .57, .4$. Note that these higher than the corresponding linear tax numbers for the same elasticity, leading us to the conclusion that if elasticity of labor supply is on average similar for the wealthy as it is for average Americans, we should have a progressive system in which rates facing top earners are significantly higher than those facing the average American.

Do the wealthy have similar elasticities of labor supply to other Americans? On the one hand, it seems unlikely that the wealthy respond to a percent change in their wages by changing their labor supply nearly as much (at least in terms of *how much* they work, as opposed to where or whether they work) a typical American does. On the other hand, the wealthy have much greater opportunities to hide their income than do most Americans and thus their effective labor supply elasticity, as observed by the tax system, may be higher.

This intuition are confirmed by Saez et al. (Forthcoming), who survey all available economic studies on the topic. They find that labor supply elasticities are very small, less than .1, at the top end of the income spectrum but are somewhat higher (perhaps around .5) in the mid-range primarily because of the adjustment of working time by women in married couples. However, elasticities of *taxable income* is much higher due to avoidance and individuals taking advantage of various deductions. Nonetheless, the authors argue that it is really the labor supply elasticity that is relevant because other sources of elasticity don't really count:

1. To the extent that elasticity comes from simple avoidance, this can be fixed by tightening tax rules and eliminating silly loopholes and deductions.

2. To the extent they come from deductions from charitable contributions, this is not a cost of higher taxes. In fact, it is a benefit as it raises charitable contributions.

3. To the extent that these come from shifting tax from year to year, these are not relevant to long-run calculations of the sort we are doing here.

Thus, based on our calculations, it seems that if we can eliminate silly deductions and loopholes, and if we accept the framework adopted here, optimal top tax rates are very high indeed, likely much higher than even President Obama is suggesting. Optimal tax rates on average Americans seem likely to be quite high as well (around 40%). These would fund large transfers to the less well-off, which would likely raise, rather than lower, the vast majority of incomes as median income ($\approx \$26k$) is much lower than mean income ($\approx \$46k$). Roughly, under such a system, about two-fifths of the wealth held by the top half of Americans, who do not earn more than a million dollars a year, would be redistributed to the bottom half and about 80% of income above a million dollars would be redistributed. Hopefully this will cause you to reflect on whether you think that the income tax system is not nearly progressive enough or whether you think the framework we have used (in which our goal is to redistribute as much as possible, but we are restrained by the deadweight loss thus generated) is flawed.

References

Diamond, Peter and Emmanuel Saez, "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, Forthcoming.

Kotlikoff, Lawrence and David Rapson, "Does it Pay, at the Margin, to Work and Save," in "Tax Policy and the Economy," Vol. 21 of *National Bureau of Economic Research Volume*, MIT Press, 2008.

Sacks, Daniel W., Betsey Stevenson, and Justin Wolfers, "Subjective Well-Being, Income, Economic Development and Growth," in "Development Challenges in a Post-Crisis World," World Bank, 2011.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *The Review of Economic Studies*, 2001, *68* (1), 205–229.

_ , Joel B. Slemrod, and Seth H. Giertz, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, Forthcoming.