**Universität Ulm**
**Fakultät für Mathematik und Wirtschaftswissenschaften**

# Optimal Taxation of Top Labor Incomes: An Interactive R Problem Set

Bachelorarbeit

in Wirtschaftswissenschaften

vorgelegt von
Jonas Send
(Matrikelnummer 760408)
am 08.01.2015

**Gutachter**

Prof. Dr. Sebastian Kranz

# Abstract

In this thesis I present the interactive *R* problem set "Optimal Taxation of Top Labor Incomes" that I have developed and that is based on the paper "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities" by Piketty, Saez and Stantcheva (2014). Using a detailed illustration of the problem set I describe its mechanisms and discuss its purpose. I also include explanations of my approach to creating certain parts of the problem set which is intended to help future developers. Concluding, I offer a brief discussion on the paper of Piketty, Saez and Stantcheva (2014) and define the academic contribution of my problem set.

# Contents

# Introduction

Economics can be defined as "the study of how society manages its scarce resources" (Mankiw (2014)). But it is not possible to perform experiments with a whole society. Hence, in economics, one is often dependent on the data that reality provides, for example via natural experiments. And in contrast to data that is collected in a laboratory, this makes it more difficult to control for external influences that are not observed and other effects that can distort results. This leads to a high importance of statistical know-how in the field of economics.
One statistical software package that helps economists tackle their statistical challenges is the open source statistical computing environment and language *R*[1] that is developed by the R Core Team (2014). To create the opportunity for students who are interested in *R* to learn how to use this statistical software and at the same time learn about economic issues, I have used the *R* package *RTutor*[2] from Kranz (2014a) to develop the interactive R problem set "Optimal Taxation of Top Labor Incomes". *RTutor* provides the tools to write an *R Markdown* file that can run an interactive offline *html* tutorial where the user is able to execute and test *R* code. My problem set enables the user to gradually reproduce the study "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities" by Thomas Piketty, Emmanuel Saez and Stefanie Stantcheva (2014)[3]. In their paper they analyse how top incomes respond to top marginal tax rates and derive optimal top tax rate formulas in a theoretical model. I chose this paper because I think that it treats a very current topic that possesses the ability to fascinate the user and because its data provides the opportunity to introduce various different tools that *R* supplies.

The problem set consists of the *R Markdown* file "Optimal Taxation of Top Labor Incomes_sol", the two data sets that are used in the tutorial, "usdata.csv" and "intdata.dta" and the text file "Optimal_Top_Labor_Taxation_Var" that contains descriptions for the variables that are used in the problem set. Running the first code chunk of the "Optimal Taxation of Top Labor Incomes_sol.Rmd" file in an *R* environment creates all the required additional files and starts the *html* offline tutorial in the default browser[4].

The main body of this thesis is the next chapter, "Problem Set Optimal Taxation of Top Labor Incomes". It is an illustration of the problem set and includes all the output that is generated while working through it[5]. While it follows the exact structure of the problem set, I added comments and further explanations. Everything that is part of the problem set as it can be viewed in a browser is framed to make it possible to easily distinguish it from commentary.
The chapter "Critical Appraisal of the Paper" contains a short analysis of Piketty et al.'s paper and names some of its propositions that might be controversial.
In the chapter "Academic Value Added" I summarize the purpose of the problem set and discuss the academic value it provides.
In the last chapter, the "Conclusion", I consider the results of my work on the tutorial and work out some possible implications for further developments of interactive R problem sets.

[1] You can find more information about *R* and a download link on http://www.r-project.org/.

[2] The package can be downloaded from *Github* on https://github.com/skranz/RTutor. You can also find helpful guides on this website that explain how to develop a problem set.

[3] Hereinafter I will refer to this article just as Piketty et al.

[4] Note that the problem set can also be solved directly in *RStudio*. However, in this thesis I will concentrate on the problem set as it is presented in a browser based interface because I mainly designed it for that use.

[5] I generated the illustration of the problem set by compiling the file "Optimal Taxation of Top Labor Incomes_output_solution.Rmd", that is also created by running the "Optimal Taxation of Top Labor Incomes_sol.Rmd" file, into *LaTeX* output. This was done by using the "Knit PDF" function that is included in *RStudio*. Note that I have manipulated the raw *LaTeX* output to some degree to get an optically more appealing result (For instance, I shortened the line lengths of some code chunks to less than 80 to make them fit).

## Problem Set Optimal Taxation of Top Labor Incomes

Name: Jane Doe

Author: Jonas Send

In their paper **"Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities"** Thomas Piketty, Emmanuel Saez and Stefanie Stantcheva (2014) analyse the response of top earners to taxes and derive optimal tax rate formulas from their findings. In this interactive offline *R* Tutorial, we are going to gradually reproduce their study and discuss it.
(The article and the corresponding public data is provided on the website of the *American Economic Association*. You can click on this link to download it.)

As it is noted at the beginning of the problem set, I sourced the article and the public data that is used for its statistical analysis from the website of the *American Economic Association*[6]. The data for the article consists of two sets of macroeconomic evidence for both the US and a group of OECD countries and of two sets of microeconomic evidence for the US and a group of OECD countries. However, the microeconomic data is confidential and not freely available. This is why the analysis that Piketty et al. conduct with this data is only shortly summarized in *Exercise 2.2* of the problem set.

### Exercise Overview

The main focus of the paper is to examine the channels through which top earners react to changes in top net-of-tax rates. Top net-of-tax rate is defined as $1 - r$, with $r$ being the top marginal tax rate on ordinary income. As often in economics, when we look at the response of people to external influences, we are going to talk about elasticities. In our case, we are going to look at three different elasticities:

- $e_1$ describes the **supply-side response** of top earners to raises of the top net-of-tax rate ("How much does real income of the top 1 percent increase in percent if the top net-of-tax rate increases by one percent?")
- $e_2$ describes **tax-avoidance responses** (changes in reported income with no influence on the real income due to changes of the top net-of-tax rate)
- $e_3$ describes **compensation-bargaining responses** (changes in taxable income due to a different level of influence top earners wield on compensation committees as a response to changes in the top net-of-tax rate)

These three elasticities combine to the overall elasticity $e$ of the taxable income. $e$ describes how much taxable top 1 percent income increases if the top net-of-tax rate rises.

For example, if the top net-of-tax rate $1 - r$ increases by $1\%$ (due to a decrease of the top marginal tax rate $r$) and the top 1 percent income increases by $0.7\%$, $e$ can be calculated as:

[6] The link to the article that is also included in the problem set via a linked phrase is https://www.aeaweb.org/articles.php?doi=10.1257/pol.6.1.230.

$$e \;=\; \frac{\dfrac{\Delta \; top \; 1 \; percent \; income}{top \; 1 \; percent \; income}}{\dfrac{\Delta \; (1-r)}{(1-r)}} \;=\; \frac{0.7\%}{1\%} \;=\; 0.7 \;.$$

Note: It is important to know the influence of the single elasticities on the overall elasticity in order to design an optimal top tax rate. For example, if the only important elasticity is $e_3$, tax rates should be close to 100% as bargaining effects only come at the expense of other income classes.

**This problem set has the following structure:**

- **Exercise 1 US evidence**
  In this chapter we take a look at empirical evidence from the US that gives us an understanding of how top tax rates and top income shares and totals are connected. This will also help us to get a first idea of the magnitude of the different mentioned elasticities. While doing so, we will learn about different tools in *R* that help us to analyse data.
- **Exercise 2 International evidence**
  Here we analyse data from various OECD countries to gather further insights into top income elasticities and optimal top tax rates. This will help us to tackle some problems that arise in *Exercise 1* and to support our findings. We will also deepen our *R* skills and look at other smart solutions for common data issues.
- **Exercise 3 Theory**
  In this exercise we consider the theoretical model that Piketty et al. develop in their paper and use it to derive a R function that will help us to calculate possible optimal top tax rates.
- **Exercise 4 Scenario Analysis**
  This exercise gives us the opportunity to look at different possible scenarios that describe top income responses and to deduce corresponding optimal top tax rates using the function that we developed in *Exercise 3.*
- **Exercise 5 Conclusion**
  This final exercise gives us the opportunity to sum up our analysis and discuss possible tax policy implications.

*Note that this is the recommended order for this problem set. It is possible to skip exercises or to start with the theory exercise before working on the evidence based exercises. However, if you are new to 'R', bear in mind that 'Exercise 2' builds on the skills acquired in 'Exercise 1'. Additionally, you have to complete 'Exercise 3.2' to be able to work on 'Exercise 4'*

Whereas Piketty el al. chose to familiarise the audience with their model straight after the introductory chapter, I decided to start my problem set with a graphical analysis of the first data set that contains various macroeconomic variables for the US between 1913 and 2008. This is to spark the interest of the user by giving him the opportunity to test first *R* functions that result in expressive plots. In order to be able to already hint to some theoretical aspects of the model that is discussed in *Exercise 3.1*, I put a short explanation of the three elasticities of top incomes into the introduction of the problem set.
After that, an overview of the structure of the problem set gives the user a first orientation. As

it is noted in the tutorial, by design it is possible to skip some exercises and slightly mix up the order. However, it is pointed out that it is recommended to work through the problem set chronologically since I believe this is the most educational way, especially for *R* beginners. The tasks in *Exercise 2* build on the skills that are acquired in *Exercise 1*. The model in *Exercise 3* is probably more comprehensible if you can already assess the different elasticities after having worked through the economic evidence. Finally, it is only possible to work on *Exercise 4* after finishing *Exercise 3.2* and the conclusion in *Exercise 5* naturally only makes sense at the end of the problem set.

### Exercise 1.1 US evidence - Graphical approach

In this exercise we are going to analyse macroeconomic data from the US for the past century by creating plots that give us a first idea of the relationship between top 1 percent income and top marginal tax rates.

#### a) Loading the required data

In order to be able to translate the data into some expressive graphics, we have to load the data into our work space. In the directory of your problem set, you should have saved the file *usdata.csv*. Use the command `read.table()` to load the data and assign it to a variable called `US.data`. Type your code into the open window below. To continue, click the *check*-button to find out whether your solution is correct. If you need help, click the *hint*-button to get some clues on how to solve the problem. If you want more information on the `read.table()`-command you can click the *info*-button below.

**Info: read.table()**
`data = read.table("tablename")` loads the data with the name *tablename* and assigns it to a data frame `data`. You have to use quotation marks for the name of the data file.

Note: Whenever you are working in an *R* environment and want to find out more about a function or a package `name` you can use `?name` or `help(name)` to display documentation and receive further information. As this is not working in this html problem set, there will always be a short introduction into new commands and packages that we utilize.

Throughout the problem set, additional information is given in so-called *info* boxes. These are boxes that unfold when you click on the title and can also be folded again. The reason for including *info* sections is that they provide the opportunity to give further information while at the same time keeping the problem set as short as possible. In this thesis, I illustrate all *info* boxes by putting them into an additional frame.

```
US.data = read.table("usdata.csv")
```

This is the first code chunk of the problem set. The code can be entered by the user and can then be checked by clicking on the *check* button. If the user has problems with solving the code

chunk, it is always possible to click on the so-called *hint* button to receive further information to help him. These help messages are also automatically generated by *RTutor*, but in most cases it makes sense to write them manually. Note that in this illustration, the hint messages are not included but for some code chunks I state them in the commentary to demonstrate how they can be structured. For example, for this code chunk the hint message reads: "'data = read.table("tablename")' is the basic construction of the command. You now only have to change the names of your variable and of the data file. And remember to use quotation marks for the file name.".
If the user has entered the correct code, the code chunk is evaluated as it would be the case in a standard *R* environment. In this case the .csv file "usdata" is loaded into the work space and assigned to the variable `US.data`. Note that this data file is not the raw data that is provided by Piketty et al. but that I have manipulated the data according to Piketty et al.'s *Stata* documentation to make it possible for the user to immediately start working on it[7].

> The data set contains different macroeconomic values for the years 1913 to 2008 such as nominal GDP, marginal tax rates for top incomes and income shares of the top 1 percent for the United States of America. If you want to take a look at the data that you just loaded, click on the button *data* you can find above. If you are interested in more detailed information of the variables we are going to work with now, see the *info* section below.

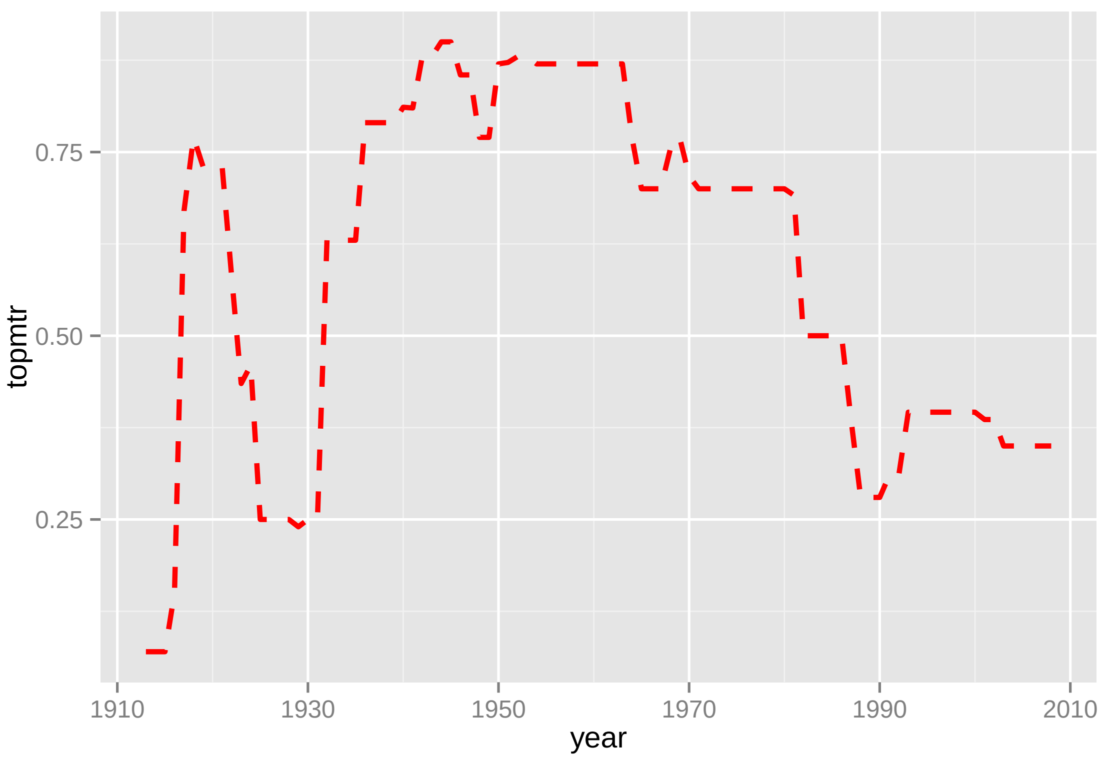
At this part of the problem set I hint to the *data* button which starts the *Data Explorer* that is part of *RTutor*. There you can look at the loaded data sets and find additional information about the data. This additional information is fed in by the .txt file "Optimal_Top_Labor_Taxation_Var" where I typed in a short description for all the variables that are used in the problem set.

> **Info: Variables of interest 1**
>
> - `top1perc` = income shares of the top 1 percent
> - `top1perc_nok` = income shares of the top 1 percent excluding realized capital gains
>
> All income share data in this data set is based on family income and we assume that individual income shares are the same as family based shares.
>
> - `topmtr` = top marginal tax rate on ordinary income; Note that this variable depicts $r$ as discussed in the introduction. That means, we can calculate the top net-of-tax rate via `1 - topmtr`.
> - `topmtr_k` = top marginal tax rate on realized long-term capital gains
>
> **b) Plotting the data**
>
> First, we are going to try to get a graphical view of the data at hand. For that, we use the package *ggplot2* which gives us a very nice array of tools to create graphs. Here is an example of

[7] The corresponding file "US_data_preparation.R" can be found in the folder "Documentation" in the file that comes with this thesis.

how to use the function `ggplot()` from the aforementioned package to plot the variable `topmtr` over time. Just click the *check* button to see the plot:

```
# load the required package 'ggplot2' from the library:
library(ggplot2)

# With 'ggplot(US.data)' we select the data that ggplot is going to use when
# looking for variables. 'geom_line(aes(year, topmtr))' adds a line which
# describes the variable 'top1perc' in dependency of the variable 'year' to
# the ggplot object. The other arguments which have to be seperated with
# commas are pretty self-explanatory; 'size = 1' leads to a thicker line.
p1 = ggplot(US.data) + geom_line(aes(year, topmtr), colour = "red",
                                 linetype = "dashed", size = 1)

# show the plot:
p1
```




This plot illustrates how the top marginal tax rate on ordinary income in the US has changed during the last century.

This part of the problem set shows two things: First, it demonstrates how the code chunks are evaluated within the *html* Tutorial as the plot is shown right beneath. Second, this code chunk is an instance of the strategy to give examples before letting the user work on a problem. This helps the user and is a great way to introduce more complex *R* commands piece by piece.

**Info: Introduction to ggplot2**

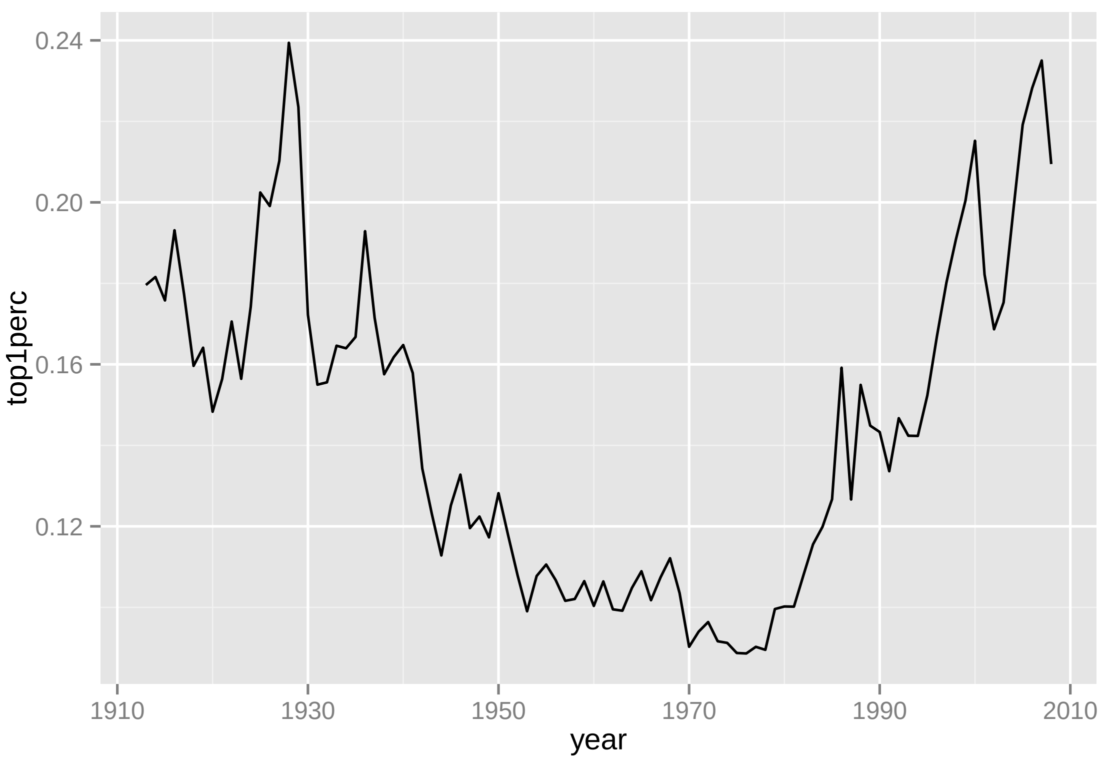
The package *ggplot2* provides a wide array of tools that allow you to create expressive and elegant plots. According to its creator Wickham (2009) it implements the grammar of graphics developed by Wilkinson (2005) which breaks up graphs into elements that can be versatilely combined.

We are often going to use the function `ggplot()` which uses this grammar to enable the user to incorporate various geometrical objects (geoms) such as lines, points, etc. in their plots.

Note: *ggplot2* also contains the function `qplot()` which stands for "quick plot". But as `ggplot()` offers a wider variety of functions and we are going to create rather complex plots, in this problem set we are only going to talk about `ggplot()`.

Now try to plot the variable `top1perc` in dependency of the variable `year` with `ggplot`. Use `geom_line()` as in the example above and save your object as `p2`. Because now we only want to have a black line you can leave out the arguments `colour`, `linetype` and `size`. Don't forget to show your plot.

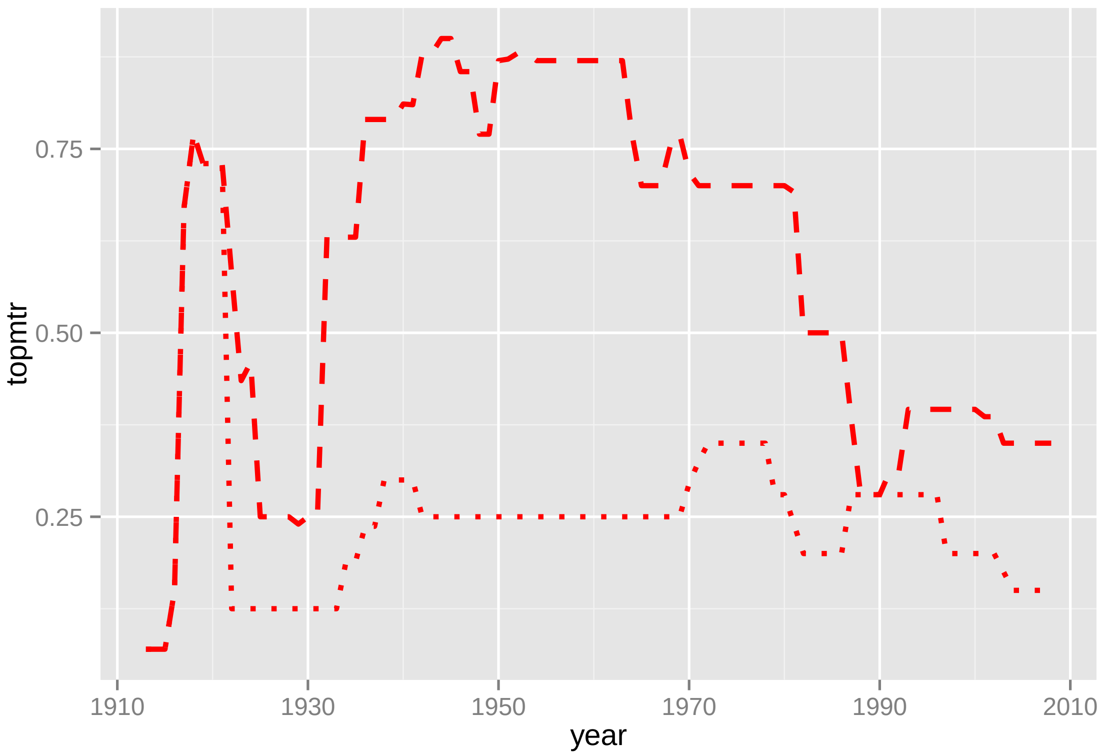
```
p2 = ggplot(US.data) + geom_line(aes(year, top1perc))
p2
```




This plot shows how the share of overall income of the top 1 percent in the US has changed between 1913 and 2008. We are going to discuss the variable's behaviour later in more detail.

You can add new geometrical objects (geoms) to your plot as shown in the following example where a dotted red line that depicts the development of the top marginal tax rate on realized long-term capital gains over time is added to `p1`:

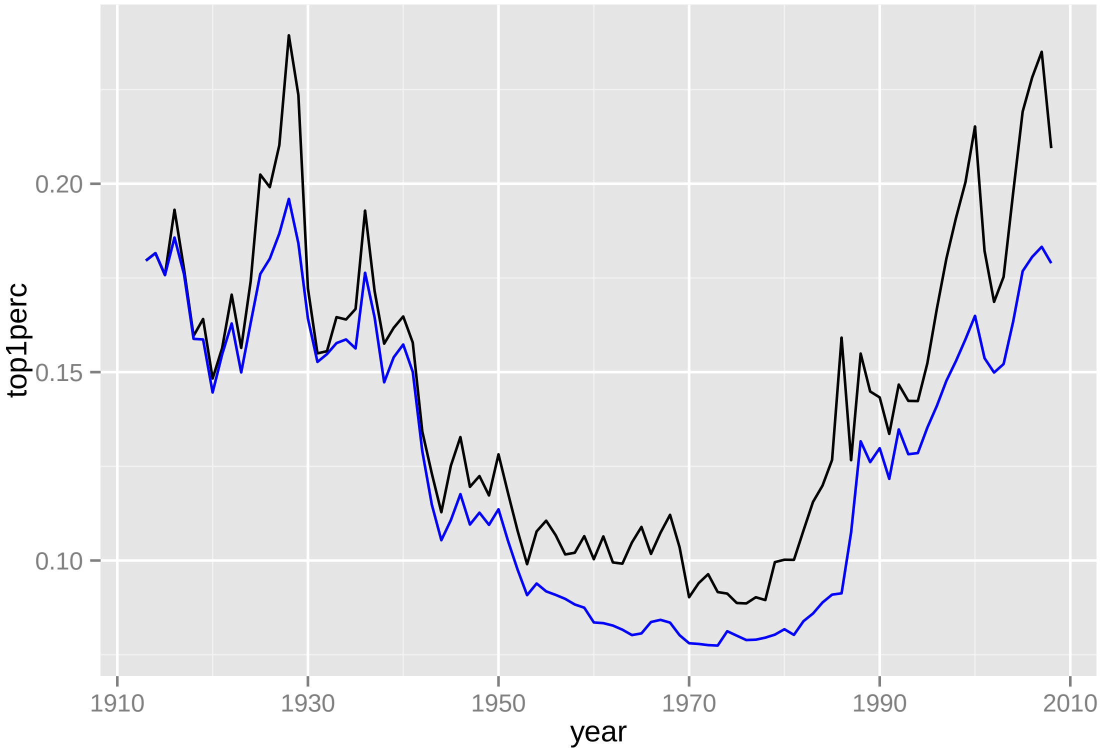
```r
p1 = p1 + geom_line(aes(year,topmtr_k), colour = "red", linetype = "dotted",
                    size = 1)
# You can add any number of geoms to a ggplot object.

p1
```




This chunk demonstrates that it is possible to use variables that were saved in past code chunks. This, of course, makes it possible to build up rather difficult plots in a couple of steps which, again, makes it easier for the user to acquire new skills. Note that it is also possible in *RTutor* to use variables that were saved in code chunks in past exercises as is shown in a later exercise.

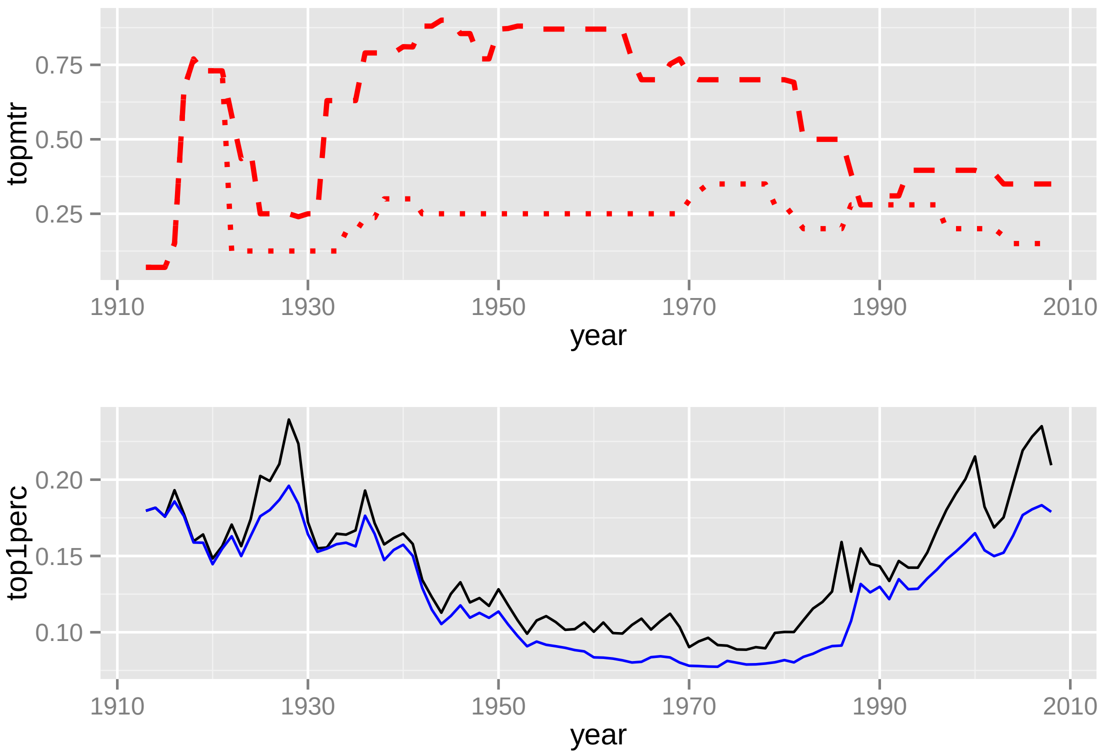
Now try to do the same and add a line which depicts the variable `top1perc_nok`, the top 1 percent income share excluding capital gains, to `p2`. As colour, use the hexadecimal code "#0000FF" which stands for the colour blue. Afterwards, show the plot.

```r
p2 = p2 + geom_line(aes(year, top1perc_nok), colour = "#0000FF")
p2
```




Now we want to look at both plots to get an impression of how the top marginal tax rates influenced the top income shares. That will give us a first chance to discuss implications for the three different elasticities that were mentioned at the beginning. For that reason we are going to use the function `grid.arrange()` from the package *gridExtra*. Take the two plots `p1` and `p2` and arrange them in one plot using the syntax `grid.arrange(plot1, plot2)`.

```
# load the required package from the library:
library(gridExtra)

# Type in your command here:

grid.arrange(p1, p2)
```




In the first plot we can see the top marginal tax rate for ordinary income $r$ as a dashed line and for realized long-term capital gains as a dotted line. The second plot shows the top 1 percent income share including realized capital gains in black and excluding realized capital gains in blue. We can see that there is a clear negative correlation between top 1 percent income share and the top marginal tax rate for ordinary income (top MTR). Especially the rise in income share and drop in top MTR since the 1980s is really significant. If this is due to a causal relationship from top MTR to top income shares, the overall elasticity $e$ is high. Also note that the share of top income that is accounted for by capital gains (the difference between the black and the blue curve) doesn't seem to be correlated to the partly immense difference between top tax rates for normal income and top tax rates for long-term capital gains. According to Piketty et al., as capital gains are the main channel for tax avoidance, this indicates that tax avoidance response $e_2$ makes up only a small amount of $e$.

While we already could get a vague idea of the dimensions of overall elasticity $e$ and tax avoidance response $e_2$ in the US, we cannot yet make a first statement about the influence of supply-side elasticity $e_1$ and compensation-bargaining elasticity $e_3$. Therefore, we are now going to plot real cash market income growth per adult of both top 1 percent incomes and bottom 99 percent incomes over time. These two values are saved in the variables `incadulttop1_n` and `incadultbot99_n`. And again, we are going to compare this plot with the top marginal tax rates.

This part of the problem set emphasizes its purpose. The user is enabled to reproduce a graph that is included in the paper with *R* commands that he learned so far and can then read through a discussion that explains the economic implications that this plot provides according to Piketty et al. So he is able to internalize the use of *R* and at the same time the economic contents of the paper "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities".
The same is true for the following last part of the first exercise that repeats most of the used

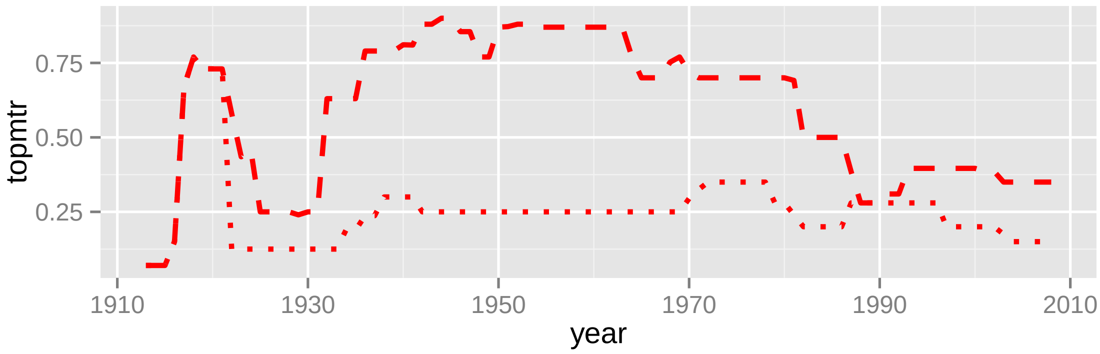
commands and provides a discussion about the graphical implications of the relation between top marginal tax rates and income growth.

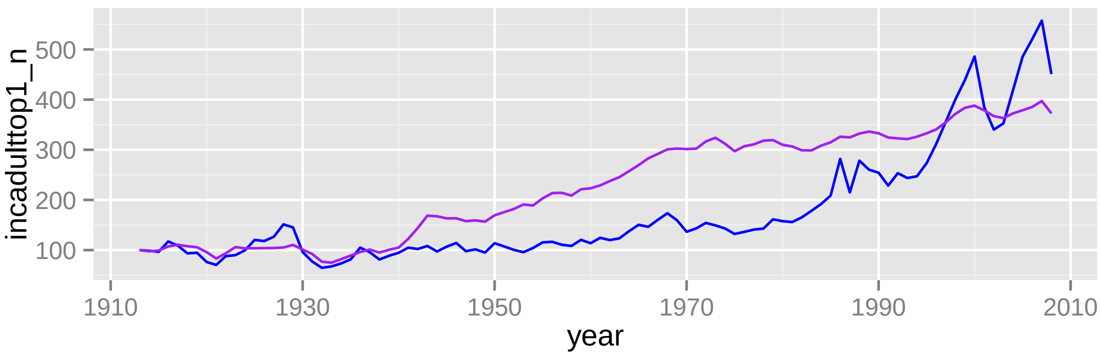
Now you have to use everything that we worked on so far. First, create a ggplot-object p3 in which you plot incadulttop1_n in dependency of the variable year as a line. Use the colour "blue". Then, add a line which depicts incadultbot99_n over time in the colour "purple" to p3. At last, plot p1 together with p3, using the command grid.arrange(). If you need help, feel free to use the *hint*-button.

**Info: Variables of interest 2**

- incadulttop1_n = income of the top 1 percent of the US population, adjusted for inflation; income in 1913 serves as base and has the value 100
- incadultbot99_n = income of the bottom 99 percent of the US population, adjusted for inflation; also in relation to 1913

```
p3 = ggplot(US.data) + geom_line(aes(year, incadulttop1_n), colour = "blue")
p3 = p3 + geom_line(aes(year, incadultbot99_n), colour = "purple")
grid.arrange(p1, p3)
```




We can see that income growth for the bottom 99 percent (purple) was highest between the beginning of the 1930s and the beginning of the 1970s, exactly the period during the last century when net-of-tax rates for the top 1 percent were low (due to a high top MTR). The growth for bottom 99 percent incomes has slowed down since about 1973, while top 1 percent incomes started to grow fast since the 1970s. This hints at the fact that the growth in top incomes might have come to the expense of bottom incomes which would translate into a high bargaining response of the top 1 percent and therefore into a high $e_3$.

By means of plotting our data, we could already draw first conservative conclusions about correlations between top net-of-tax rates and top 1 percent incomes. We also were able to get a first idea of the influence the three different elasticities have on the overall elasticity. But while it is nice to have plots of data to get a first impression, one should never draw final conclusions about causal relationships between different factors just by looking at graphs. That is why, in the next two exercises, we are going to do some quantitative and statistical analysis of our data. To move to *Exercise 1.2* you can click on the *Go to next exercise...*-button in the bottom left as a shortcut.

**Info: Correlation does not imply causation!**
Just because two factors are highly correlated, that doesn't mean that there is a causal relationship between them. They could both be influenced by an unknown externality. If you want to see some examples for high correlations that clearly don't indicate a causal relationship, take a look at Spurious Correlations, a site that illustrates this problem in a rather entertaining way.

*You can find the corresponding graphics and the associated discussions on pages 16f., 20 and 22 of the paper.*

As well as reproducing and discussing the results of the paper, I try to introduce the user to some basic economic rules. This is why I include an *info* section about correlation and causation at this point. Another benefit of mentioning problems that still exist in the analysis that has been done so far, is the fact that this might motivate the user to work on the next exercise to be able to solve them.
At the end of each exercise, I name the pages of the paper that it is based on to enable the user to read the corresponding remarks of Piketty et al. To move on to the next exercise, the user can click on the respective tab on top of the *html* page or on the *Go to next exercise...* button on the bottom left, as mentioned in the problem set.

### Exercise 1.2 US evidence - Simple estimations

In this exercise, we want to get a simple, quantitative estimate of the overall elasticity $e$ for the period from 1960 to 2008, the last year of our data.

#### a) Preparing the data

To do that, we need the average of the first five years (1960-1964) and of the last five years (2004-2008) of our variables `topmtr`, `top1perc` and `top1perc_nok`.

Here is an example of how to use the function `filter()` from the package *dplyr* to isolate the data from the years 1960 to 1964. Click the *edit*-button to activate the code chunk and then proceed using the *check* button, as usual.

**Info: filter() and dplyr**
`filter()` expects two arguments: a data frame and one or more conditions with Boolean operators to filter that data frame by. For example, `filter(US.data, year > 1975)` provides a new data frame that contains all rows of `US.data` where the variable `year` is bigger than 1975.
The command `filter()` is part of the package *dplyr*, a collection of very useful tools to manipulate and work with data frames. We are going to use a variety of other functions from *dplyr* in this problem set as well.

In *R*, there are always many different ways of doing something to get the same results. I tried to introduce the user to commands and packages that are commonly used and provide elegant solutions for statistical tasks, because one objective of the problem set is to give the user tools that he can also use for other purposes. In this instance, one could also utilize the standard function `subset()`, but I think `filter()` from the *dplyr* package by Wickham and Francois (2014) offers a wider array of easy possibilities to manipulate data, especially when tasks get more complex.

```
# first, we load our data again
US.data = read.table("usdata.csv")

# second, we load the required package from the library:
library(dplyr)

# then, we can apply the 'filter()'-function to our data frame:
filter(US.data, year >= 1960 & year <= 1964)
```

```
##   year  top1perc top1perc_nok    gdp topmtr topmtr_k incadult_nok
## 1 1960 0.10034580   0.08356590 526400   0.87     0.25     19.30383
## 2 1961 0.10640656   0.08337601 544700   0.87     0.25     19.70269
## 3 1962 0.09949906   0.08273675 585600   0.87     0.25     20.55049
## 4 1963 0.09916510   0.08163936 617700   0.87     0.25     21.17397
## 5 1964 0.10479104   0.08020751 663600   0.77     0.25     22.23220
##   incadulttop1_nok incadultbot99_nok incadulttop1_n incadultbot99_n
## 1         161.3142          17.86938       113.7370        223.2339
## 2         164.2732          18.24238       124.6085        229.0873
## 3         170.0281          19.04062       119.9461        237.6475
## 4         172.8629          19.64175       123.4476        245.4996
## 5         178.3190          20.65557       137.6674        257.4621
```

You can see that the `filter()`-command provided us with a new data frame that contains only the rows for the years 1960-1964. We can now access the columns that we want to analyse by using the syntax `data.frame$column.name`. Then we can calculate the average of the required variable by using the command `mean()`. In the next example, you can see how to combine these three functions to receive the aforementioned averages for the variable `topmtr`.

```
topmtr_start = mean(filter(US.data, year >= 1960 & year <= 1964)$topmtr)
topmtr_end = mean(filter(US.data, year >= 2004)$topmtr)

# To get an idea of the variables that we calculated, we show them:
topmtr_start
```

```
## [1] 0.85
```

```
topmtr_end
```

```
## [1] 0.35
```

In this chunk we created two new variables. `topmtr_start` is the average of the top marginal tax rate in the US for the years 1960-1964. `topmtr_end` is the same for the years 2004-2008.

The two code chunks above are another example of using demonstrations of how the commands work first before giving the user the task to utilize them. In this instance I split up the example into two chunks to make it easier for the user to comprehend how to apply the different functions. Sometimes it is also recommendable to show the variables that are computed to give the user an idea of their dimension.

Use the same commands and syntax to compute the average of `top1perc_nok` for the years 1960-1964 as `top1_start_n` and for the years 2004-2008 as `top1_end_n`. Then do the same for the variable `top1perc` and safe the results as `top1_start_k` and `top1_end_k`, respectively.

```
# assign the averages of 'top1perc_nok' for 1960-1964 to 'top1_start_n' and
# for 2004-2008 to 'top1_end_n'


# assign the averages of 'top1perc' for 1960-1964 to 'top1_start_k' and for
# 2004-2008 to 'top1_end_k'


top1_start_n = mean(filter(US.data, year>= 1960 & year<= 1964)$top1perc_nok)
top1_end_n = mean(filter(US.data, year >= 2004)$top1perc_nok)
top1_start_k = mean(filter(US.data, year >= 1960 & year <= 1964)$top1perc)
top1_end_k = mean(filter(US.data, year >= 2004)$top1perc)
```

### b) A first comparison

We now arrange the computed variables in the data frame `Comparison.data` and show it using the command `grid.table()` from the *gridExtra* package (the package that we have already utilized to arrange different plots):

```r
# This code chunk creates a new data frame 'Comparison.data' which contains
# three columns for the three different variables that we just summarized
# and two rows for the two periods. Note that 'row.names = c(...)' can be
# used to manually name the rows of your data frame and that 'c(a,b,c)'
# creates an array containing the values 'a', 'b' and 'c'.
Comparison.data = data.frame("top_mtr" = c(topmtr_start, topmtr_end),
                             "top1_n" = c(top1_start_n, top1_end_n),
                             "top1_k" = c(top1_start_k, top1_end_k),
                             row.names = c("1960-1964", "2004-2008"))

# 'round(data, digits = x)' rounds all the values in the data frame 'data'
# to x digits after the comma. 'grid.table(data)' shows the data frame
# 'data' as a table.
grid.table(round(Comparison.data, digits=3))
```

For tasks that only appear once or are too complex to teach the user in a few simple steps, I use prepared code chunks as I do for the examples. This helps the user to work through the problem set in a timely manner and saves him from frustration. However, I still add comments that explain how the commands work in case the user is interested to use them for his own purposes.

| | top_mtr | top1_n | top1_k |
|---|---|---|---|
| *1960–1964* | 0.85 | 0.082 | 0.102 |
| *2004–2008* | 0.35 | 0.177 | 0.218 |

We see that the top marginal tax rate `top_mtr` dropped drastically from 85% in the beginning of the 1960s to 35% at the end of the last decade. During the same period, top 1 percent income shares excluding capital gains `top1_n` more than doubled from 8.2% to 17.7% and top 1 percent income shares including capital gains `top1_k` also increased strongly from 10.2% to

21.8%. As did the plots we generated in *Exercise 1.1*, these numbers also indicate a strong negative correlation between top marginal tax rates and top incomes. We can use these numbers to calculate a first, very simple estimate for our overall elasticity $e$ of top incomes.

**c) Estimation**

We are going to estimate $e$ using the formula

$$\frac{log(top\ 1\ percent\ income\ end)\ -\ log(top\ 1\ percent\ income\ start)}{log(1-top\ MTR\ end)\ -\ log(1-top\ MTR\ start)}.$$

All the mathematical formulas that are included in the problem set are written in *LaTeX*. This demonstrates one of the big benefits of working with *R Markdown*: You can include *LaTeX* and *html* code directly[8].

Try to use this formula to calculate the estimated overall elasticity for the top 1 percent income share excluding capital gains as `e_estimate_n` and including capital gains as `e_estimate_k`. Note: the command to compute the natural logarithm is `log()`. Also, don't forget to use all the required brackets. If you happen to have problems solving this one, using the *hint*-button would be a very good idea. You might want to scroll up a little to remember the names of the variables that we created in the first two exercises that you have to use here.

**Info: Why do we use 'log()' and '(1 - top MTR)'?**
If we use the rule

$$log(x)\ -\ log(y)\ =\ log\left(\frac{x}{y}\right)$$

we can transform

$$\frac{log(top\ 1\ percent\ income\ end)\ -\ log(top\ 1\ percent\ income\ start)}{log(1-top\ MTR\ end)\ -\ log(1-top\ MTR\ start)}$$

into

$$\frac{log\left(\frac{top\ 1\ percent\ income\ end}{top\ 1\ percent\ income\ start}\right)}{log\left(\frac{1-top\ MTR\ end}{1-top\ MTR\ start}\right)}.$$

[8]To enter *LaTeX* equations in an *R Markdown* file you simply put it between two "$" for inline equations and between two "$$" for display equations. *html* code can be entered directly into the file.

If we denote $top\ 1\ percent\ income\ start$ by $top\ 1\ percent\ income$ and $1 - top\ MTR\ start$ by $1 - r$ and use the rearrangements

$$\frac{top\ 1\ percent\ income\ end}{top\ 1\ percent\ income\ start} = 1 + \frac{\Delta\ top\ 1\ percent\ income}{top\ 1\ percent\ income}$$

and

$$\frac{1 - top\ MTR\ end}{1 - top\ MTR\ start} = 1 + \frac{\Delta\ (1-r)}{(1-r)}$$

we get

$$\frac{log\left(1 + \dfrac{\Delta\ top\ 1\ percent\ income}{top\ 1\ percent\ income}\right)}{log\left(1 + \dfrac{\Delta\ (1-r)}{(1-r)}\right)}.$$

Finally, if we utilize $log(1+x) \approx x$, for small x, we obtain

$$\frac{\dfrac{\Delta\ top\ 1\ percent\ income}{top\ 1\ percent\ income}}{\dfrac{\Delta\ (1-r)}{(1-r)}}$$

which is exactly our definition of $e$. This means that using logarithmic formulas helps us to estimate our overall elasticity for top incomes.
(We are also going to use this property in the next chapter for our linear regressions.)

In addition, we use $(1 - top\ MTR)$ because we are interested in the response of top incomes to the net-of-tax rate rather than to the top marginal tax rate.

This is a good example for the purpose of *info* boxes. Some users might not be interested in the reason behind using logarithmic values for this estimation and for the regressions of the next exercise or already possess this knowledge. These users might lose interest if they had to read through longer mathematical derivations. However, it is still an important economic background information that might be helpful for other users in this problem set and also for other economic issues.

```r
# assign e_estimate_n

# assign e_stimate_k

# show e_estimate_n

# show e_stimate_k

e_estimate_n = (log(top1_end_n) - log(top1_start_n)) /
               (log(1 - topmtr_end) - log(1 - topmtr_start))
e_estimate_k = (log(top1_end_k) - log(top1_start_k)) /
               (log(1 - topmtr_end) - log(1 - topmtr_start))
e_estimate_n
```

```
## [1] 0.520623
```

```r
e_estimate_k
```

```
## [1] 0.5173257
```

For developing the formula for `e_estimate_n` the hint message is the exact solution. For `e_estimate_k` the hint message reads "All you have to do is take the first line and replace 'e_estimate_n' with 'e_estimate_k'". Because this code chunk can be rather difficult to solve, I provide the solution for the first part in order not to frustrate the user. However, I do not do this for the second part as the problem set still is supposed to be a stimulating challenge.

For both income including and income excluding capital gains, the estimated elasticity is around $0.52$. This means two things. First, if the rise in income shares is only due to the decrease in the top marginal tax rate, then the real overall elasticity $e$ is high with a little over $0.5$. Second, the two elasticities are almost the same. As capital gains are the main channel for tax avoidance, this implies that $e_2$ is close to zero.

Now it's time for you to answer a first short question. Fill in your answer in place of "???" and remove the comment in the corresponding line.

```r
# Question: If the tax-avoidance elasticity component e2 was a substantial
# fraction of the overall elasticity e, would responses for income including
# capital gains be substantially...
# a) ...smaller or
# b) ...bigger
# than for income excluding capital gains? (considering a change in top MTR
# on ordinary income)
#
# answer = "???" # Assign "a" or "b" to 'answer' and remove the comment

answer = "a"
```

Because this question can be quite difficult for the user to answer, not only do I provide the answer in the *hint* message but I also added an *info* box that explains why the first option is the right solution of this question.

You can click on the following *info* box to read an explanation of the answer to this quiz.

**Info: Tax-avoidance responses**
Let's consider a situation with no tax-avoidance responses. If top MTR on ordinary income $r$ increases, ordinary income decreases according to the overall elasticity $e$. And as there is no tax avoidance response, capital gains also decrease (as they are also dependent on productivity and bargaining). This means that ordinary income decreases by virtually the same rate as income including capital gains.

Now consider a situation with strong tax avoidance. If top MTR on ordinary income increases, again, ordinary income decreases according to $e$. But as there is a strong tax avoidance response, a substantial part of this missing income is shifted to income that is tax-favoured, i.e. capital gains. So a resulting increase in capital gains offsets parts of the decrease in ordinary income. And that means that ordinary income decreases by a higher rate than income including capital gains. (And this translates into a higher $e$ for income excluding capital gains than for income including capital gains.)

Of course, the number we calculated is a very vague estimate, so in the ensuing exercises, we are going to try to produce some more reliable figures.

*This partial exercise is based on pages 16 and 18 of the paper.*

## Exercise 1.3 US evidence - Linear regressions

### a) Elasticity estimations

Now, we want to to check the assumptions that we have gathered so far with some linear regressions. In contrast to *Exercise 1.2*, we are going to use the whole time series from 1913 to 2008.
First, we are going to use the following simple linear regression:

$$log(top\ 1\ percent\ income\ share)\ =\ \alpha\ +\ e\ \cdot\ log(1\ -\ top\ MTR)\ +\ \epsilon$$

(Note: We use logarithmic values in order to obtain elasticity estimates. For further information see the *info* section *Why do we use 'log()' and '(1 - top MTR)'?* in *Exercise 1.2*.)

This example shows you how to fit this regression on top income share excluding capital gains using the command `lm()`:

**Info: The linear regression model and 'lm()'**
In a linear regression model, we assume that a dependent variable (or regressand) is linearly related to several explanatory variables (or regressors). In a standard model it is also assumed that the standard deviation of the error term $\epsilon$ is constant and does not depend on the value of the regressors. The most basic technique to obtain estimates for the unknown parameters that describe the relation between regressors and the regressand (in our case $e$) is the Ordinary Least Squares (OLS) method. The method of OLS minimizes the sum of squared residuals which means it tries to minimize differences between the observations and the predictions of the model. (cf. Hayashi (2000))

In *R*, the command `lm()` can be used to fit linear models. As default, it uses OLS. The syntax needed is `lm(formula, data, ...)`. For example, if you want to fit a model of the linear relationship between columns `a` and `b` and a column `c` in a data set `data` you use `lm(c ~ a + b, data)`. You can assign lm-objects to variables and you can get a more detailed description of your linear model using `summary(lm(...))`.

```
# First, we load our data again:
US.data = read.table("usdata.csv")

# Then, we do our first linear regression:
US.reg1 = lm(log(top1perc_nok) ~ log(1 - topmtr), data = US.data)
```

If it is not declared otherwise in the settings (we are going to look at an example for this in *Exercise 4*), variables are only saved for their respective partial exercise. This is why, here, the variable `US.data` has to be newly assigned. This has three advantages: First, not all the variables have to be available at the same time which relieves the computer. Second, it is more transparent for the user because he can observe more easily which variables he is working with. And third, it enables the user to work on the exercises in a self-chosen order.

Try to do the same regression on top income share including capital gains `top1perc` and assign it to the variable `US.reg2`.

```
US.reg2 = lm(log(top1perc) ~ log(1 - topmtr), data = US.data)
```

Now we can look at the two regressions utilizing the command `showreg()` from the package 'regtools':

**Info: 'showreg()' and the package 'regtools'**
The package *regtools* provides tools to present regression results both in R and as Latex or HTML output. It combines the functionality of a couple of packages when dealing with regressions. We are going to use the included function `showreg()` to obtain a more expressive and elegant presentation for our regressions than the standard `summary(lm())` command offers. You are going to learn about various options this function provides in several examples and exercises.

```r
# load the required package from the library:
library(regtools)

# We pass 'showreg()' a list with our two regressions and add names to the
# two models using 'custom.model.names':
showreg(list(US.reg1, US.reg2),
        custom.model.names = c("capital gains excluded",
                               "capital gains included"))
```

```
##
## =================================================================
##                   capital gains excluded  capital gains included
## -----------------------------------------------------------------
## (Intercept)       -1.84 ***               -1.69 ***
##                   (0.05)                  (0.04)
## log(1 - topmtr)    0.25 ***                0.26 ***
##                   (0.04)                  (0.03)
## -----------------------------------------------------------------
## R^2                0.29                    0.38
## Adj. R^2           0.29                    0.38
## Num. obs.         96                      96
## =================================================================
## *** p < 0.001, ** p < 0.01, * p < 0.05
```

We find the estimator for $e$ in our regression in the line *log(1 - topmtr)*. The value in brackets underneath is the calculated standard deviation of our estimator. The stars behind represent the significance level of our estimator. The three stars for *log(1 - topmtr)* in the first column mean that if the true value of $e$ was zero then the probability of estimating a coefficient of $0.25$ or higher would be below $0.1\%$. (The higher the estimate for a coefficient is compared to its standard deviation, the more significant is the influence of the regressor on the regressand.) Using this type of regression, we get an elasticity of around $0.25$, again with no significant difference whether we include capital gains or not.

As a next step, we want to include a time trend in our regression in order to test for the influence time has on our estimations. Our regression has now the following form:

$$log(top\ 1\ percent\ income\ share) = \alpha + e \cdot log(1 - top\ MTR) + c \cdot time + \epsilon$$

**Info: Controlling for a variable**
Controlling for a variable, in our case the variable *time* (or, to be precise, `year`), means to separate out the effect of an independent variable from the effects of the other variables on the regressand (cf. Lewis-Beck et al. (2003)). To implement this in a regression you have to add this particular independent variable to the regressors.

To implement this regression, you can copy the code we used to get `US.reg1` and `US.reg2`. Then add the variable `year` to the regressors using the concatenator `+`. Safe these new time-trend-including regressions as `US.reg3` for the variable `top1perc_nok` and `US.reg4` for the variable top1perc. Then use showreg() to show the results. Adopt the manual model names from the last example.

```r
US.reg3 = lm(log(top1perc_nok) ~ log(1 - topmtr) + year, data = US.data)
US.reg4 = lm(log(top1perc) ~ log(1 - topmtr) + year, data = US.data)
showreg(list(US.reg3, US.reg4),
        custom.model.names = c("capital gains excluded",
                               "capital gains included"))
```

```
##
## ===============================================================
##                   capital gains excluded  capital gains included
## ---------------------------------------------------------------
## (Intercept)        7.89 ***                3.81 *
##                   (1.68)                  (1.54)
## log(1 - topmtr)    0.30 ***                0.29 ***
##                   (0.04)                  (0.03)
## year               0.00 ***                0.00 ***
##                   (0.00)                  (0.00)
## ---------------------------------------------------------------
## R^2                0.48                    0.46
## Adj. R^2           0.47                    0.45
## Num. obs.         96                      96
## ===============================================================
## *** p < 0.001, ** p < 0.01, * p < 0.05
```

Again, we receive a significant positive elasticity with around $0.3$. Again, with no notable difference between the two income shares. Our results are therefore robust to including a time trend. Piketty et al. conclude that this means that there is presumably a strong link between top marginal tax rates and top 1 percent shares and that tax avoidance effects (expressed through $e_2$) tend to be small, just as we expected when looking at the summarizing plots from *Exercise 1.1*.

**Info: Further insights**

- You might have already noticed that the elasticities we get from our regressions are lower than the ones we got from our simple estimations from *Exercise 1.2*. Piketty et al. explain this with the fact that before the 1970s the average marginal tax rate for the top 1 percent was smaller than the statutory tax top rate that we consider. That means that an increase or a decrease in top MTR didn't actually affect the whole top 1 percent which leads to an underestimation of the overall elasticity $e$.
- Our estimation of $e_2$ might be biased downward if tax avoidance opportunities would have declined throughout our time span. However, as it is explained in the paper, it seems more likely that the exact opposite has happened with an increasing availability of, for example, offshore accounts and fringe benefits.
- Note that Piketty et al. do find strong evidence for large tax-avoidance responses in the short run. But their analysis leads them to state that in the long run the tax-avoidance elasticity "appears to be small (say, $e_2 < 0.1$)".

### b) Effect of top MTR on income growth

In *Exercise 1.1* we looked at the influence marginal top tax rates have on income growth of the top 1 and bottom 99 percent. We are going to do the same using regressions, again, to check how the different channels of income response contribute to our overall elasticity. Our regressions will only use data excluding capital gains because there is again no significant difference between the two types of income, as is shown in the paper.

Including a time trend, we are going to regress the influence of the top net-of-tax rate `1 - topmtr` on three different income variables in our data set: average real income `incadult_nok`, top 1 percent real income `incadulttop1_nok` and bottom 99 percent real income `incadultbot99_nok`, all excluding capital gains. If you want to get more details on these variables, feel free to use the data explorer.

Our regression now has the form:

$$log(income) = \alpha + e \cdot log(1 - top\ MTR) + c \cdot time + \epsilon$$

The regressions for `incadulttop1_nok` and `incadultbot99_nok` are already given, you only have to assign the regression for `incadult_nok` to the variable `US.reg7`. It's not a big challenge, but the standard `lm()` regression is a powerful tool which you can't practice using enough.

```
US.reg5 = lm(log(incadulttop1_nok) ~ log(1 - topmtr) + year, data = US.data)
US.reg6 = lm(log(incadultbot99_nok)~ log(1 - topmtr) + year, data = US.data)

# Do the last one for 'incadult_nok':

US.reg7 = lm(log(incadult_nok) ~ log(1 - topmtr) + year, data = US.data)
```

So far, we have only used the standard linear regression which uses the assumption that the variance of the error term is constant (known as homoscedasticity). This might not be true and might simulate significance of results that are not significant. In their paper, for their regressions on the US data Piketty et al. use a more robust version, a Newey-West regression with 8 lags, to be exact. This type of regression is put into use to deal with autocorrelation and heteroskedasticity. To find out more, click the *info*-button below.

**Info: Newey West**
Besides dealing with heteroskedasticity (the opposite of homoscedasticity), in our data we might also face the problem of autocorrelation. We speak of autocorrelation if the data in a time series is influenced by its own historical values. In our data it is very compelling to expect this phenomenon as top 1 percent income share in one year probably influences top 1 percent income share in the next year.

To consider these two things in our regression model we are going to use the Newey-West estimator for the standard deviations. This estimator was first introduced by Newey and West (1987). As this estimator checks for both heteroskedasticity and autocorrelation, you'll find that the estimated standard deviations increase.

To apply standard errors from Newey West, we utilize the possibility to pass `showreg()` a list of covariance matrices. To get these matrices, we use the function `NeweyWest()` from the *sandwich* package. Take a look at the code in the following example to comprehend it, then click the *check*-button as usual.

**Info: 'NeweyWest()' and the 'sandwich' package**
The *sandwich* package serves to provide different estimators that are heteroskedasticity-consistent (HC) and heteroskedasticity and autocorrelation consistent (HAC).
One of these HAC estimators is the Newey West estimator. We can obtain the according covariance matrix using the function `NeweyWest(x, ...)` where `x` is a fitted model object. In our case, to obtain the estimator as described by Newey and West (1987) we set the argument `prewhite` to FALSE and to consider autocorrelation up to eight years back we set the argument `lag` to 8.

```r
# load the required package from the library:
library(sandwich)

# Create a Newey West covariance matrix for each regression model and assign
# them to the list 'cov.list':
cov.list = list(NeweyWest(US.reg5, lag = 8, prewhite = FALSE),
                NeweyWest(US.reg6, lag = 8, prewhite = FALSE),
                NeweyWest(US.reg7, lag = 8, prewhite = FALSE))

# Summarize the three models in the list 'reg.list':
reg.list = list(US.reg5, US.reg6, US.reg7)

# Show the three regressions, using the Newey West cov. matrices. Again,
# we use custom model names. 'digits = 3' rounds all values to three digits
# after the comma. 'include.adjrs = FALSE' and 'include.rsquared = FALSE'
# surpress the lines 'R^2' and 'Adj. R^2' for better clarity.
showreg(reg.list , custom.model.names = c("Top 1 percent real income",
                                          "Bottom 99 percent real income",
                                          "Average real income"),
        digits = 3, vcov.li = cov.list, include.adjrs = FALSE,
        include.rsquared = FALSE)
```

```
##
## ====================================================================
##                   Top 1 % income  Bottom 99 % income  Average income
## --------------------------------------------------------------------
## (Intercept)      -19.000 ***     -33.077 ***         -31.492 ***
##                   (3.451)         (2.901)             (2.648)
## log(1 - topmtr)    0.265 ***      -0.080 *            -0.036
##                   (0.046)         (0.039)             (0.035)
## year               0.013 ***       0.018 ***           0.018 ***
##                   (0.002)         (0.001)             (0.001)
## --------------------------------------------------------------------
## Num. obs.         96              96                  96
## ====================================================================
## *** p < 0.001, ** p < 0.01, * p < 0.05
```

Note that I have shortened the output of this code chunk to make it fit. Of course, in the problem set the names of the different models are in line with what is set in the argument `custom.model.names` of the `showreg()` command.

The results confirm our finding from exercise *a)* that top 1 percent income is strongly positively linked to the top net-of-tax rate. $e$ is highly significant. In addition, bottom 99 percent income is significantly (at the 5% level) negatively related to the top net-of-tax rate. Piketty et al. argue that this suggests that increases in top incomes come at the expense of bottom incomes (an effect called "trickle up"), indicating a high $e_3$. Average incomes are not significantly linked to top net-of-tax rates, implying that the supply-side response of top earners to changes in top marginal tax rates $e_1$ makes up only a small portion of the overall response $e$ (if top earners responded to raises in top net-of-tax rates primarily by increasing productivity, then that should have a positive influence on GDP).

At the end of this exercise, let us do a little quiz. Remember, you just need to replace the "???"'s with the right answers.

```
# Question 1: Considering the implications that Piketty et al. detect in
# their analysis so far, which one of the three elasticities seems to be the
# biggest in the US:
# a) e1 (supply-side response),
# b) e2 (tax-avoidance response) or
# c) e3 (compensation-bargaining response) ?
# answer1 = "???" # Assign "a", "b" or "c" to 'answer1' and remove the
#  comment
#
# Question 2: Look at our last regressions. Is GDP growth likely linked to
# the top marginal tax rate?
# answer2 = "???" # Assign "yes" or "no" to answer2 and remove the comment
#
# Question 3: Can we be sure yet, that our preliminary conclusions reflect
# reality?
# answer3 = "???" # Assign "yes" or "no" to answer3 and remove the comment

answer1 = "c"
answer2 = "no"
answer3 = "no"
```

For more information, especially on *question 3* of our little quiz, click the *info*-section below.

**Info: Reverse causality and external effects**

While we could already draw some conclusions about the dimension and the composition of the elasticity $e$ of top incomes, we still face a couple of problems that prevent us from passing our definite verdict on optimal top tax rates:

- So far, we have assumed that the link between top net-of-tax rates and top income shares is causal. But there is a chance that we face reverse causality, meaning that, actually, top income shares influence top net-of-tax rates. This might be the case if higher income empowers the top 1 percent to influence policy via lobbying or funding of campaigns in order to lower the top marginal tax rate.
- Slower GDP growth coincides with an increase in top net-of-tax rate. This slower economic growth could very well be unrelated to cuts in top MTR and influenced by an external effect. This could lead to an overestimation of compensation-bargaining elasticity $e_3$ and an underestimation of supply-side elasticity $e_1$ as positive effects of a higher top 1 percent productivity could be shrouded by and the decline in bottom 99 percent income growth could be due to diminished GDP growth. (This is a rather compelling argument as growth did slow down in a lot of OECD countries since the 1970s.)

Besides the problems that you can read about in the *info* section above, keep in mind that so far, we only have analysed data from the US. If we want to obtain universal insights into top income elasticities and optimal top tax rates we have to look at further evidence.

*For the regressions and the associated discussion compare pages 16 and 18-22.*

For this first exercise I tried to follow the same structure that Piketty et al. use in their paper and that is also generally helpful when looking at data. First, I help the user visualize the data and explain first implications that can be detected (following Piketty et al.'s remarks). Then, the user is enabled to calculate a first estimation of the overall elasticity of top incomes that gives him a quantitative understanding of the economic issue. And lastly, the graphical implications and the calculations are examined using statistical analysis.
Again, at the end of the exercise I mention pending problems and point to following exercises for possible solutions.

### Exercise 2.1 International evidence - Top income shares

After we have taken a look at data just from the US in *Exercise 1*, we are next going to analyse international macroeconomic evidence in order to be able to substantiate our findings. We are going to deepen the *R* skills that we have acquired in *Exercise 1* and also consider some more sophisticated ways to study and visualize data.

For the second exercise I use a slightly different structure than for *Exercise 1*. While the first exercise was separated into one part where the user translates the data into plots and another part where the data is statistically analysed, in *Exercise 2* I use the first part to cover the relation between top marginal tax rates and top income shares and the second one to cover the influence top tax rates have on GDP growth to check for the magnitude of supply-side responses. This means that the user does the same analysis for the international evidence that he does for the US evidence. But the new structure is supposed to make the problem set more exciting.

#### a) Using a Google motion chart

First, we need to upload the new data *intdata.dta* which should also be saved in the directory of your problem set. This time, the data is stored in the *.dta*-format which is used to store data by the statistical software *Stata*. To load a *.dta* file we have to use the command `read.dta()` from the package *foreign*. This command is used with the same syntax as `read.table()`.

**Info: 'read.dta()' and the foreign package**
The package *foreign* was built to enable the user to upload data that was stored by one of several statistical analysis software (for example, *SPSS* and *Stata*). It also contains functions that allow you to save data in the file format of the mentioned programs.
In their paper, Piketty et al. analyse their data using the very popular statistical software package *Stata*. That is why their data is provided in the *.dta*-format that is used by *Stata* to store data. If you want to upload such a file into your *R* environment, you have to use the command `read.dta()` which is used with the same syntax as the `read.table()` command that we introduced in the very first exercise: `read.dta("tablename")`.

Try to upload the file *intdata.dta* into your work space using the syntax `read.dta("filename")` and assign it to the variable `Int.data`.

```
# First, we have to load the required package
library(foreign)

# Then, upload the data:

Int.data = read.dta("intdata.dta")
```

Like the data file "usdata.csv" that I used for the first exercise of the problem set, "intdata.dta" also is a modified version of the raw data set that Piketty et al. provide. Strictly following the

*Stata* documentation of their paper, I prepared the data to the extent that the user is able to immediately start with data analysis. I saved the prepared data set as a .dta file[9] to show the user how to import data from foreign file types.

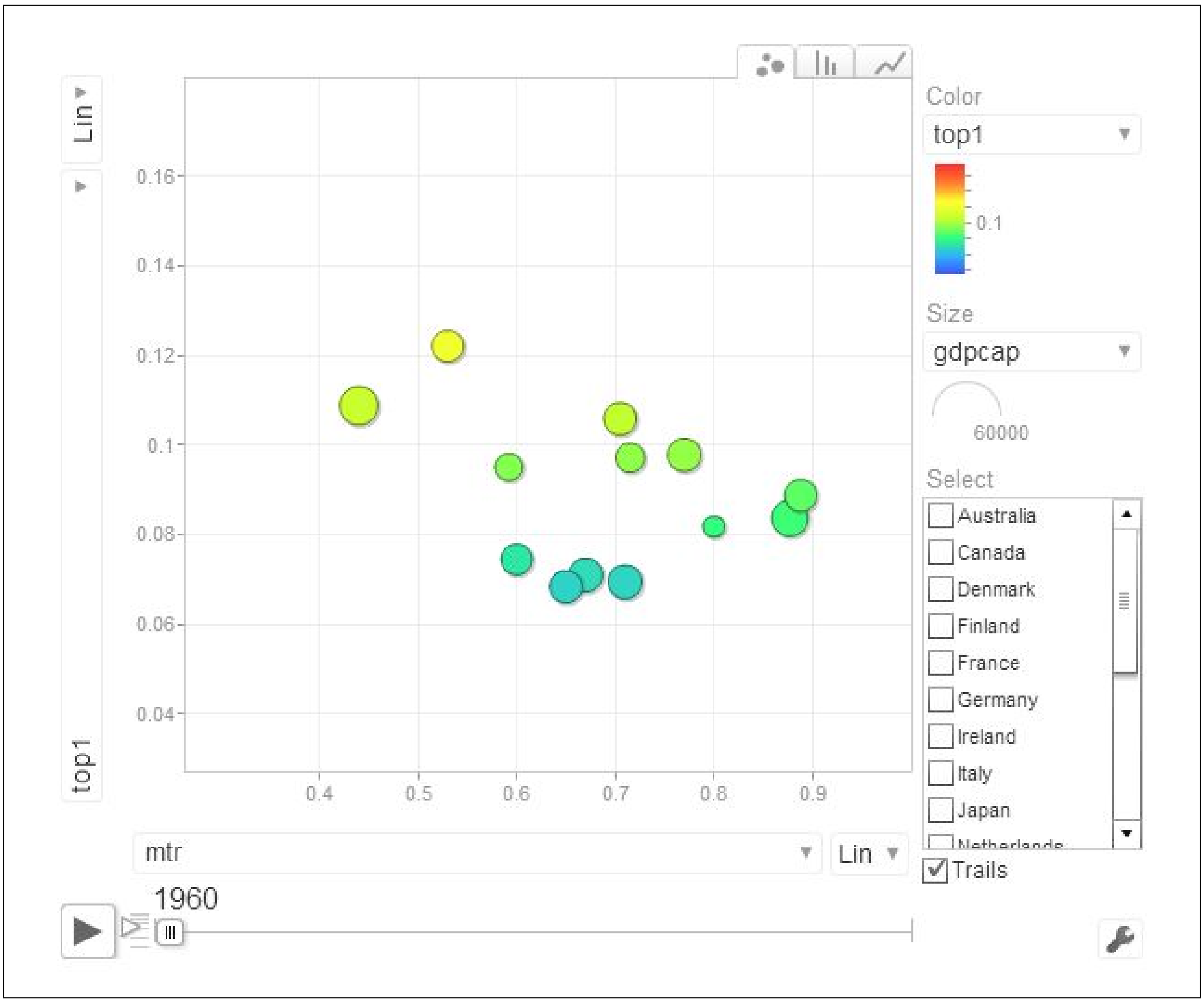
This data set contains the three macroeconomic values marginal tax rates for top incomes as `mtr`, income shares of the top 1 percent `top1` and real GDP per capita `gdpcap` from 18 different OECD countries for the years 1960 to 2010. For some variables, the data is not complete over the whole time span for some countries. See the *Data Explorer* for more information.

To get a first overview of how top marginal tax rates influenced top 1 percent income shares in the different countries we are going to use a motion plot that is provided by the package *googleVis* that allows you to create a lot of different Google plots. This package is a great tool set to illustrate your data which is definitely worth exploring. As a start, you can try to comprehend the code in the next example chunk.

**Info: googleVis**
The *googleVis* package offers the possibility to visualize *R* data frames with interactive Google Charts. The output of a *googleVis* function is html code which can be opened in a new window in your browser or embedded into existing html documents and websites. It is a great tool to dive into data and in many cases by far exceeds the possibilities of visualization of information that more common ways to plot provide.
We are going to use maybe the most popular kind of Google chart: The Motion Chart. To do so, we utilize the function `gvisMotionChart()`. To get an idea of how this function works, take a look at the example below.

```
# First, we load the needed package
library(googleVis)

# Then, we create a motion chart using the command 'gvisMotionChart()'. We
# set 'idvar = "country"' because we are interested in the data for each
# country and 'timevar = "year"' because 'year' is the variable that depicts
# time in our dataset. We use 'xvar = "mtr"' and 'yvar = "top1"' because we
# want to know more about the influence of top MTR on top 1 percent incomes.
# 'colorvar = "top1"' and 'sizevar = "gdpcap"' set default values for our
# plot. In addition, we use our data only until 2005 because after this
# point in time we don't have sufficient information for some countries.
mp = gvisMotionChart(subset(Int.data, year <= 2005), idvar = "country",
                     timevar = "year", xvar = "mtr", yvar = "top1",
                     colorvar = "top1", sizevar = "gdpcap")

# Finally, we simply plot the motion chart
plot(mp, tag = "chart")
```

[9] In the folder "Documentation" you can find the file "Int_data_preparation" that shows exactly how I manipulated the data to obtain the file "intdata.dta".




Because this *Google motion chart* is a dynamic *html* object, I include it in this illustration as a screenshot of the problem set[10]. In the problem set, all the variables that are used for the x-axis, the y-axis and the color and size of the bubbles can be set manually. If the user clicks the play button, the motion plot runs through the data for the included years.

Each bubble in the graph depicts one country and its relationship between top MTR and top 1 percent income share. If you click the play-button in the bottom left, you can see how that relationship developed over time for the different countries. You can play around a little bit with the graph. For example, click on different bubbles to see what country they represent or change the variables for *x*, *y*, *color* and *size*.

Now, deploying the motion chart, we want to do a little quiz:

[10]To include a *Google Chart* in an *R* problem set, the argument "results" of the code chunk has to be set to "'asis'" in order for *R Markdown* not to further evaluate the *html* output.

```
# Question 1: If you take a look at the year 1960, can you see a strong
# relation between top MTR and top 1 percent income?
# answer1 = "???" # Assign "yes" or "no" to 'answer1' and remove the comment
#
# Question 2: Take a look at the year 2005. Is there now a strong
# relationship?
# answer2 = "???" # Assign "yes" or "no" to 'answer2' and remove the comment
#
# Question 3: Comparing the development of Japan and Canada between 1960 and
# 2005, which one of the two countries seems to have a higher elasticity of
# top income shares? (Tip: Select the two countries, activate 'Trails' and
# let the chart play through the whole time span.)
# a) Canada
# b) Japan
# answer3 = "???" # Assign "a" or "b" to 'answer3' and remove the comment

answer1 = "no"
answer2 = "yes"
answer3 = "a"
```

Analyzing the motion chart closely, the following lessons emerge:

- The distribution of top marginal tax rates across OECD countries was very widespread in 1960 (between about 45% and 90%) and has become narrow in 2005 (between about 40% and 60%). At the same time, the dispersion of top income shares has become much broader (the difference between the minimum and maximum of top income share grew from about 5.5 percentage points in 1960 to about 12 percentage points in 2005).

- While during the 1960s and 1970s there is no plain correlation visible between top net-of-tax rates and top 1 percent income share, beginning in the 1980s we can see more and more a very strong positive correlation between these two variables. The situation we can observe in 2005 is consistent with our findings from *Exercise 1*.

- The development across OECD countries is very heterogeneous, implying that the overall elasticity strongly depends on the institutional reality in each country. We will look at some elasticity estimations for individual countries at the end of this partial exercise.

### b) Let's back it up with statisitcs!

So far, today the relationship between top net-of-tax rate and top income share internationally leads to the same conclusions as the US evidence. Again, we want to support these conclusions we drew using charts with regression analysis. We are going to use the following linear regression including country fixed effects to test the influence of top net-of-tax rates on top 1 percent income shares in our international data:

$$log(top\ 1\ percent\ income\ share) = \alpha + e \cdot log(1 - top\ MTR) + b \cdot country + \epsilon$$

**Info: Country fixed effects**
To include country fixed effects in a regression means to control for the influence the variable *country* has on our dependent variable. Fixed effects, in contrast to random effects, describe a situation were we have data from all manifestations of a variable that are of interest. This is true for the variable *country* in our case as we have data for all 18 countries that we are interested in. Opposite to testing for the variable *time*, this leads to as many new coefficients as there are individuals in the data. In our case, this means that we get 18 new estimators for the fixed effect of each country. Naturally, each estimator equals zero for 17 countries and an estimated value for one of them.

Implement this regression on our data set `Int.data` using the `lm()` command, just as you did in *Exercise 1.3*. Note that you can add the variable `country` to the regressors the same way as you did with the variable `year`. Top net-of-tax rate can be obtained using `(1 - mtr)`. Save the regression as `Int.reg1`.

```
Int.reg1 = lm(log(top1) ~ log(1 - mtr) + country, data = Int.data)
```

In the motion chart we were able to observe that the strength of the correlation between top net-of-tax rates and top income shares increased over time. That's why we want to do the exact same regression we have just done for the two time spans until 1980 and from 1981. To do that, use the `filter()`-command to pass `lm` the correct data and save the regression for the first period as `Int.reg2` and for the second period as `Int.reg3`.

```
Int.reg2 = lm(log(top1) ~ log(1 - mtr) + country,
              data = filter(Int.data, year <= 1980))
Int.reg3 = lm(log(top1) ~ log(1 - mtr) + country,
              data = filter(Int.data, year >= 1981))
```

Let's take a look at the results from our three regressions. Remember how we used Newey West standard errors in *Exercise 1.3*? This time we are going to use another type of standard errors that is very often used to obtain a robust regression by utilizing a standard function of the `showreg()` command.

**Info: Using standard robust errors with 'showreg()' - HC1 as an example**
You don't necessarily have to pass `showreg()` a covariance matrix (as we did for the Newey West) to get a robust regression. If you set the `robust` argument to TRUE you can choose between various different types of robust standard errors. For this purpose, you can set the argument `robust.type` to one of the following values: "HAC", "cluster", "HC1", "HC2", HC3","HC4" or "NeweyWest".

We are going to use "HC1", which describes a simple degrees of freedom correction as defined by MacKinnon and White (1985). This type of standard errors is often used to get a robust regression, for example, as a standard in *stata*.

```r
# We set 'robust = TRUE' and 'robust.type = "HC1"' to obtain robust standard
# errors. Further, we use 'omit.coef = "country"' to leave out all 18 lines
# with the word 'country'.
showreg(list(Int.reg1, Int.reg2, Int.reg3),
        custom.model.names = c("1960-2010", "1960-1980", "1981-2010"),
        digits= 3, robust= TRUE, robust.type = "HC1", omit.coef = "country")
```

```
##
## ====================================================
##                1960-2010     1960-1980     1981-2010
## ----------------------------------------------------
## (Intercept)    -2.450 ***    -2.813 ***    -2.209 ***
##                (0.035)       (0.055)       (0.045)
## log(1 - mtr)    0.314 ***     0.007         0.626 ***
##                (0.025)       (0.039)       (0.044)
## ----------------------------------------------------
## R^2             0.637         0.680         0.770
## Adj. R^2        0.628         0.660         0.761
## Num. obs.     774           292           482
## ====================================================
## *** p < 0.001, ** p < 0.01, * p < 0.05
```

We notice that the estimated elasticity was very low and insignificant during the first period from 1960 to 1980 and very high at around $0.6$ in the period from 1981 to 2010. This suggests different behavioural responses due to institutional changes, as do the differences between countries. If you look at the whole period from 1960 to 2010 you obtain an estimated elasticity of about $0.3$. For both the whole time span and the later phase, elasticities are highly significant.

This means, that for the past decades our international evidence is consistent with our US evidence. This strongly implies that there is in fact a strong positive correlation between top net-of-tax rates and top 1 percent income shares. Hence, the overall elasticity $e$ can be, especially for the recent past, estimated at a high level of about $0.5$, as Piketty et al. do in their paper.

**Info: Differences between countries**
We could already infer from our plot that there are significant differences in elasticity across countries, implying that the institutional set-up influences the top income elasticity of a country. Piketty et al. give a few examples for this phenomenon: While overall the elasticity for the whole period can be estimated at around $0.3$, in English-speaking countries, especially in the US and the UK, elasticity is around $0.5$ and in countries such as Japan or Sweden it is close to zero.

*This partial exercise is based on pages 3 and 22-26 of the paper.*

### Exercise 2.2 International evidence - Economic growth

#### a) One more plot

Again, we want to check if the high responsiveness of top incomes is due to a change in economic productivity and GDP and hence due to a high supply-side elasticity $e_1$. To do this, we first prepare our data to be able to create a plot. In order to do this we are going to use four different functions that are provided by the *dplyr* package:

- `group_by()` takes a data frame and returns a data frame grouped by a certain variable. This command is used with the syntax `group_by(table.name, variable.name)`.
- `do()` gets passed a grouped data frame `data` and performs a computation on each group. It is a general purpose complement to the specialized manipulation functions such as `filter()`. This command uses the syntax `do(grouped.data.name, function.name(.))`. `.` is used as a placeholder for the grouped data. `do()` always returns a data frame.
- `first()` and `last()` are specialized versions of the command `nth()` and can be utilized to return the first and last element, respectively, of a column.

The next example shows you how we can use these functions to create the data frame `Int_growth.data` which, for each country, contains the variable `dmtr_perc` which depicts change in top MTR over the time span of our data in percentage points and `gdp_growth` which depicts average annual real GDP growth for the same period in percent.
Of course, as there are a lot of new functions in this code chunk simultaneously, it might be a little hard to understand how the single commands work. But this example is primarily designed to give you an impression of what is possible to accomplish with the *dplyr* package. Just try to comprehend it as good as you can, using the descriptions given above.

```
# Again, we upload our data:
Int.data = read.dta("intdata.dta")

# Then, we create a new data frame that contains the variables we need:
Int_growth.data = do(group_by(Int.data, country),
                     data.frame(dmtr_perc= 100 * (last(.$mtr)-first(.$mtr)),
                                gdp_growth = 100 * ((last(.$gdpcap) /
                                          first(.$gdpcap))^(1 / 50) - 1)))
```

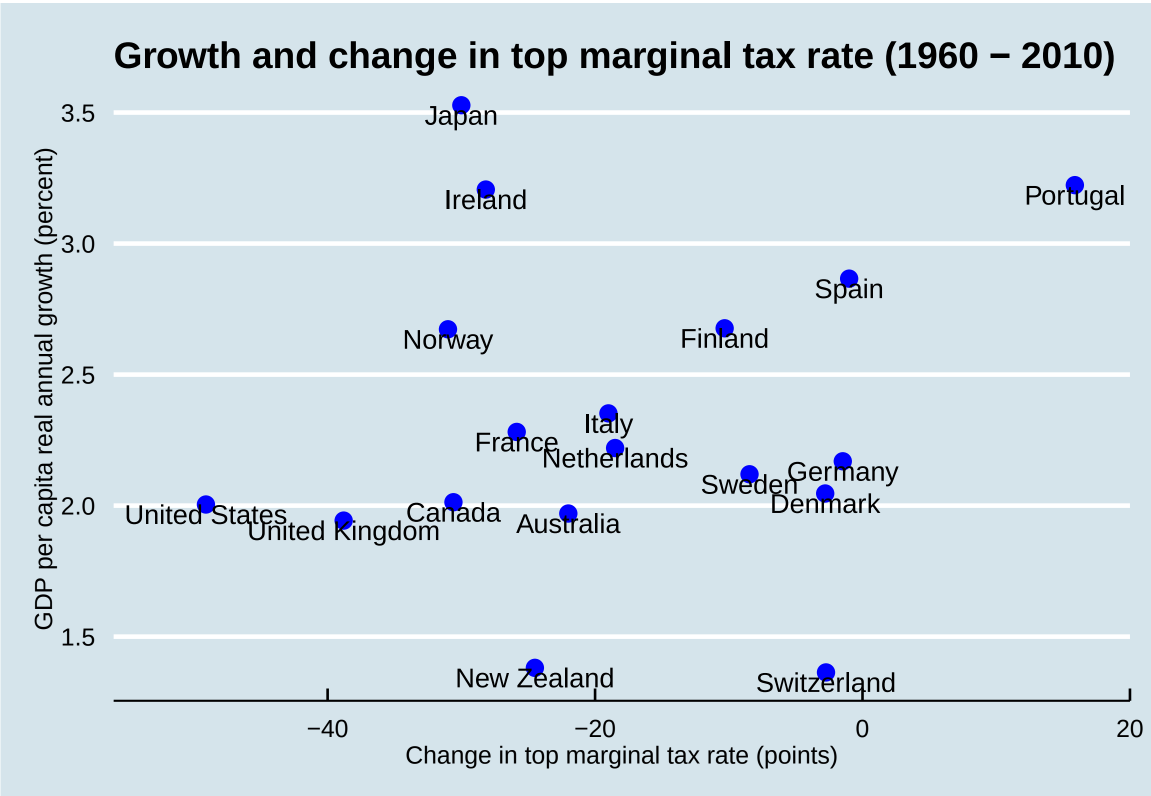
Use the *data*-button to take a look at `Int_growth.data`. You can see that it has three columns and 18 rows, containing the overall change in top MTR and average annual GDP growth for each country.

Now, let's see if you still remember the things that we've learned in *Exercise 1.1*: Create the ggplot object `p1` that uses the new created data and depicts average GDP growth `gdp_growth` in dependency of overall top MTR change `dmtr_perc` as blue points that have the size `3.5`. Note: points can be depicted with `geom_point` instead of `geom_line`. Don't forget to add `size` to its arguments.

```
p1 = ggplot(Int_growth.data) + geom_point(aes(dmtr_perc, gdp_growth),
                                          colour = "blue", size = 3.5)
```

Before we plot `p1` we are going to add expressive labels to our plot. In addition, we are going to use the theme that is used by the weekly newspaper "The Economist" for their plots. This function is provided by the package *ggthemes*. (Again, this example is intended rather as a stimulus than for you to completely internalize.)

```
# First, we add a title, and a label for both the x- and the y-axis:
p1= p1+ ggtitle("Growth and change in top marginal tax rate (1960 - 2010)")+
          xlab("Change in top marginal tax rate (points)") +
          ylab("GDP per capita real annual growth (percent)")

# Add a label for each country at the position where the points are.
# 'Vjust = 1' shifts the labels a little downward so that they are directly
# below the points.
p1 = p1 + geom_text(aes(x = dmtr_perc, y = gdp_growth, label = country),
                    size = 4, vjust = 1)

# Change the range of the x-axis to get all labels into the plot:
p1 = p1 + coord_cartesian(xlim = c(-56,20))

# Load the package 'ggthemes' from the library and add the Economist theme
# to the plot:
library(ggthemes)
p1 = p1 + theme_economist()

# Finally, show the plot:
p1
```




If we look at the plot, there is no significant correlation visible between cuts in top tax rates and economic growth. (This also holds true for the sub periods, as Piketty et al. show in the Appendix of their paper.) If a substantial part of the high responsiveness of top incomes to top tax rates would be due to supply-side responses, decreases in top tax rates should also lead to higher GDP growth. However, this is not the case in our data. This indicates a high compensation-bargaining elasticity $e_3$ and a low supply-side elasticity $e_1$. This is especially true if we assume that the influence of tax avoidance responses, and therefore $e_2$, is small as we have found in our US evidence.

### b) One more regression

One last time, we want to test our conclusions with regression analysis. We are going to use the following regression over the whole time span:

$$log(real\ GDP\ per\ capita) = \alpha + \beta \cdot log(1 - top\ MTR) + c \cdot time + b \cdot country + d \cdot time \cdot initial\ GDP\ per\ capita + \epsilon$$

As you can see, not only do we implement a time trend and country fixed effects, we also include the interaction of initial GDP per capita with a time trend. This incorporates the catch-up effect initial GDP might have on GDP growth for poor countries.

**Info: Growth Theory**
For our plot above, we used simple average real GDP growth. In most growth theories, it is suggested that poorer countries can profit from a catch-up effect and tend to grow faster than rich countries do, for example due to diminishing returns of deployed capital. (An idea that was pioneered in a large part by the work of Sollow (1956).) This effect could have distorted our plot and, as a consequence, our conclusions. That's why we are testing for the influence of initial GDP on GDP growth in the regression at hand.

The next example demonstrates how to use the command `mutate()` from the *dplyr* package to get the logarithm of GDP per capita for the first year for each country. First, we group `Int.data` by the variable `country` using the `group_by()` function. Then, we pass this data frame to `mutate()` as a first argument and determine in the second argument what should be done with each group: We want to compute the logarithm of the first entry for gdpcap for each country, which we manage by using the function `first()`, and assign it to `ln_gdp_start`. The `mutate()` function creates a new data frame, containing the old data and the additional column `ln_gdp_start` which we save as `Int.data`. Last, we calculate the mean of the logarithm of the GDP in 1960 for the 18 different countries and assign it to `ln_gdp_start_avg`.

```
# Use mutate to add the new column 'ln_gdp_start' to the data frame
# 'Int.data':
Int.data = mutate(group_by(Int.data, country),
                  ln_gdp_start = log(first(gdpcap)))

# Compute the average of initial GDP:
ln_gdp_start_avg = mean(Int.data$ln_gdp_start)

# show the variable:
ln_gdp_start_avg
```

```
## [1] 9.351129
```

Now it's your turn, again. Use `mutate()` to add a column to Int.data with the name `gdp_start_year` which depicts the product of `(ln_gdp_start - ln_gdp_start_avg)` and `(year - 1960)`. Note that you have to group the data by country, again.

```
Int.data = mutate(group_by(Int.data, country),
                  gdp_start_year = (ln_gdp_start - ln_gdp_start_avg) *
                                   (year - 1960))
```

The purpose of the variable `gdp_start_year` is to test for the influence initial GDP has on GDP growth. It is negative if initial GDP is below average and vice versa. (Again, the use of the *Data Explorer* can be useful.)

Now we can do the regression that we looked at above. We use the same robust standard errors as we did in *Exercise 2.1*. Besides implementing a couple of arguments of `showreg()` to make the result optically more appealing, this time we are going to use the html-output of `showreg()`. All we have to do is set the argument `output` to "html".

```
# First, we do the regression, also using the new variable we created:
Int.reg2 = lm(log(gdpcap) ~ log(1 - mtr) + year + country + gdp_start_year,
              data = Int.data)

# Then, we show that regression using 'showreg()'. Again, we set
# 'robust = TRUE' and 'robust.type = "HC1"'. We omit all coefficients that
# include the word 'country'. We also add a title, a caption and custom
# coefficient names. Then we set 'output = "html"' in order to receive
# html-output:
showreg(list(Int.reg2), custom.model.names = c("1960-2010"), digits = 3,
       robust = TRUE, robust.type = "HC1", omit.coef = "country",
       title = "Effects of top MTR on real GDP per capita", caption =
       "Including a time trend, country fixed effects and a GDP time trend",
       custom.coef.names =
         c(NA, "(1 - r)", "time trend", rep(NA, 17), "time x initial GDP"),
       output="html")
```

**Effects of top MTR on real GDP per capita**

| | 1960-2010 |
|---|---|
| (Intercept) | -35.682*** |
| | (0.601) |
| (1 - r) | -0.018 |
| | (0.011) |
| time trend | 0.023*** |
| | (0.000) |
| time x initial GDP | -0.012*** |
| | (0.001) |
| $R^2$ | 0.948 |
| Adj. $R^2$ | 0.947 |
| Num. obs. | 918 |

***p < 0.001, **p < 0.01, *p < 0.05

Including a time trend, country fixed effects and a GDP time trend

Again, this is a screenshot of the *html* output that is displayed in the problem set[11]. While this may not look more elegant in the problem set, I decided to demonstrate this option to give the user an idea of the variety of possibilities that the function `showreg` from the package *regtools* by Kranz (2014b) offers.

We can see that the effect of top net-of-tax rates $\beta$ is slightly negative, but it's not significant. That is why, in their paper, Piketty et al. "conservatively conclude that low top tax rates do not have any detectable positive impact on GDP per capita".
At this point, we can state that it seems unlikely that the major part of the overall elasticity $e$ of top income shares to top net-of-tax rates is made up by the supply-side elasticity $e_1$. This leads Piketty et al. to the statement that our "macro-level analysis appears to be more consistent with our nonconventional bargaining model than with the standard model used in tax analysis".

While reverse causality still remains a possibility, we can find in our international data that in all countries that faced a large top marginal tax rate cut and a surge in top 1 percent income share, the increases in top income tend to follow the tax rate cuts.

The fact that we looked at various different countries diminishes the probability that we are underestimating the supply-side elasticity $e_1$ due to external effects that are independent from top tax rates and influence GDP.
However, we still face the possibility that deviations of GDP growth from its trend not caused by top tax rates are correlated with the evolution of top tax rates. For example, countries could tend to cut top tax rates when they expect GDP growth to slow down. This, again, would lead to an underestimation of $e_1$. This is tackled by Piketty et al. by also looking at micro evidence where this problem is not an issue. (Find out more at the bottom of this page.)

[11] As this is *html* output the argument "results" of the code chunk has to be set to "'asis'".

Note: Piketty et al. also use this type of regression to get a simple estimate of the different elasticities which will also influence our calculations of the optimal marginal top tax rate $r^*$. To find out how this can be done, take a look at the *info* section below.

**Info: Estimations of elasticities**

If we do the last regression on the later period from 1981 to 2010 we obtain a $\beta$ of $0.008$. If we then assume that the top 1 percent income share $\pi$ at the beginning equals $0.1$ this can be translated into $e_1 = 0.08$ because in the supply-side model we have $\beta = \pi \cdot e_1$. If we assumed that there are unrecorded tax avoidance effects we would have $e_1 + e_2 = 0.08$. Of course, the standard error around this estimate is large but even if we assumed that $e_1 + e_2$ was twice as high it would still make up less than $40\%$ of our estimated elasticity $e = 0.5$. This leads to a compensation-bargaining elasticity $e_3$ of at least $0.3$.

*In their paper, Piketty et al. go on to analyse microeconomic evidence from both the US and OECD countries. They rule out some problems and counter arguments and consolidate the findings that we had. We will not further devote ourselves to the analysis of that data. However, you can find a short summary of their analysis in the info-box below.*

**Info: Microeconomic evidence**

While we focus on macroeconomic evidence in this problem set, Piketty et al. also consider empirical micro-level evidence from CEO pay "to directly investigate whether CEO pay responds to top tax rates and whether this response is due to bargaining effects rather than productive effort".

First, they look at data of CEO pay and firm performance from the US for the period from 1970 to 2010. They consider a model by Bertrand and Mullainathan (2001) that describes how CEO pay depends on luck and performance and test whether rewards for luck are higher when top tax rates are lower. They find strong evidence of pay for luck and that pay for luck has been stronger in the low tax period after 1986 than in the high tax period before 1987. Combined with their finding that workers' wages do not show any pay for luck, this is consistent with the bargaining model where a high top income elasticity is due to a high compensation-bargaining elasticity.

Then, they go on to analyse a dataset that combines detailed information on CEO pay from 14 countries in 2006, information on stock ownership, firm performance and firm governance. They find a strong, negative correlation between total pay and top tax rates across countries. This relation stays almost the same when controlling for firm performance, implying that nearly none of the effect of taxes on pay is due to changes in productivity. In well-governed firms, the effect of top net-of-tax rates on CEO pay is smaller which is also consistent with the bargaining model, as well-governed firms should be more able to control rent-seeking channels. Finally, they find that the elasticity for bonuses and equity is extremely large while for salaries it is rather small.

In conclusion, their additional results from microeconomic evidence support the compensation-bargaining hypothesis and are consistent with their findings from the macroeconomic evidence that we have reproduced.

*You can find the analysis of microeconomic evidence on pages 29-37 of the paper.*

As it was stated earlier, the analysis of microeconomic evidence that Piketty et al. conduct could not be included in this problem set because the data they use is confidential. To find out how the data for their paper was constructed, see the *PDF* file "2011_281AEJPol_ReadMe"[12].

After we have analysed empirical evidence, in the next exercise we are going to take a look at an economic model with which we are going to try to describe our findings. This will help us to derive possible optimal top tax rates and policy implications.

*This partial exercise is based on pages 21f. and 26-28. There you can also find further regressions Piketty et al. perform (Including a bootstrapping method which they deploy as a robustness check on their results). The results and implications stay the same.*

[12] All the files of the article the problem set is based on can be found in the folder "Article_Piketty_et_al". For the mentioned readme file open the folder "data".

**Exercise 3.1 Theory - Studying a model**

*In this exercise we are going to analyse a model that includes the three elasticities we have already talked about. This exercise is not mandatory and you can skip it if you don't feel like reading it through. Just go to Exercise 3.2 where, at the beginning, you can find a short summary of the most important findings of the model that is developed on this page.*
*But it's still recommended to work through this theoretical exercise at some point because it gives you more profound economic insights into the topic.*

As promised in the beginning, now we take a deeper dive into the theoretical model Piketty et al. develop in their paper. While they consider a single model for each of the three scenarios (supply-side, tax-avoidance, compensation-bargaining) and combine them in a final formula for optimal top tax rate $r^*$, we are going to develop a model that combines the three income responses and will lead to the same results.
Further derivations and information are given in several *info*-sections.

I designed this problem set so that it is possible to skip *Exercise 3.1* because some users might not be interested in the theoretical framework that Piketty et al. developed and only want to work through empirical evidence and its implication. This is also why I start *Exercise 3.2* with a short summary of the result of this exercise.
In their paper, Piketty et al. first introduce the standard supply-side model that explains the response of top incomes to top tax rates solely with changes in productivity effort. Then they combine this model with tax avoidance and compensation bargaining separately. Finally, they derive an optimal top tax rate but do not explicitly show how they combine the two separate models. For the problem set I developed a way to combine these three aspects in one model right from the start and derive the same results from it that Piketty et al. deduce. This is mainly to make the problem set more compact and more clear for the user as all implications are derived from the same model.
Another value I added to the problem set is that I derived all formulas step by step while in the paper some of this is done implicitly. Again, this was done to offer the user a comprehensible problem set. Most of these derivations are wrapped up in *info* boxes to preserve compactness.

We denote by $z$ taxable earnings and by $T(z)$ the non-linear tax schedule. Our top marginal tax rate above a given threshold $\bar{z}$ is still $r$. The government maximizes the social welfare function $W = \int G(u_i)\, dv(i)$, where $G()$ is increasing concave, $u_i$ is the utility of individual $i$ and $dv(i)$ is the density mass of people of individuals of type $i$. We assume that the average social marginal welfare weight for top bracket income earners is zero. This means that the government sets top MTR $r$ to maximize tax revenues among top incomes.

**Info: Welfare weight zero**
The social marginal welfare weight on individual $i$ is defined as $g_i = G'(u_i)u_{ci}/p$. For top bracket income earners, we assume that this is on average zero. This allows us to obtain an upper bound on the optimal top tax rate. However, as we will discuss later, we can also integrate a positive social welfare weight $g$ on marginal consumption of top bracket earners.

**a) Deriving elasticities**

Individuals can put effort $h_i(y)$ to generate real income $y$. (This emulates the standard supply-side model as in Saez (2001).)
To describe bargaining compensation effects, individual $i$ receives a fraction $\eta$ of his real product $y$ and can put effort $k_i(\eta)$ into increasing compensation. This leads to income $\eta y$ and to bargained earnings $b = (\eta - 1)y$.
We also include a simple model that captures tax avoidance: We consider an environment where top earners can shelter a part $x$ of their real income $\eta y$ so that taxable income is $z = \eta y - x$. $z$ is taxed at marginal tax rate $r$ while sheltered income is taxed at $t$ lower then $r$. Sheltering an amount of income $x$ causes costs $d_i(x)$ for the individual $i$.
(Note: All three effort functions $h_i(y)$, $k_i(\eta)$ and $d_i(x)$ are $\geq 0$, increasing and convex in their respective variable and normalized so that $h_i'(0) = k_i'(0) = d_i'(0) = 0$.)

Adding all this together leads to a utility function for individual $i$:

$$u_i(c, y, x, \eta) = c - h_i(y) - d_i(x) - k_i(\eta),$$

where $c = \eta y - rz - tx + R$ is disposable after tax income and $R = r\bar{z} - T(\bar{z})$ denotes the virtual income coming out of the non-linear tax schedule.

If we differentiate the function $u_i(c, y, x, \eta)$ with respect to the variables that can be influenced by the individual we get the following three first order conditions:

- $h_i'(y) = \eta(1 - r)$,
  describing the optimal real income effort choice $h_i'(y)$ in dependency of top marginal tax rate $r$ and share $\eta$,
- $k_i'(\eta) = y(1 - r)$,
  describing the optimal bargaining effort $k_i'(\eta)$ in dependency of top MTR $r$ and real income $y$ and
- $d_i'(x) = r - t$,
  describing the optimal income sheltering effort $d_i'(x)$ in dependency of the margin between top MTR $r$ and sheltered income tax $t$.

**Info: Derivation 1**
Inserting the formula for disposable after tax income $c$ into $u_i(c, y, x, \eta)$ and using $z = \eta y - x$ and $R = r\bar{z} - T(\bar{z})$, we get:

$$u_i(y, x, \eta) = \eta y - r(\eta y - x) - tx + r\bar{z} - T(\bar{z}) - h_i(y) - d_i(x) - k_i(\eta)$$

Differentiating $u_i(y, x, \eta)$ with respect to real income $y$ and setting to zero, we obtain:

$$\frac{\partial u_i}{\partial y} = \eta - r\eta - h_i'(y) \stackrel{!}{=} 0$$

$$\Leftrightarrow h_i'(y) = \eta(1 - r)$$

Differentiating $u_i(y, x, \eta)$ with respect to fraction $\eta$ and setting to zero, we obtain:

$$\frac{\partial u_i}{\partial \eta} = y - ry - k_i'(\eta) \stackrel{!}{=} 0$$
$$\Leftrightarrow k_i'(\eta) = y(1 - r)$$

Differentiating $u_i(y, x, \eta)$ with respect to sheltered income $x$ and setting to zero, we obtain:

$$\frac{\partial u_i}{\partial x} = r - t - d_i'(x) \stackrel{!}{=} 0$$
$$\Leftrightarrow d_i'(x) = r - t$$

Let's do a quiz to internalize these first findings.

```
# Question 1: How do the three different efforts react to an increase in r
# (ceteris paribus)?
# Assign each of the three variables which depict the three different
# efforts to the values "+" if you think effort will also increase, "-" if
# you think effort will decrease and "=" if you think effort will stay the
# same:
# h = "???"
# k = "???"
# d = "???"
# Assign "+", "-" or "=" to the three effort variables and remove the
# comment
#
# Question 2: What happens to sheltered income x if we set r = t?
# a) x = 0
# b) x = y
# c) Can't say
# answer = "???" # Assign "a", "b" or "c" to 'answer' and remove the comment

h = "-"
k = "-"
d = "+"
answer = "a"
```

Because in this theoretical exercise there are no chunks with code to edit for the user, I included two rather comprehensive quizzes to keep the user engaged.

We can now also express the three elasticities that we used throughout this problem set so far in formulas:

$$e_1 = \frac{\frac{dy}{y}}{\frac{d(1-r)}{1-r}}$$

is the **supply-side elasticity**, also called first elasticity or real elasticity. It describes by how much real economic productivity changes if the net-of-tax rate changes. Piketty et al. call this the only real limiting factor for optimal top tax rates as both tax avoidance and compensation bargaining is harmful.

$$e_2 = s_2 \cdot e = \frac{\frac{dx}{z}}{\frac{d(r-t)}{1-r}}$$

describes the **tax-avoidance elasticity**. It depicts by how much sheltered income changes in relation to reported taxable income if the margin between top MTR and tax on sheltered income changes in relation to the top net-of-tax rate. Piketty et al. define tax avoidance as "changes in reported income due to changes in the form of compensation but not in the total level of compensation". Tax avoidance depends vitally on the tax system and its opportunities for avoidance. In our model, $e_2$ makes up the fraction $s_2$ of $e$.

**Info: Tax avoidance**
According to Stiglitz (1985), tax avoidance can be distinguished into three main categories (For which we are going to mention some examples named by Piketty et al. in their paper):

- Postponement of taxes: An example can be reductions in current cash compensation for deferred compensation such as stock-options or future pensions
- Tax arbitrage across individuals facing different tax brackets (or the same individual facing different marginal tax rates at different times): Tax induced transactions between individuals within a family, etc.
- Tax arbitrage across income streams facing different tax treatment: Reducing cash compensation for fringe benefits, changes in the form of business organization such as shifting profits from the individual income tax base to the corporate tax base and re-characterization of ordinary income into tax favoured capital gains (Arguably the main channel for tax avoidance and the one that we concentrated on in *Exercise 1*)

The fourth category that we include in our definition of tax avoidance is tax evasion such as using offshore accounts. As Slemrod & Yitzhaki (2002) put it, "the distinguishing characteristic of evasion is illegality".

Tax-avoidance opportunities can either be pure creations of the tax system such as different taxation of different income forms or real enforcement constraints such as the problem of taxing profits from informal businesses. As we assume that the social welfare weight on top incomes equals zero, tax avoidance is harmful and the government aims to close tax-avoidance opportunities (A fact that we are also going to look at mathematically at the end of this page).

$$e_3 = s_3 \cdot e = \frac{\frac{db}{z}}{\frac{d(1-r)}{1-r}}$$

is the **compensation-bargaining elasticity**, or just bargaining elasticity. It describes by how much bargained earnings change in relation to reported taxable earnings if net-of-tax rate changes. Bargained earnings describe the difference between pay and marginal economic product. In our model, $e_3$ can be calculated as $e$ times the fraction $s_3$.

**Info: Compensation bargaining**
Pay of top earners may not be equal to marginal economic product due to their power to influence compensation committees. Compared to lower income earners, top earners generally have more opportunities to influence their pay (Especially when you think about executives in big companies). When net-of-tax rates decrease, top income earners have a higher incentive to positively influence their income. Bargaining of top income earners is harmful as the bargaining aggregate is zero and higher bargained incomes of top earners comes at the expense of lower incomes and because the government sets a higher welfare weight on lower incomes.
This so-called rent-seeking externality that bargaining creates will also be incorporated in the optimal tax that we are about to derive.

**Info: Derivation 2**
The fraction of the behavioural response of taxable income $z$ to top net-of-tax rate $1-r$ due to tax avoidance $s_2$ can be expressed as

$$s_2 = \frac{\frac{dx}{d(r-t)}}{\frac{dy}{d(1-r)} + \frac{db}{d(1-r)} + \frac{dx}{d(r-t)}}$$

which can be, using $d(r-t) = -d(1-r)$ and $y + b - x = z$, transformed into

$$\frac{\frac{dx}{d(r-t)}}{\frac{dz}{d(1-r)}}.$$

Now, we can proof that $s_2 \cdot e = e_2$ because

$$s_2 \cdot e = \frac{\frac{dx}{d(r-t)}}{\frac{dz}{d(1-r)}} \cdot \frac{\frac{dz}{z}}{\frac{d(1-r)}{1-r}} = \frac{\frac{dx}{z}}{\frac{d(r-t)}{1-r}} = e_2.$$

The same can be shown for $s_3$ and $e_3$, as

$$s_3 = \frac{\frac{db}{d(1-r)}}{\frac{dy}{d(1-r)} + \frac{db}{d(1-r)} + \frac{dx}{d(r-t)}} = \frac{\frac{db}{d(1-r)}}{\frac{dz}{d(1-r)}}$$

and

$$s_3 \cdot e = \frac{\frac{db}{d(1-r)}}{\frac{dz}{d(1-r)}} \cdot \frac{\frac{dz}{z}}{\frac{d(1-r)}{1-r}} = \frac{\frac{dx}{z}}{\frac{d(r-t)}{1-r}} = e_3.$$

These three elasticity components combine to the overall elasticity $e$ in the following form:

$$e = \frac{y}{z}e_1 + e_2 + e_3 = \frac{\frac{dz}{z}}{\frac{d(1-r)}{1-r}}.$$

Again, we can see in the formula what we have already discussed earlier: The overall elasticity $e$ describes by how much taxable income increases if the top net-of-tax rate increases.

**Info: Derivation 3**
Proof, that $e = \frac{y}{z}e_1 + e_2 + e_3$:

$$\frac{y}{z}e_1 + e_2 + e_3 = \frac{y}{z} \cdot \frac{\frac{dy}{y}}{\frac{d(1-r)}{1-r}} + \frac{\frac{dx}{z}}{\frac{d(r-t)}{1-r}} + \frac{\frac{db}{z}}{\frac{d(1-r)}{1-r}} = \frac{\frac{dy+db-dx}{z}}{\frac{d(1-r)}{1-r}} = \frac{\frac{dz}{z}}{\frac{d(1-r)}{1-r}} = e$$

## b) Deriving the optimal top tax rate

Now, we want to take a look at the tax earnings of the government and see how we can develop an optimal top marginal tax rate from our findings.

The government aims to maximize taxes collected from taxpayers in the top bracket which are, in our model, described by the function

$$T = r(z - \bar{z}) + tx - N \cdot E(b).$$

$E(b)$ depicts the average bargained earnings in the economy. $N$ describes the size of the population. As we assume that the government adjusts the demogrant $-T(0)$ to fully offset $E(b)$, we have to subtract $E(b)$ from taxes collected from top bracket tax payers. (Note that this is a strong assumption. You can see a little discussion on this topic in the *info*-section below.)

**Info: Offset of bargaining effects**
When we looked at compensation bargaining, we considered the fact that bargaining always comes at the expense of others. If top earners earn an additional bargained earning $b$ then $E(b) = b/N$ denotes the average bargained earnings in the economy. Piketty et al. assume that the government offsets $E(b)$ by adjusting the demogrant $-T(0)$ of its non-linear income tax schedule for $N$ individuals. It de facto subsidizes taxpayers before applying its income tax to compensate for lost income due to bargaining of top earners.

This simplification is possible because we assume that $E(b)$ affects all individuals uniformly. This might not be the case. Piketty et al. show in the appendix of their paper that their model is also applicable if we assume that bargaining effects other individuals differently and that only top earners bargain.

Another assumption that Piketty et al. implicitly made is that the government is able to observe $b$. (This is not trivial and can arguably be seen as questionable: Compensation bargaining is only possible because companies can't perfectly observe real economic product of their employees. It is hard to see why then the government should possess this ability.)

If we differentiate tax collected from taxpayers in the top bracket $T$ with respect to top MTR $r$ and apply the findings we gathered on elasticities above, the **optimal top marginal tax rate** can then be calculated as:

$$r^* = \frac{1 + t \cdot a \cdot e_2 + a \cdot e_3}{1 + a \cdot e},$$

where $a$ depicts the Pareto parameter of the top tail of the distribution $z/(z - \bar{z}) > 1$. (Note that it is not trivial to know $a$, $t$, $e$, $e_2$ and $e_3$. This is why we have looked at extensive empirical evidence to be able to make estimations of optimal tax rates using this formula.)

**Info: Derivation 4**

If we differentiate $T = r(z - \bar{z}) + tx - N \cdot E(b)$ with respect to top MTR $r$ and set to zero, we obtain

$$\frac{\partial T}{\partial r} = (z - \bar{z}) + r\frac{\partial z}{\partial r} + t\frac{\partial x}{\partial r} - \frac{\partial (N \cdot E(b))}{\partial r} \stackrel{!}{=} 0.$$

This, we can translate into

$$z - \bar{z} = r\frac{dz}{d(1-r)} - t\frac{dx}{d(r-t)} - \frac{db}{d(1-r)},$$

using $dr = -d(1-r)$, $dr = d(r-t)$ and $N \cdot E(b) = b$.

If we use the formulas for the fractions $s_2$ and $s_3$ (or, alternatively that $s_2 \cdot e = e_2$ and $s_3 \cdot e = e_3$) we can derive

$$s_2\frac{dz}{d(1-r)} = \frac{dx}{d(r-t)}$$

and

$$s_3\frac{dz}{d(1-r)} = \frac{db}{d(1-r)}.$$

If we utilize these two terms in our equation we get

$$z - \bar{z} = (r - t \cdot s_2 - s_3)\frac{dz}{d(1-r)}.$$

Employing the formula for overall elasticity $e$, we can rearrange this to

$$\frac{z - \bar{z}}{z} = \frac{e(r - t \cdot s_2 - s_3)}{(1-r)}$$

and because

$$\frac{z - \bar{z}}{z} = \frac{1}{a}$$

we get

$$r \cdot a \cdot e - t \cdot a \cdot s_2 \cdot e - a \cdot s_3 \cdot e = 1 - r$$

which we can, using $s_2 \cdot e = e_2$ and $s_3 \cdot e = e_3$, finally rearrange to

$$r^* = \frac{1 + t \cdot a \cdot e_2 + a \cdot e_3}{1 + a \cdot e}.$$

With the derivation of the optimal top tax rate $r^*$ we have reached the main proposition of Piketty et al.'s model. We want to take this opportunity to do another short quiz to analyse how changes in the different variables have an impact on this optimal top tax rate.

```
# Question 1: If we experience an increase in t, how does that influence our
# optimal top tax rate r* if we assume that e2 > 0?
# a) r* increases, too
# b) r* decreases
# c) r* stays the same
# answer1 = "???" # Assign "a", "b" or "c" to 'answer1' and remove the
# comment
#
# Question 2: How does an increase in e1 influence r* (ceteris paribus)?
# a) r* increases, too
# b) r* decreases
# c) r* stays the same
# answer2 = "???" # Assign "a", "b" or "c" to 'answer2' and remove the
# comment
#
# Question 3: How does an increase in the bargaining elasticity e3 (e.g. due
# to more bargaining possibilities) influence r* (ceteris paribus)?
# a) r* increases, too
# b) r* decreases
# c) r* stays the same
# answer3 = "???" # Assign "a", "b" or "c" to 'answer3' and remove the
# comment

answer1 = "a"
answer2 = "b"
answer3 = "a"
```

Great, now we have worked through the model that we will use to explain the observations we can make in our data. Before you move on, you can click the *info*-sections below if you want to find out how we can prove that in the optimal case $t^* = r^*$ or if you are interested in what would happen in our model if we had a positive social welfare weight on consumption of top earners.

**Info: Optimal tax policy**

So far, we assumed that the tax on sheltered income $t$ is given. But how should $t$ be set if that wasn't the case and the government tries to maximize tax revenue from top earners $T$?

If we differentiate $T = r(z - \bar{z}) + tx - N \cdot E(b)$ with respect to $t$ and set to zero and use the equation $z = \eta y - x$, we get

$$\frac{\partial T}{\partial t} = -r\frac{\partial x}{\partial t} + x + t\frac{\partial x}{\partial t} \stackrel{!}{=} 0.$$

We can use the fact that

$$\frac{dx}{dt} = -\frac{dx}{d(r-t)}$$

to obtain

$$x + (r-t)\frac{dx}{d(r-t)} = 0.$$

Because, by definition, $x \geq 0$, $r \geq t$ and $\dfrac{dx}{d(r-t)} \geq 0$, this can only be true if $t^* = r^*$.

This means that, if the government has the power to change tax on sheltered income $t$, the government should set $t$ equal to top MTR $r$ to "improve efficiency and its ability to tax top incomes" as Piketty et al. (2014) find in their paper. This is consistent with our statement earlier that tax avoidance is harmful and the government aims to close tax-avoidance opportunities.

**Info: positive social welfare weights**

In the model we looked at so far the government put a social welfare weight of zero on marginal consumption of top earners. If we also allow for positive weights $0 \leq g < 1$ our formula for optimal top marginal tax rate becomes

$$r^* = \frac{1 - g + t \cdot a \cdot e_2 + a \cdot e_3}{1 - g + a \cdot e}.$$

You can see that optimal top tax rate $r^*$ decreases with increasing welfare weight $g$:

$$\frac{\partial r^*}{\partial g} = \frac{a(t \cdot e_2 + e_3 - e)}{(1 - g + a \cdot e)^2} \leq 0$$

because $e \geq t \cdot e_2 + e_3$.

This is intuitive, as an increasing $g$ leads to an increasing positive welfare impact of top incomes above the top tax threshold $\bar{z}$ and therefore the government has less incentive to reallocate those top incomes.

*The derivations and interpretations in this partial exercise follow pages 5-16 of the paper.*

### Exercise 3.2 Theory - Developing a function

In this exercise, we want to implement a short function that can calculate the optimal top marginal tax rate $r^*$ that we have derived in *Exercise 3.1*:

$$r^* = \frac{1 + t \cdot a \cdot e_2 + a \cdot e_3}{1 + a \cdot e}$$

**Info: If you have skipped Exercise 3.1**
In *Exercise 3.1* we devised formulas for the utility function of top earners and for the function that describes tax revenue the government receives from top earners. We were able to use these formulas to gradually develop the just mentioned formula for an optimal top marginal tax rate the government can set to create maximum tax earnings among top earners and therefore create optimal welfare:

$$r^* = \frac{1 + t \cdot a \cdot e_2 + a \cdot e_3}{1 + a \cdot e}$$

The elasticities that are used in this formula should be familiar from the first exercise:

- $e$, as usual, depicts our overall elasticity of top taxable incomes to top marginal tax rate $r$ above a certain income threshold $\bar{z}$. It holds that $e = (y/z)e_1 + e_2 + e_3$.
- $(y/z)e_1$ is the supply-side elasticity multiplied by the ratio between real product $y$ and real taxable income $z$. It describes that fraction of the overall change of taxable income that is due to a change in real economic productivity.
- $e_2$ denotes the tax-avoidance elasticity, i.e. that part of the overall elasticity that is due to sheltering income from ordinary income taxes.
- $e_3$ depicts the compensation-bargaining elasticity that describes that part of the overall elasticity which is due to compensation bargaining effects.

We can also see two new variables:

- $t$ denotes tax on sheltered income. (Sheltered income is income that isn't taxed with the ordinary marginal income tax rate $r$ and thus describes tax avoidance effects.)
- $a$ depicts $z/(z - \bar{z})$, the Pareto parameter of the top tail of the distribution.

We want to start with an example. You probably know the pq-formula:

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p}{2} - q},$$

which is used to calculate the two zeros of a monic quadratic equation of the form

$$x^2 + p \cdot x + q = 0.$$

We can translate the calculations for the two solutions of the pq-formula into the following code:

```
-p/2 + sqrt((p/2)^2 - q)
-p/2 - sqrt((p/2)^2 - q)
```

We use the function `sqrt()` to compute the square root and the syntax `x^y` to get `x` to the power of `y`.

In the next code chunk you can see how the pq-formula is implemented as a function that returns the two solutions using the R-computations we just introduced. Run the code to upload the function into our working environment.

```r
pq.formula = function(p, q){
  return(c(-p/2 + sqrt((p/2)^2 - q), -p/2 - sqrt((p/2)^2 - q)))
}
```

You can see that the function gets the two variables p and q passed in the header of the function. Then, it computes the two zeros and concatenates them using the function `c()`. The result, an array with two values, is returned with the command `return()`.

Let's see how we can use that function. Suppose, we have the following quadratic equation:

$$x^2 + 2x - 3 = 0$$

Now we can use `pq.formula` with the values $p = 2$ and $q = -3$ to compute the two zeros of this equation:

```r
pq.formula(2, -3)
```

```
## [1]  1 -3
```

The function returns the two zeros $x_1 = 1$ and $x_2 = -3$. You can see that the function uses the variables we pass it in the same order as they are defined in the header of the function.
Throughout this problem set we have already used a couple of different functions, like `read.table()`, `log()` or `filter()`, but now we also know how to program our own functions that we can then use in the same way.

**Info: What about no zeros?**
As you might have noticed, our function only works when there actually are zeros to be computed. If there are none, our function will return error messages. To avoid this we can use an if-clause that only runs its code if its condition is evaluated as true:

```r
pq.formula = function(p, q){
  if((p/2)^2 - q > 0){
    return(c(-p/2 + sqrt((p/2)^2 - q), -p/2 - sqrt((p/2)^2 - q)))
  }
}
```

But as we don't need this for our own tax function, we can ignore this.

Of course, this *info* box is not relevant for the understanding of the paper itself, but it prevents the user from possibly being confused and also serves to show the user how an *if* statement can be deployed in *R*.

Let's utilize this to implement a function for our optimal top marginal tax rate $r^*$ step by step.

Again, take a look at the formula that we have derived:

$$r^* = \frac{1 + t \cdot a \cdot e_2 + a \cdot e_3}{1 + a \cdot e}$$

As you can see we need the following variables to be able to obtain a solution: $e$, $e_2$, $e_3$, $t$ and $a$.

Now, implement the formula for $r^*$ as code, using the variables `e`, `e2`, `e3`, `t` and `a` (You don't have to write a function, yet).

```
# First, we set all the variables to default values. We are going to put in
# more meaningful numbers later.
e = 0.3
e2 = 0.1
e3 = 0.1
t = 0.2
a = 1.2

# Then, you can type in your formula here:

(1+t*a*e2+a*e3)/(1+a*e)
```

```
## [1] 0.8411765
```

Because in *RTutor* it is not possible to check the correctness of a formula, I set the variables that are used to default values in order to be able to examine if the user has typed in the right code[13].

Now, we write a function named `top.tax` that gets passed the variables we need in the following order: $e$, $e_2$, $e_3$, $t$, $a$. Then, the function returns the optimal top tax rate, using the computation that we wrote in the last code chunk.

```
top.tax = function(e,e2,e3,t,a){
  return((1+t*a*e2+a*e3)/(1+a*e))
}
```

For this code chunk I use an example instead of letting the user implement the function. It would probably be pedagogically more sensible to let the user do the work here. However, this is not possible because *RTutor* is also not able to check the correctness of a function. And as

[13] *RTutor* mainly checks the output of a code chunk. And because formulas without using actual values and functions have no real output, it is not possible to check them.

the function `top.tax()` is used further on in the problem set it is crucial that it is implemented correctly.

Now it's time to test our function. Let us assume that the variables have the following values: $e = 0.35$, $e_2 = 0.1$, $e_3 = 0.05$, $t = 0.25$ and $a = 1.8$. Call the function `top.tax` with these values. Make sure to use the values in the right order or, alternatively, assign the variables manually: `(e2 = 0.1, e = 0.35, ...)`.

```
top.tax(0.35, 0.1, 0.05, 0.25, 1.8)
```

```
## [1] 0.696319
```

As you can see, the values we used for the different elasticities, taxes and parameters lead to an optimal top tax rate of about $70\%$.

But of course, these numbers are only fictional. In the next exercise we are going to use numbers that should be closer to reality.

**Exercise 4 Four Scenarios**

*You must have finished Exercise 3.2 to be able to work on this exercise.*

After gathering implications from economic evidence and developing a theoretical framework, we finally want to apply our findings and derive implications for tax policy. For this purpose we look at four possible scenarios that Piketty et al. consider that explain the surges in top income shares and the coinciding cuts in top marginal tax rates that we could observe in our data. In addition, we use the function we have developed in *Exercise 3.2* to calculate possible corresponding optimal top tax rates that maximize tax revenue from top bracket earners.

Note that in the last two exercises I do not use *info* boxes. This is due to the fact that I think, in contrast to the first three exercises, all the information is substantially important for the user to understand.

**a) Skill-Biased-Technological-Change Scenario**

In this scenario, the increase in top income shares over the last recent decades is explained by the fact that technological progress has been skill-biased and has favoured top earners with high qualifications. This means that highly talented individuals were able to increase their economic productivity while the rest of the workers' productivity more or less stagnated. This effect could coincide with cuts in top marginal tax rates and lead to an overestimation of top income elasticity $e$.

Piketty et al. come to the conclusion that "this scenario cannot explain why only some OECD countries have experienced a surge in top income shares and why that surge has been highly correlated with the drop in top marginal tax rates" as all OECD countries have experienced similar technological change. They also point out that this scenario cannot explain the strong correlation between CEO pay and top net-of-tax rates in their microeconomic evidence.

**b) Supply-Side Scenario**

This scenario describes a situation where decreasing top tax rates lead to higher top income shares because top earners raise their work effort. This corresponds to the classic supply-side story that was proposed by Lindsey (1987) and Feldstein (1995).

In their paper, Piketty et al. consider an overall elasticity $e = 0.5$, a figure that is consistent with the estimates that we were able to produce with our statistical analysis of US and international evidence.
In the pure supply-side scenario, there is no tax avoidance and no compensation bargaining. This leads to $e = e_1 = 0.5$ and $e_2 = e_3 = 0$. As Pareto coefficient $a$ we are going to use the one that Piketty et al. identify for the United States: $a = 1.5$. And for alternative tax on sheltered income $t$ we are going to use $t = 0.2$. (A value that is hard to estimate, as Piketty et al. admit, but in our case is perhaps reasonable. Note that in this case $t$ can be chosen randomly, because it is multiplied with $e_2 \cdot a = 0$.)

Now, call the function `top.tax(e,e2,e3,t,a)` with the corresponding values to find out which optimal top tax rate our model suggests for this scenario. Assign your result to the variable `top_tax_supply` and show it.

```
# assign 'top_tax_supply'

# show the variable

top_tax_supply = top.tax(0.5, 0, 0, 0.2, 1.5)
top_tax_supply
```

```
## [1] 0.5714286
```

As you can see, in our model the values that we chose for the supply-side scenario lead to an optimal top tax rate $r^* = 57\%$.

As mentioned earlier, for *Exercise 4*, the user must have already worked through *Exercise 3.2*. This is because, for the calculations of the optimal top tax rate for the three elasticity scenarios, the function `top.tax()` is used and therefore must have been implemented. To use a variable that was created in another exercise, you have to make use of the possibility to import it in the settings[14].
Note that for this exercise it would not be necessary to save the calculated values in variables. But it still makes sense to let the user do it as it can be a good habit to always store one's results to make them accessible and retraceable.

**c) Tax-Avoidance Scenario**

This scenario posits that the high overall top income elasticity $e$ is due to a large avoidance elasticity $e_2$. That means that increasing tax rates lead top earners to exploit loopholes which allow them to report less ordinary income while not necessarily reducing their economic product.

This translates into a high tax-avoidance elasticity $e_2$. Piketty et al. consider a scenario with the following values: overall elasticity $e = 0.5$, tax-avoidance elasticity $e_2 = 0.3$, compensation-bargaining elasticity $e_3 = 0$, Pareto coefficient $a = 1.5$ and tax on sheltered income $t = 0.2$.
Again, call the function `top.tax(e,e2,e3,t,a)` with these values and assign the result to the variable `top_tax_avoidance`.

```
# assign 'top_tax_avoidance'

# show the variable

top_tax_avoidance = top.tax(0.5, 0.3, 0, 0.2, 1.5)
top_tax_avoidance
```

[14] To see how this can be done, take a look at the lines 1402-1404 of the file "Optimal Taxation of Top Labor Incomes_sol.Rmd".

```
## [1] 0.6228571
```

We can observe that, now, we obtain a revenue-maximizing top tax rate of $r^* = 62\%$ which is higher than in the standard supply-side scenario. This is due to the fiscal externality that is created by the fact that some of the income that is shifted away from ordinary income can still be taxed as sheltered income with the tax rate $t$.

Note: In their paper, Piketty et al. show that if the government is able to broaden the base and increase tax enforcement to reduce tax-avoidance elasticity $e_2$ to $0.1$ (while keeping $e_1 = 0.2$ constant), in our model we obtain an optimal top tax rate $r^*$ of $71\%$. This makes it clear that, in the tax-avoidance scenario, the government first has to close loopholes before being able to further increase top tax rates.

### d) Compensation-Bargaining Scenario

In this scenario, lower top tax rates lead to higher top income shares because they motivate top earners to bargain more aggressively for higher compensation. In their paper, Piketty et al. come to the conclusion that this scenario is most consistent with the evidence that we also looked at in this problem set and suggest the following estimates: $e = 0.5$, $e_2 = 0$ and $e_3 = 0.3$. The Pareto coefficient $a = 1.5$ and tax on sheltered income $t = 0.2$ again stay the same. Piketty et al. call these estimates their "preferred estimates" while also mentioning that "these estimates are not sharply identified but they illustrate the critical importance of the decomposition of the overall elasticity into three elasticities."

One last time, translate these estimates into a top marginal tax rate $r^*$ utilizing our function `top.tax()`. Assign the result to the variable `top_tax_bargaining` and show it.

```
# assign 'top_tax_bargaining'

# show the variable

top_tax_bargaining = top.tax(0.5, 0, 0.3, 0.2, 1.5)
top_tax_bargaining
```

```
## [1] 0.8285714
```

In our model, you can observe that this scenario leads to the by far highest optimal top tax rate $r^*$. This is due to the fact that we assumed that compensation bargaining is harmful and hence is tried to be avoided by the government.

In this exercise we could see what top tax rates the different scenarios and their respective composition of top income elasticity $e$ lead to in the model of Piketty et al. In the next and also last exercise we are going to further discuss these findings.

*This scenario analysis is based on pages 4 and 37-39.*

## Exercise 5 Conclusion

### a) Scenario Summary

The following table sums up the results from our calculations from *Exercise 4*:

| Supply-Side Scenario | Tax-Avoidance Scenario | | Compensation-Bargaining Scenario |
|---|---|---|---|
| | narrow tax base | broad tax base | |
| $e_1$=0.5<br>$e_2$=0.0<br>$e_3$=0.0 | $e_1$=0.2<br>$e_2$=0.3<br>$e_3$=0.0 | $e_1$=0.2<br>$e_2$=0.1<br>$e_3$=0.0 | $e_1$=0.2<br>$e_2$=0.0<br>$e_3$=0.3 |
| **r*=57%** | **r*=62%** | **r*=71%** | **r*=83%** |

This table depicts the optimal top tax rates $r^*$ that result from the application of the model of Piketty et al.
Remember that for all estimations in this table we used $t = 0.2$ and $a = 1.5$.

This is a screenshot of *html* output that is directly created by *R Markdown* when *RTutor* is started in the browser[15]. I decided to use *html* code directly to obtain a more elegant result and to avoid another example code chunk that would distract the user from the conclusion. Note that the *html* code for this table was partly created by using the website TablesGenerator.com (n.d.).

We can understand from this table that it is very important for the government to know the composition of the top income elasticity $e$ in order to be able to set the optimal top marginal tax rate $r^*$ that maximizes top tax returns. Using the Pareto coefficient $a = 1.5$ and the tax on sheltered income $t = 0.2$ that Piketty et al. identify for the US, the results vary from $57\%$ under the supply-side scenario to up to $83\%$ under the compensation bargaining scenario. Note that top MTR is recently slightly below $40\%$ in the US according to the Internal Revenue Service (2013).
The table also shows the importance of tax base broadening under the tax-avoidance scenario that was mentioned in *Exercise 4*.

If we follow the implications from our evidence and the conclusions that Piketty et al. draw from them, the compensation-bargaining scenario is the most likely one to explain the strong correlation between top net-of-tax rates and top income shares. (Remember that US evidence suggested a small tax avoidance elasticity $e_2$ and both US and international evidence indicated that supply-side elasticity $e_1$ is not the main driving factor of income responses. These results are also consistent with the findings Piketty et al. derive from their microeconomic evidence that are summarized at the end of *Exercise 2.2*.) This would mean that the surges of top income shares in some countries over the last decades can mainly be explained by top earners bargaining more aggressively for higher compensation.
If we consider this scenario we obtain high top tax rates. For example, in the US we get an optimal top tax rate of about $83\%$ which is more than double the recent tax rate.

[15] See lines 1471-1503 of the file "Optimal Taxation of Top Labor Incomes_sol.Rmd".

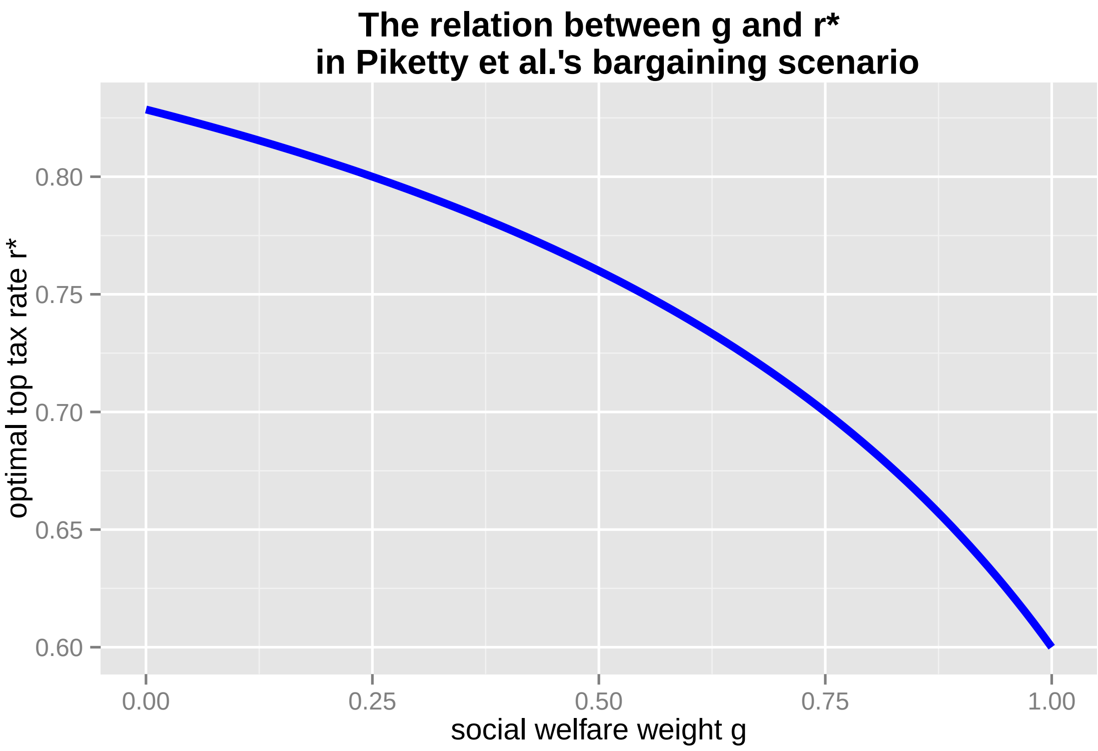
(In the pure compensation-bargaining scenario, where bargaining elasticity $e_3 = e$, we would in fact receive an optimal top tax rate $r^* = 100\%$.)

## b) Further Outlook

Just as Piketty et al. do in their paper, we have focused on a revenue maximizing top tax rate and considered a marginal social welfare weight on top bracket earners $g$ of zero. But welfare weights may vary between societies. This may again depend on which of the three elasticities seems to be the most important: If we consider a pure supply-side scenario, "pay is fair by definition and hence a zero weight can only be justified by strong redistributive motives", as Piketty et al. state. And also in other scenarios it might not be trivial to assume only revenue maximizing motives.

The following code chunk is designed to show you how social welfare weight on top earners $g \in [0, 1)$ can influence optimal top tax rate $r^*$ in the compensation bargaining scenario that we have considered. First, we implement the function `top.tax.g(e,e2,e3,t,a,g)` which is an extension of the function that we have developed in *Exercise 3.2* and follows the formula that Piketty et al. derive in their paper that also includes $g$ (See the last *info* box of *Exercise 3.1*). Then, we use the function `stat_function()` from the *ggplot* package to plot `top.tax.g(e,e2,e3,t,a,g)` in dependency of $g$ while using all the other values from our compensation-bargaining scenario ($e = 0.5$, $e_2 = 0$, $e_3 = 0.3$, $a = 1.5$ and $t = 0.2$). (Again, this code serves mainly as an incitation - Just try to comprehend it and maybe test it for your own purposes.)

```
# First, we implement the function 'top.tax.g()':
top.tax.g = function(e,e2,e3,t,a,g){
  return((1-g+t*a*e2+a*e3)/(1-g+a*e))
}

# We use 'stat_function()' to plot r* for values of g between 0 and 1. Note
# that we have to first create a dummy data set. Again, we use some of the
# arguments of 'ggplot()' to make the plot better-looking:
p = ggplot(data.frame(x = c(0, 1)), aes(x)) +
    stat_function(fun = function(x) top.tax.g(0.5,0,0.3,0.2,1.5,x),size=1.5,
                  colour = "blue") + xlab("social welfare weight g") +
    ylab("optimal top tax rate r*") +
    ggtitle("The relation between g and r*
    in Piketty et al.'s bargaining scenario") +
    theme(plot.title = element_text(size = 14, face = "bold"))

# Finally, we show the plot:
p
```




You can see that, depending on the social welfare weight $g$, in the compensation-bargaining scenario with $e_3 = 0.3$ the optimal marginal top tax rate $r^*$ varies between $60\%$ for high welfare weights and $83\%$ for low welfare weights. This example serves to demonstrate that in order for a government to determine the optimal top tax rate $r^*$ it not only needs to know the decomposition of the income elasticity $e$ but also has to identify the welfare weight on top incomes $g$ that prevails in the society.

### Some space for testing

At the end of this problem set, here is one last code chunk where you can play around a little bit with different values for the variables that we have discussed throughout the different exercises. Simply change the values the function `top.tax.g()` is called with and see what optimal top tax rate $r^*$ the model of Piketty et al. suggests. (Note that according to Piketty et al., the prevailing Pareto coefficient for many European countries is $a = 2$.)

```
top.tax.g(e=0.5, e2=0, e3=0.3, t=0.2, a=2, g=0)
```

```
## [1] 0.8
```

I decided to include this code chunk to give the user the opportunity to test some own values and see what optimal tax rates are suggested by the model that was studied in the problem set. To make sure that the user is able to change the results of the code chunk without receiving error messages the code is embedded in a *task_notest* environment[16].

[16] Compare lines 1551-1553 of the file "Optimal Taxation of Top Labor Incomes_sol.Rmd".

**Congratulations, you have successfully worked through this problem set!**
We have studied a framework that describes top income responses to top marginal tax rates and derived optimal top tax rates for different scenarios. We also looked at macroeconomic evidence to comprehend which of these scenarios possibly describes real economic phenomenons. And, of course, we looked at various R solutions that help us tackling such problems.
Hopefully, you have enjoyed this problem set and have gathered a lot of different skills that motivate you to approach other economic issues.

*For this last exercise, compare pages 38-40 of the paper.*

---

## References

### Academic Papers and Books


- Bertrand, Marianne, & Sendhil Mullainathan. (2001). "Are CEOs Rewarded for Luck? The Ones without Principals Are". Quarterly Journal of Economics 116 (3), p. 901-32. DOI: 10.1162/00335530152466269

- Feldstein, Martin. (1995). "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act". Journal of Political Economy 103 (3), p. 551-72.

- Hayashi, Fumio. (2000). "Econometrics". Princeton University Press. ISBN: 978-0-691-01018-2

- Internal Revenue Service (Department of the Treasury). (2013). "Your Federal Income Tax". Publication 17, Catalog Number 10311G.

- Lewis-Beck, Michael S., Alan E. Bryman, & Tim F. Futing Liao. (2003). "The Sage encyclopedia of social science research methods". Sage Publications. ISBN: 978-0-761-92363-3

- Lindsey, Lawrence B. (1987). "Individual Taxpayer Response to Tax Cuts: 1982-1984: With Implications for the Revenue Maximizing Tax Rate". Journal of Public Economics 33 (2), p. 173-206. DOI: 10.1016/0047-2727(87)90073-9

- MacKinnon, James G., & Halbert White. (1985). "Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties". Journal of Econometrics, 29(3), p. 305-325.

- Newey, Whitney, & Kenneth D. West. (1987). "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix"." Econometrica, 55, issue 3, p. 703-08.

- Piketty, Thomas, Emmanuel Saez, & Stefanie Stantcheva. (2014). "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities". American Economic Journal: Economic Policy, 6(1), p. 230-71. DOI: 10.1257/pol.6.1.230

- Saez, Emmanuel. (2001). "Using Elasticities to Derive Optimal Income Tax Rates". Review of Economic Studies 68 (1), p. 205-29. DOI: 10.3386/w7628

- Slemrod, Joel, & Shlomo Yitzhaki. (2002). "Tax avoidance, evasion, and administration". Handbook of public economics, 3, p. 1423-1470. DOI: 10.1016/S1573-4420(02)80026-X
- Solow, Robert M. (1956). "A contribution to the theory of economic growth". The Quarterly Journal of Economics, 70 (1), p. 65-94. DOI: 10.2307/1884513
- Stiglitz, Joseph E. (1985). "The General Theory of Tax Avoidance". National Tax Journal, Volume XXXVIII, No. 3, September 1985, p. 325-338. DOI: 10.3386/w1868
- Wickham, Hadley. (2009). "ggplot2: Elegant Graphics for Data Analysis". Springer. ISBN: 978-0-387-98140-6
- Wilkinson, Leland. (2005). "The Grammar of Graphics". Springer. ISBN: 978-0-387-28695-2


### R and R packages


- Arnold, Jeffrey B. (2014). ggthemes: Extra themes, scales and geoms for ggplot. R package version 2.1.0. http://CRAN.R-project.org/package=ggthemes
- Auguie, Baptiste. (2012). gridExtra: functions in Grid graphics. R package version 0.9.1. http://CRAN.R-project.org/package=gridExtra
- Gesmann, Markus, & Diego de Castillo. (2011). googleVis: Using the Google Visualisation API with R. The R Journal, 3(2):40-44, December 2011.
- Kranz, Sebastian. (2014). regtools: Tools for presenting regressions results. R package version 0.1.
- Kranz, Sebastian. (2014). RTutor: Creating R exercises with automatic assessment of student's solutions. R package version 0.7.1.
- R Core Team. (2014). foreign: Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, Weka, dBase, .... R package version 0.8-61. http://CRAN.R-project.org/package=foreign
- R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/
- Wickham, Hadley. (2009). ggplot2: Elegant graphics for data analysis. Springer New York.
- Wickham, Hadley, & Romain Francois. (2014). dplyr: A Grammar of Data Manipulation. R package version 0.3.0.2. http://CRAN.R-project.org/package=dplyr
- Zeileis, Achim. (2004). sandwich: Econometric Computing with HC and HAC Covariance Matrix Estimators. Journal of Statistical Software 11(10), 1-17. URL http://www.jstatsoft.org/v11/i10/.


### Websites


- Vigen, Tyler. (2014). Spurious Correlations. http://tylervigen.com/.

At the end of the problem set I reference all the publications, packages and websites that are used in the problem set. Note that in the chapter "References" of this thesis I only name the sources that are used in the commentary of this chapter or in one of the other chapters.

## Critical Appraisal of Piketty et al.'s Paper

At least since the release of Piketty's book "Capital in the Twenty-First Century" (2014), inequality can be considered one of the most popular topics in modern economics. One fact that fuels discussions concerning this issue is arguably the surge of top income shares in many western countries. Examples of large taxpayers responses in the 1980s (cf. Slemrod (1995 and 1996)) has led to a broad literature analysing the relationship between top tax rates cuts and top income share boosts. For one thing, surges in top incomes have been explained with technological change and globalization favouring high skilled top earners[17]. For another thing, one explanation is the standard supply-side story as proposed by Lindsey (1987) and Feldstein (1995) where the increase in top incomes is a result of an increase in economic productivity as a response to lower marginal tax rates. Another possible response channel of top incomes is tax avoidance which describes the phenomenon that tax payers shift income away from ordinary taxed income when ordinary tax rates rise[18]. Another narrative is that decreasing taxes have led top earners to bargain for higher incomes.
As I summarize and present in the problem set, in their paper "Top Taxation of Top Labor Incomes: A Tale of three Elasticities", Piketty et al. introduce a model following a welfarist approach[19] that captures supply-side, tax-avoidance and compensation-bargaining responses of top incomes to top marginal tax rates. The macro and microeconomic evidence both from the US and from a set of OECD countries they present is consistent with the bargaining scenario. According to them, their findings suggest low tax-avoidance effects and show that top tax rate cuts are not associated with higher economic growth. They also argue against the skill-biased explanation because international and micro evidence speak against it.

There are some points in the paper of Piketty et al. that might have room for improvement that I am going to discuss in the following.
Naturally, as the different elasticities that are utilized in the paper are estimated using evidence from certain countries and certain years, these elasticities might not be valid for the whole possible range of tax rates[20]. Therefore, Piketty et al.'s calculations of the optimal top tax rates might be biased in a certain direction.
Piketty et al. include tax evasion in their definition of tax avoidance. However, tax evasion leads, naturally, to no tax return when in contrast pure tax avoidance still leads to earnings via the tax on sheltered income $t$. You can also assume that tax evasion is not statistically recorded. This could mean that they underestimate tax-avoidance elasticity $e_2$ as they define it and consequently overestimate compensation-bargaining elasticity $e_3$ and supply-side elasticity $e_1$, as lower taxes could have lead to higher ordinary incomes due to less tax evasion. However, this would not change their formula for optimal top tax rates because tax evasion creates no fiscal externality as pure tax avoidance does[21].
Some estimates of the paper that are most probably estimated accurately could be documented in more detail. For example the tax on sheltered income $t$ they use for the calculations of optimal top marginal tax rates in the last chapter is 20%. They admit that this value is hard to estimate.

[17] Gabaix and Landier (2006), for example, explain the increase in CEO pay with the increase in firm value in a neoclassical model.

[18] See Slemrod and Yitzhaki (2002) for an introduction of this approach to the standard model.

[19] This means that the government's objective is to maximize a weighted sum of individual utilities. For a list of alternative approaches see Piketty and Saez (2013).

[20] As Saez, Slemrod and Giertz (2012) put it for income elasticities that they use: "An elasticity estimated around the current tax system may not apply to a hypothetical large tax change". (Note, however, that due to the fact that Piketty et al. look at evidence from several decades they were able to analyse rather large tax changes.)

[21] Additionally, some might argue that today there are more tax evasion possibilities due to better access to offshore accounts etc. and that it is therefore unlikely that tax evasion has declined. This argument is also brought forward by Piketty et al.

However, they give no explanation why they use this exact estimation[22]. As Pareto coefficient $a$ they use 1.5 because according to them this is the prevailing coefficient in the US[23].
In their paper, Piketty et al. assume that bargained earnings $b$ come uniformly at the expense of all other individuals and that they are offset by the government by adjusting the demogrant $-T(0)$. While they show in the appendix of the paper that it is not a necessity for the model that bargaining effects individuals uniformly they do not discuss whether it is possible for the government to actually observe $b$. In the problem set I hint to the fact that this can be seen as questionable[24] [25]. If we assume, however, that the government sets the same welfare weight on the incomes that are reduced due to bargaining of top earners that it does on tax earnings from top earners, the implications of the model and the formula for optimal top tax rates $r^*$ stay the same. This is due to the fact that we could then again include $N \cdot E(b) = b$ in the formula for tax earnings from top earners $T$ and to the fact that $b$ is not a part of the formula for $r^*$.
Lastly, Piketty et al. detect in their international evidence that income elasticities seem to be different in different countries which leads them to state that "the elasticity likely depends on the institutional set-up of each country"[26]. This implies that there are other factors that influence income elasticities and contradicts a universal cross-country elasticity of top incomes. However, Piketty et al. use their cross-country study to derive the figure 0.5 for their elasticity estimate that they use to calculate possible optimal top tax rates for different scenarios. A further discussion of possible institutional influencing factors could have helped to establish clarity in this matter.
Despite these issues, the main proposals of their paper, namely that the decomposition of income elasticity is important and that compensation bargaining seems to have a large influence on top income responses, remain mostly unaffected.

Taking everything into account, the paper of Piketty et al. provides a new theoretical framework that includes the three most important response channels of top earners concerning taxes and substantiate it with compelling evidence both macroeconomic and microeconomic. Piketty and Saez can be considered ranking among the leading researchers in the analysis of income inequality and optimal taxation and it will be thrilling to see how their paper influences future research.
As I mentioned earlier, I think the article "Optimal Taxation of Top Labor Incomes" and its currentness motivate students to work through the problem set and to gain deeper insights into the topic.

[22] In the US the tax on capital gains on assets held less than one year is 43.4% and on assets held longer than one year 23.8% if you include the 3.8% surtax imposed by the Patient Protection and Affordable Care Act, according to the Tax Foundation (2014).

[23] To be fair, a derivation of this value can be found in Diamond and Saez (2011)

[24] Compensation bargaining is only possible because companies cannot perfectly observe real economic product of their employees. It is hard to see why then the government should possess this ability.

[25] Note that I did not discuss most of the issues I name here in the problem set because I think that it would rather confuse the user than help him gaining a deeper insight.

[26] As they find, the same is true when you look at different periods of their international macroeconomic evidence.

## Academic Value Added

The problem set I created gives students the opportunity to deepen their knowledge in a couple of different areas.
First, it enables them to interactively learn the language *R* and the use of some of its additional packages. Utilizing step-by-step examples and giving them further information on the packages in *info* sections, I provide them with the tool set to learn the use of *R* by solving short code chunks. Additional examples show the user other methods that can help him to solve economical problems.
Second, the problem set presents a summary of the paper "Optimal Taxation of Top Labor Incomes" by Piketty et al. and gives the student the opportunity to reproduce the results that are obtained in the article step by step. Occasional quizzes test the user's understanding of the paper's theories and help him to stay involved.
Lastly, at several points throughout the problem set I give the user additional information that helps him acquire basic economic and econometric knowledge. Examples are an explanation of the relation between correlation and causality, an introduction to linear models and their fitting and an outlook to the use of different robust standard errors.

To keep the problem set as compact as possible, I explicitly combined the three different models that Piketty et al. only implicitly merge in their formula for optimal top tax rate. Additionally, I present the complete derivations that are to a large degree only hinted at in the paper. As it is explained in the commentary of the chapter above, this serves to make the problem set as comprehensible as possible.
For my problem set I changed the order that is used in the paper. In contrast to Piketty et al., in the problem set the student first analyses macroeconomic evidence of the influence of top tax rates on top income shares and then looks at a model that captures this relationship. I did this to make the problem set engaging right from the start. As it was mentioned above, for the analysis of the microeconomic evidence of the paper I offer the user a short summary as its data is not publicly available.
To make the calculation of the top tax rates for the different scenarios that is done in the paper more transparent and interesting for the user, *Exercise 3.2* of the problem set helps the user to develop a function that can perform this calculation. In *Exercise 4* this function is utilized by the user to obtain the same optimal tax rates for the different scenarios that are presented by Piketty et al. An extended version of this function is also used in the conclusion of the problem set, *Exercise 5*, to illustrate how different social welfare weights on top incomes influence optimal top tax rates. Finally, in the last code chunk of the problem set, the user is enabled to use the extended version to try out new values for the different variables that are discussed. This offers the user a simple possibility to play around with his own values and helps him to further understand the connections between the different variables and the optimal top tax rate that is implied by the model of Piketty et al.
At the beginning of *Exercise 2.1*, I introduce the user to the package *googleVis* by Gesman and Castillo (2011) and use its function to create a *Motion Chart* of the international macroeconomic data to give the user an interactive graphic tool to understand the relation between top tax rates and top income shares over the last half century.

Combining all these aspects, after working through the problem set the student should have understood the main implication of the paper of Piketty et al. but also should have acquired the economic knowledge and the *R* skill set to tackle other economic questions.

# Conclusion

*RTutor* is a convincing example of the vast possibilities that *R* and *R Markdown* have to offer. It provides a framework to create interactive tutorials that enable the student to acquire knowledge about *R*, economic methods in general and specific economic articles in particular at the same time.
Furthermore, the same is true for creating an *R* problem set. Developing exercises to help students understand the article "Optimal Taxation of Top Labor Incomes" also helped me gain deeper insights into the topic it embraces. Additionally, it was a enriching opportunity to consolidate my *R* skills and to consider new *R* packages that will help me in the future. I recommend trying to develop an interactive problem set and hope that my thesis is helpful for those who follow this recommendation.

*RTutor* is still in development. It will be intriguing to see what additional functions and possibilities it will have to to offer in the future. Hopefully, there will soon be a broad selection of problem sets that give students the chance to discover economic articles.

My problem set "Optimal Taxation of Top Labor Incomes" will soon be made available as a package on my GitHub account[27].

[27] You can find my account under the link https://github.com/JonasSend.

## References

**Diamond, Peter, & Emmanuel Saez**. (2011). The case for a progressive tax: from basic research to policy recommendations. The Journal of Economic Perspectives, 165-190. DOI: 10.1257/jep.25.4.165

**Feldstein, Martin**. (1995). "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act". Journal of Political Economy 103 (3), p. 551-72.

**Gabaix, Xavier, & Landier, Augustine**. (2006). Why has CEO pay increased so much? (No. w12365). National Bureau of Economic Research. DOI: 10.3386/w12365

**Gesmann, Markus, & Diego de Castillo**. (2011). googleVis: Using the Google Visualisation API with R. The R Journal, 3(2):40-44, December 2011.

**Kranz, Sebastian**. (2014a). "RTutor: Creating R exercises with automatic assessment of student's solutions". R package version 0.7.1.

**Kranz, Sebastian**. (2014b). "regtools: Tools for presenting regressions results". R package version 0.1.

**Lindsey, Lawrence B**. (1987). "Individual Taxpayer Response to Tax Cuts: 1982-1984: With Implications for the Revenue Maximizing Tax Rate". Journal of Public Economics 33 (2), p. 173-206. DOI: 10.1016/0047-2727(87)90073-9

**Mankiw, N. Gregory**. (2014). "Principles of Economics". Cengage Learning. ISBN: 978-1285165875

**Piketty, Thomas**. (2014). "Capital in the 21st Century". Harvard University Press. ISBN: 978-0674430006

**Piketty, Thomas, & Emmanuel Saez**. (2013). "Optimal labor income taxation". Handbook of Public Economics. Vol. 5: p. 391-474.

**Piketty, Thomas, Emmanuel Saez, & Stefanie Stantcheva**. (2014). "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities". American Economic Journal: Economic Policy, 6(1), p. 230-71. DOI: 10.1257/pol.6.1.230

**R Core Team**. (2014). "R: A language and environment for statistical computing". R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

**Saez, Emmanuel, Joel Slemrod, & Seth H. Giertz**. (2012). "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." Journal of Economic Literature 50 (1): 3-50.

**Slemrod, Joel, & Shlomo Yitzhaki**. (2002). "Tax avoidance, evasion, and administration". Handbook of public economics, 3, p. 1423-1470. DOI: 10.1016/S1573-4420(02)80026-X

**TablesGenerator.com**. (n.d.). "Tables Generator". Retrieved December 27, 2014, from http://www.tablesgenerator.com/

**Tax Foundation**. (2014). "Federal Capital Gains Tax Rates, 1988-2013". Retrieved December 29, 2014, from http://www.http://taxfoundation.org/article/federal-capital-gains-tax-rates-1988-2013

**Wickham, Hadley & Romain Francois**. (2014). dplyr: A Grammar of Data Manipulation. R package version 0.3.0.2. http://CRAN.R-project.org/package=dplyr