UNIVERSITY OF INNSBRUCK

MASTER THESIS

German Michel from fit to fat? Evidence on health effects of the "Bundeswehrreform"

Author:
Markus MAIER
00816200

Supervisor:
Prof.Martin HALLA

March 22, 2018

Prologue

The following master thesis is a quantitative analysis of the Bundeswehrreform in 2011, the idea for this paper is based on own experiences during military service and by suggestions from professor Martin Halla regarding the used data model. As the Bundeswehrreform was a huge issue in the German public, some parts of the paper are also inspired by different newspaper articles and the public discussion that took place prior to 2011. As the data searching process was difficult, a huge thanks to professor Martin Halla for his patience regarding the data retrieving process.
Also, the model turned out to posses some problems in terms of valid-solid controls, which delivered some odd results at the beginning. However, discussing the topic with fellow students offered a new perspective at the model and allowed me to include strong controls. Therefore, a special thanks to Felix Keller and his support along the way in regards of findings suitable controls. Also, the members of Martin Halla's course deserve a thanks due to their great feedback and suggestions.
Lastly, I like to thank everybody who either helped me with the thesis or gave me emotional support throughout the process of creating the work, thank you guys.

Abstract

Many papers on mandatory drafting found various effects among young conscripts, their health outcomes and chances in later life, however only a few investigated the effects of reforms abolishing such drafting. This paper investigates the health effects of the Bundeswehrreform of 2011 using SOEP (Social Economic Panel) data, with a Difference-in-Difference approach on male birth cohorts (female are the control) before and after the reform took place for those being 18 or 19 years old at the time of the reform.
However, I find no evidence that the Bundeswehrreform had any significant effect on health outcomes of young men. Neither the share of smokers nor the average Body-Mass-Index has changed with the end of military drafting.

Contents

List of Tables

List of Figures

1. Introduction

After the fall of Nazi Germany, Germany's military was abolished and large parts of young men during that period did not have to serve in the military. However, the events that followed the second World War such as the nuclear crisis and the separation of Western and Eastern Germany lead to a new threat and therefore Western Germany introduced its mandatory drafting system. The Bundeswehr should act as a defensive shield in case of an invasion from the parties of the Warsaw pact (including the Eastern German NVA) and it was designed as a mixture between drafted soldiers and those who would serve as professionals.

Until the German reunion in 1990, the military drafting system did not change too much (only serving time was adjusted) and was not seen as old fashioned or outdated in the public. But after the end of the Sowjetunion, the German public and politicians changed their opinion on the military drafting system quite often.
Most politicians[1] being in favour of the mandatory drafting argued, that a military with a mandatory draft system is cheaper and will help to recruit talented young men as future officers and soldiers. They also stated that a mandatory system might benefit young men in their future life as discipline and courage (for those doing civil work) are skills, which are also searched for by companies.

However, many German economists[2] argued, that this might be true in the short term and only looking at the business side of the reform, but in the long run the costs for young people are much higher than only the time spent at the military. Most of them either lose their first salary or might not be able to study directly after their service. Also, a professional army is more efficient and costs less in the long term as fewer soldiers are needed to maintain the same military strength.

[1]For example Reinhold Robbe (Wehrbeauftragter, German politician responsible for the needs of German Soldiers), `http://www.sueddeutsche.de/politik/sechs-monate-wehrdienst-rekruten-sitzen-nur-rum-1.152746`Sueddeutsche article

[2]Collection of different arguments/data and references on the topic, visit `https://www.diw.de/sixcms/detail.php?id=284379`DIW homepage

Because large parts of the public and politicians were in favour of the data provided by different economists, mandatory drafting was abolished in 2011 after nearly 20 years of discussion.

The reform that took place in 2011 is the main subject of my master thesis and provides a natural experiment for birth cohorts born before and after the reform, that would be drafted to serve either in the military or in a civil institution. It is the first paper to directly investigate the health effects of the reform in 2011 on young men in Germany, as most papers on the German Military focused on the behaviour during the military.

I use a Difference-in-Difference approach (Diff-in-Diff) to examine smoking and the Body-Mass-Index (BMI) behaviour of the reform for all men, as the drafts were non-random and do not allow for selection, while women are the control group. The Diff-in-Diff approach delivers the possibility to identify a casual effect of the reform on health outcomes and splits the young people in two groups to compare (before and after the reform). The men group before the reform consists about 75% of non-drafted men and of 25% drafted men, while the after group has no drafted men. As the used data are birth cohorts and lack the numbers at the exact year of the reform, the Diff-in-Diff design had to be preferred over a RDD (Regression Discontinuity Design).
The reason to include BMI and smoking as outcomes, is the overall effect on people in terms of their economic and overall situation, people who are fitter and smoke less, tend to be less prone to certain diseases (obesity and lung cancer), which should benefit the overall population.

While the focus lies on the health effects of the reform, I also investigate the exact moment when the reform really changed the behaviour of young men. The German-Drafting-System does not rely on random drafts, but on calling in every young men and check for their fitness and therefore young men are able to avoid drafting, especially if controls are less strict regarding an upcoming reform.

2. Background literature

Yet papers from recent years, which give some insight information on how service itself can affect young men and their health outcomes, tend to investigate random draft lottery systems, that weren't used by German military. Those draft lotteries normally assign a number to every young man and call in new conscripts up to a specific number, either completely random or in cohorts, which can't be observed by the young men.

The Vietnam draft lotteries in America are the most prominent ones. Siminski and Ville (2011) looked on long term mortality of draftees from Australia serving in Vietnam (data collected between 1994 and 2007) and found no significant increased mortality compared to peers, which could be linked with increased smoking or heavier men, if there had been an effect. Angrist et al. (2011) on the other hand looked at the long-term household stability among Vietnam War draftees in the United States and did not find any difference in unemployment and disability rates, which is a surprising effect as the observed soldiers were in actual combat.

Angrist (1990) also showed in an earlier paper that young conscripts earn less money directly after their service. Still in the long run their earnings converge to those of their peers and they do not suffer from negative health outcomes. Contrary to Angrist et al. (2011) findings, Autor et al. (2011), show that soldiers, who served longer throughout the Vietnam era, have a higher chance to suffer from PTSD (Post-traumatic Stress Disorder) and have either higher disability rates or are more prone to unemployment.

Further findings by Galiani et al. (2011), Hjalmarsson and Lindquist (2016), Siminski et al. (2013) and Rohlfs (2010) suggest, that conscripts in Argentina, Sweden, Australia and veterans from the Vietnam War are more violent and have a larger record of criminal behaviour and activities. Especially young men from lower income groups, according to Hjalmarsson and Lindquist (2016), are more prone to an increase of their criminal activity, although the amount of sick days at work decreased. Galiani et al. (2011) suggest, that gun violence and street gang related crimes are more common under conscripts than under their peers, due to the fact, that they had gun training and therefore have lower entry costs for gun related crimes. All those conscripts were in actual combat (Swedish conscripts excluded) and were drafted via a random draft lottery.

However, looking at nonrandom drafts and conscripts, that were not in combat and only had basic military training the findings are a bit different. Bauer et al. (2009), who used a RDD (Regression-Discontinuity-Design) design to look at the unemployment and wage rates of young German conscripts, that were more likely to be drafted for the new founded "Bundeswehr" back in 1950es concluded, that young men received a wage "premium" during their service. Still the wage "premium" effect got insignificant, when controlled for selected groups (low income and/or high income groups). Moreover, in another study by Bauer et al. (2014), young German men being drafted tend to finish their university careers more often than peers, that did not go to a military institution. But those young men being drafted for the German military are less open, have less agreeableness characteristics and show more neurotic behaviour, according to Schult and Sparfeldt (2015). Although the effects are small, they can also be found in other studies, regarding the same psychological spectrum, as Jackson et al. (2012) shows in his study about military psychology. The effects are also present in a wider sample of young men joining the army.

Furthermore, the general psychological health seems to be affected by military service itself. A paper by Bedard and Deschênes (2006) shows that veterans from the Second World War and Korean War suffer from PTSD (Post-traumatic Stress Disorder), but also from diseases such as lung cancer and heart failure, that can be traced back to military induced smoking. Same applies to conscripts, that weren't involved in the wars, regarding lung cancer and heart failure.
The overall findings on health outcomes and general characteristics of young men - either being drafted via random or non-random drafts - are different. Some do suffer systematically from psychological problems, while most do not have any long term problems with their service regarding their health or financial outcomes.

Therefore, the further extend of the paper will explain the German Drafting System and some reforms that took place before 2011 in detail. If you are familiar with the system itself, you can skip the section and look at the main part. There I briefly explain the design of a Difference-in-Difference approach and the parallel trend assumption. The main part consists of the empirical strategy, the data that is used for the analysis and the main analysis of the reform.

3. German Military Draft System and reforms prior to 2011

The military system in Germany has some unique features. First men can choose between serving in the army or being at a civil institution (usually a hospital or something related to it). Second, they have the option to deny military and civil service as a whole for certain severe reasons (only a small fraction) and or excluded after examination for physical or psychologically reasons, if they do not fit[3].

3.1. Important reforms affecting the experiment

After the reunion of 1990, German military had five important military reforms. The first reform took place simultaneously to the 2+4 contracts[4], which also fixed the troop strength and provided Germany a peace treaty. The 2+4 contracts were several contracts between Germany and former occupiers, such as France, the U.S.A, England, Russia and smaller nations or 3rd parties involved in the second World War.
Moreover the 2+4 contracts defined the final borders, after the inclusion of Eastern Germany and Germany's decline of former regions (Ostpreußen, Danzig etc.). The contracts also fixed other geographic and economic causalities still pending from the second World War such as the polish borders and payments that Germany had to pay countries, it invaded during the Second World War.
According to the 2+4 contracts reform period, young men over 18 years were drafted to serve for 12 months (15 month before 1990 in western Germany), more detailed information on the Bundeswehrhomepage[5]. After 1990 three more reforms took place, the serving duration was cut in 1996 to 10 months,

[3] `http://www.zentralstelle-kdv.de/pdf/09-Anlage-3-1.pdf`Drafting report, detailed information on drafting
[4] `http://www.krr-faq.net/2u4.htm`Contracts in detail
[5] `https://www.bundeswehr.de` Homepage of the German Military

2002 to 9 months and in 2010 to 6 month, shortly before the whole mandatory military service came to an end in 2011.

Table 1 on page 26 gives a small overview on the reforms and its outcomes for the mandatory serving time. As shown in Table 1, serving time was constantly decreasing before it was abolished in 2011, causing a crowding out effect for men that were mentally and physically fit enough to join the army. Table 2 on page 26, shows the numbers between 2000 and 2009. Although the supply of young men did not differ that much between the different birth cohorts, troop strength of conscripts declined and numbers for 2010 and 2011 are already missing, since the Bundeswehr was in some drastic change two years before the end of its mandatory drafting in 2011.
The fact, that drafting varied over time, is important for the main analysis, because the drafting was more sensitive with each reform. Young men that would have been in the drafting pick, prior to a troop strength reform, weren't selected at all with each new reform.

3.2. How does drafting work in detail?

Drafting in the German military is rather complex, because there are a lot of exception for certain people and people can to choose whether or not to serve in the military or at a social institution.
As a German teenager you normally receive a letter from your nearby "Kreiswehrersatzamt", which invites you to attend a medical and psychological screening. Before attending the screening process, you have the option to deny service completely or only military service. Denying service completely is only granted if you have two brothers, that already served in the army, if your ancestor had been victims of the holocaust during the Second World War or if your current mental or psychological state is unstable, so that an invitation makes no sense at all, further information is provided by the Kreiswehrersatzamt[6].

The process at the "Kreiswehrersatzamt" follows the guideline portrayed at Table 3 on page 27, young men are separated into 4 different "fitness" groups.

[6] `http://www.zentralstelle-kdv.de/pdf/09-Anlage-3-1.pdf`Drafting report, detailed information on drafting

Men in group one must be at least 179 cm and should not be taller than 195 cm, they also should have perfect eyesight, no allergies and no overweight. Men in group two can be under 179 cm and over 195 cm, but not over 206 cm and under 159 cm. People with minor medical issues, such as glasses, little overweight and allergic diseases are accepted. Men in group four are typically those, who face current medical problems, such as broken bones, depression or their parents need personal assistance. Group five consists of people with severe medical issues, such as being to overweight, too tall, or small, being blind, deaf or mute. Several medical or psychological problems do not add up, only the most severe issue counts for being selected or not.
The selection process delivers an inconsistent selected sample, I am using the term inconsistent, because only a fraction of the people that are fit for military service had to undergo the process at the "Kreiswehrersatzamt", see Table 1. Furthermore, the randomness occurs for all drafted men, both serving in the military and those denying military service, but who work for a civil institution, see Graphic 1 on page 39.

To explain the inconsistent selection in more detail, look at Table 4 on Page 4, which shows that the drafting in Germany was inconsistent and just got more selective with each new reform (also fit men weren't even invited sometimes, see Schneider (2003)). Especially after 2000 young men that were limited in their fitness, but could serve in the military or for the community, were treated as generally unfit, since category T3 in Table 27 was abolished and everyone from that category was put into T5. For further reading, the paper by Schneider (2003) gives a more detailed look onto the drafting process, showing moreover the inconsistency of it.

4. Data and Empirical Framework

The data used throughout the paper, is a modified data set of the Socio Economic Panel (SOEP long, data for birth years 1980-2014, version 29, 2016, doi:10.5684/soep.v30), which is a longitudinal data set with representative individual samples. Each household and individual have a specific id and can be traced throughout the data set. The SOEP data includes roughly 15.000 households with around 40.000 individuals, who are interviewed either each year or every second year. Smoking and BMI for example are examined every second year, the gathered structure has the following pattern; interviewed at age 18 (born 1980), interviewed at age 19 (born 1981), interviewed at age 18 (born 1982), interviewed at age 19 (born 1983) and so on.

To get reliable data for an Difference-in-Difference approach before and after the reform, the panel structured data had to be transformed into birth cohorts before and after the reform, which are equal in their social structure.
The transformed birth cohort data set contains young people born between 1980 to 1994, which were asked if they declined or accepted military work or were never drafted at all. The control group consists of females in the same birth cohorts, as military or community service was only mandatory for men. The data for the treated also includes young men, that didn't do military or community service. All men (treated) and women (control) are the same age in Data Set 1, either 18 or 19 years old, when they were asked about their BMI and smoking behaviour. The cut-off is either for people born at 1990 (reform of 2011) or 1984 (reform of the T groups, see section 2).

4.1. Theoretical framework

Diff-in-Diff Estimation: The approach which will be used throughout the thesis is the so-called Difference in Difference approach. To get a better overview, Table 5 on page 27 shows you the standard estimation strategy. The basic idea are two-time periods, before and after a treatment (in this case the Bundeswehrreform in 2011), with two different groups.
While one group is the treated (male) and the other one the control group (female), now the difference between each group in the specific time will be measured and afterwards the differences that have been calculated for both time periods will also be subtracted from one another, resulting in the Difference-in-Difference estimator. If everything has been done the right way, we will receive the Average-Treatment-Effect (ATE) of the treatment, which should look like the following.

$$\hat{\delta} = ((\overline{y_{11}} - \overline{y_{12}}) - (\overline{y_{21}} - \overline{y_{22}})), \tag{1}$$

, where delta is the estimator (Diff-in-Diff), y_{11} and y_{12} the difference of control and treated in the second period and y_{21} and y_{22} the difference of the two in the first period. The Difference-in-Difference approach is an easy way to identify the casual effect of a treatment inside a natural experiment, or in this case of a public policy intervention. As it measures the ATE it allows us to verify or neglect an effect of the Bundeswehrreform on health outcomes and therefore if the intervention had any consequences on young German men.

Parallel trend: The parallel trend assumption requires stable parallel pre-cut-off trends. Those are important to see if the control (female in my case) behaves in a similar way as the treatment group (male in my case). Also in order to provide consistency, the parallel trend assumption allows us to see autocorrelation and dips, see Ashenfelter (1978). To better understand the concept of the parallel trend, Figure 2 on page 40 shows a standard linear parallel trend. As one can easily see, the trends should be similar before the cut-off for the treatment and control group, which is required to have a reliable control.

Outcome variables for the parallel trend: The basic idea for my outcome variables in the parallel trend figure is to take the averages of the different BMIs, (detailed explanation of the BMI, see Eknoyan (2008)), for each birth cohort and sex (male as treated, female as control), which is calculated in the following fashion;

$$BMI = \frac{m}{h^2} \tag{2}$$

$$mean_{BMI} = \frac{\sum \frac{m}{h^2}}{n} \tag{3}$$

where m is the body mass in kilogram and h the body height in centimetres of each person and the mean BMI is calculated for each birth cohort and sex. The share of smokers, is basically just the mean of the dummy variable smoking over the different birth groups and sex.

4.2. Initial descriptive and graphical intuition

The first descriptive output is Table 6 on page 28, which shows a sample size for a selected data set (to determine the right cut-off). Men before 1990 are all

part of the "forced" labour group (mentioned in Figure 1 on page 41), while men after 1990 mostly weren't selected for community or military service).
However, the columns for forced labour sample sizes show, that due to the military reform effects, mentioned in Table 4 on page 27, the forced labour sample size declined from 1985 to 1990 rather rapidly, almost eight years before the system was abolished in 2011. The data here is the same, like the one that is used later in the main regression, however as each sample for each birth-cohort is matched for military and civil service, its sample size is much lower (because many don't give information about their service). The gathered information is important to determine the right cut-off of the different reforms that took place.
The table also goes in line with Figure 3 on page 41, which is an official statistic of the Bundeswehr and the conscript that were called to duty. As one can easily see, each reform forced a drop-in conscripts called for duty, and in 2010 the numbers were already really low.
The second graphical intuition can be observed at Figure 4 on page 42, which shows the correlation between sex (male 1, female 0) and BMI. The BMI increased over time for both sexes, while the difference between both sexes stayed quite similar, the confidence intervals seem to be the same with an exception in 1993. Looking at the same correlation for smoking in Figure 4 on page 43 we can observe a contrary trend, where smoking decreased and the share of smokers is more evenly distributed over both sexes.
Both figures deliver a preview of the parallel trend, which is the next step to look on and to determine pre-cut-off trends. Figure 6 on page 44 shows the parallel trend for both health outcomes (the birthyear cut-offs are at 1984 and 1990) of Data Set 1. Although the reform was between birth cohorts 1992 and 1993, only an insignificant fraction of young men served after 1990, because of the fading out effect of prior reforms (which could lead to an identifying problem).
Taking a deeper look at the figure we can see, that the left-hand side represents the average smoking behaviour, while the right-hand side shows the average BMI. The grey line stands for female, while the black one for male, a marker highlights each dip. I implemented both cut-offs into the graphs, to get a better overview and as a graphical reference point. It seems like both trends are stable and display a somewhat expectable trend. While the share of smokers declines for both sexes, the average BMI for both groups goes up. So overall, we can conclude that young men and women decreased their

consumption of cigarettes, but eat a lot more and got heavier.
The trends go in line with the findings from the WHO[7], which shows that young people tend to smoke less in most European countries (if you want to look up Germany, search by strg+f as pages aren't displayed in the report). Also, the BMI trend shows, that young people tend to eat more unhealthy and exercise less frequently or not at all, see Di Cesare and Bentham (2016) for worldwide trends on BMI.

In summary, the parallel trend seems to be realistic in terms of data and cohort behaviour, and therefore it full-fills the prior assumption of parallel linear trends and goes in line with the overall behaviour of young people and their share of smokers and mean of BMI during that time.

4.3. Model and output

As stated in the introduction, the model is a classic Diff-in-Diff model with separate cut-offs at 1984 and 1990 (because of the small sample size close to 1992), which will be compared via treated (male) and untreated (female) birth cohorts.

The formula is as the following:

$$Smoking_t^c = \beta_0 + \beta_1 * (time_t) + \beta_2 * (male^c) + \beta_3 * (did_t^c) + X'\beta + \epsilon_t^c \quad (4)$$
$$BMI_t^c = \beta_0 + \beta_1 * (time_t) + \beta_2 * (male^c) + \beta_3 * (did_t^c) + X'\beta + \epsilon_t^c \quad (5)$$

where:

$Smoking_t^c$= Smoking as a dummy variable for each person, either yes or no.

BMI_t^c= The individual BMI for every person.

$time_t$ = Time variable for the treatment (before 1990/1984 (0) or after 1990/1984 (1)).

[7]`http://apps.who.int/iris/bitstream/10665/156262/1/9789241564922_eng.pdf?ua=1`WHO report on health outcomes

$male^c$ = Treatment variable for the different sexes (male treat (1), female non-treat (0)).

did_t^c = Interaction variable between time and treatment (time*male)

X' = Vector of control variables such as region and migration background

β_0 = Constant of the regression

β_1 = First coefficient: The estimated post-treatment effect that is common to both the treatment and control groups;

β_2 = Second coefficient: The estimated difference between the treatment and control groups that is constant over time;

β_3 = Third coefficient: The Difference-in-Differences estimator of the treatment effect, the average effect on share of smokers and the BMI effect of each additional year added to the eligibility age;

β = Control coefficient: Effect of control variables (region, migration background an household income)

ϵ_t^c = Residual

The control variables I included are region, migration, and household income. Region is defined by East and West Germany, normally one would choose a level for each state, but as the sample size is not so big, smaller states would not be included and furthermore would only have small sample sizes, making the dummy somewhat weak for interpretation. Also, the differences between Eastern and Western Germany are still prominent in wage rates, and therefore the region dummy should be a better solution.

Migration is splintered into to three groups; natives (people that were born in Germany and have German parents), indirect migration (people that were born in Germany, but whose parents are immigrants) and people that aren't born in Germany. Migration should be a strong control variable, as most people migrating from different countries tend to be different from the native population. Every regression includes robust standard errors, I also tried to exclude unrealistic outliers for the BMI such as BMIs of 150 or above, which were falsely gathered in the interviews and delivered unrealistic trends.

Looking at the first output in Table 7 on 29 for age groups 18-19, we can easily see, that the cut-off at the 1990 birth cohorts offers a zero effect for smoking and the BMI, only the controls such as region, migration and log household income deliver a significant change for BMI and smoking, but as the Diff-in-Diff estimator is insignificant, it is not possible to identify the casual effect of those changes.

However, we can see, that the control variables are well specified; region has a positive effect (Western Germany is the zero dummy) and an increasing household income has a negative effect on smoking. The controls (household income excluded) for the outcome variable BMI seem to be insignificant overall and thus other factors do not influence BMI. This also goes in line with the strength of the model as the BIC is worse for the bigger model. But even with a significant result, the R^2 of the regression and therefore the explained variation would be rather small.

The fact that neither the BIC nor the R^2 provide a satisfying answer might arise from the fact that the intervention of the reform was not strong enough and therefore the model does not fit the data. Maybe school reforms, which took place in the late 2000es for both sexes had a greater effect on the unhealthy eating of young people, as-well as the several reforms on smoking were stronger in regards of reducing the smoking consumption.
To verify or neglect the zero effect of the reform, we can look at the second output Table 9 on page 30, which displays a slightly different result as the previous table, although the cut-off at the birth cohorts is 1984 this time. However, this time the share of smokers decreases and is significant, the controls for migration region and household income are all significant, which takes away the ability to identify the casual effect of the reform. An also the fit of the regression is once again not very high. We should also take into consideration, that the sample of birth-cohorts before 1985 is a lot smaller than between 1985-1990, which could lead to sample size problems regarding the effects of smoking on young men.

To address this problem the third output Table 9 on page 9 shows a regression over selected birth cohorts, where the more sensitive selected group (mentioned in the reform section, cohorts between 1985-1990) of young men is excluded and only young men before 1985 and after 1990 are taken into consideration. As seen by the output, the result for the Diff-in-Diff estimator is insignificant and again the controls for smoking are significant and shrink the BIC for the smoking regression, while they do not for the BMI regression.

Furthermore the 95% confidence interval of the Diff-in-Diff estimator for all three variation of the model for age groups 18-19 is close at zero for most estimates. Taking a look at Table 10 on page 32 we can see, that both the upper and lower bound for smoking is really narrow, while the confidence interval for the BMI is a lot wider, which probably results from measurement errors due to the fact that BMI is more biased (people are only interviewed and might lie about their weight), also BMI is a lot more unstable during time (remember the outcomes variables are only collected every second year). Smoking on the other hand isn't biased in terms of the measurement (only in terms of truth of the declaration), and should stay stable over time.

Therefore, we can conclude the Diff-in-Diff estimator delivers a zero effect for our analysis and at this stage we should conclude that the reform had no impact on young men regarding their health status.

4.3.1. Robustness checks, Data Set 2

To further evaluate the zero effect, I will look at the same models, but this time for the young people which were either 20 or 21 at the time of the interview, to see if there is an effect or not in another data set. Looking at the trend assumptions (Figure 7 on page 45) of Data Set 2, we can see a rather similar behaviour for the older age groups. The dips are behaving pretty much the same and we have a negative trend for smoking and a positive trend for BMI once again.

Looking at Table 11 on page 33, we can observe an increasing effect for smoking for the cut off at 1990 without birthyear-cohort selection. However, the effect is only significant in Data Set 2 and therefore we should be very cautious about effects, that could also be the consequences of differing sample sizes. The controls behave pretty much the same as in the previous model, with household income being the most prominent. Furthermore, the fit does not seem to be very good and the BIC only decreases slightly with the bigger models. Therefore, we can conclude that there is a zero effect on BMI of the reform in this table.

Table 12 on page 35 provides a surprising result, as it neglects the findings with the younger age groups at Table 8, as we can't observe any significant result at the 1984 cut-off. Therefore, we can see that in opposition to the findings of Data Set 1 the robustness check does not confirm an effect at the 1984 cut-off. Furthermore, the controls for region and household income are once again significant, showing us that income and region (Eastern, Western Germany) influence smoking the most and BMI is only strongly affected by income. This goes in line with general trends in the population, that poor people in Germany tend to smoke more and that poor people overall tend to be heavier.

While Table 11 and 12 includes all birth-cohort table 15 on page 37 only those between 1980 and 1984, as-well as 1990 till 1994. However, if we take a brief look at the output, the result goes in line with that of Table 11 on page 33, the diff-in-diff is once again significant, however only on the (p>0.5) margin. Therefore we should be careful about an identified casual effect. Maybe the observed significant result also originates from a measurement error collected in the older birth cohort.

Once again, the 95% confidence interval of the Diff-in-Diff estimator for all three variations of the model for age groups 20-21 behaves similar to those of Data Set 1. While taking a look at Table 14 on page 36 we can see, that both the upper and lower bound for smoking is really narrow, while the confidence interval for the BMI is once again a lot wider. As for the significant results for smoking the bounds are less narrow and both positive, however as stated above this might result from wrongly measured samples.

4.3.2. Robustness checks, Data Set 3 - reform in 2000 - older men

The second last part of the output section is a data set for people being interviewed at the age of 22 and 23, between 1980 and 1990. Figure 8 on page 46 shows the parallel trend for that specific group After looking at the parallel

trend for the older age groups between 1980 and 1990, we can look at the related regression table.
As shown by Table 15 on page 37, we once again see a zero effect result for the Diff-in-Diff estimator. In contrast to the findings of Data Set 1, the cut-off at 1984 for the ages 22 and 23 for each birth-cohort yields no significant result and in contrast to Data Set 2 we can't see an increase of smoking. As also shown by Table 16 on page 38 the confidence intervals are pretty much the same in terms of tightness, the bound for smoking is close to zero, while the BMI bounds are much wider, probably resulting from inconsistent gathered information.
So once again we can conclude, that the controls are significant for smoking and only household income for the BMI, yet as the estimator offers a zero effect result, the cut-off as an intervention did not affect young men and it is not possible to identify a casual effect.

5. Conclusion

Therefore the main conclusion of my thesis is, that the Bundeswehr reform and the end of the mandatory drafting had a zero effect on young men regarding their health outcomes, young men did not behave differently than their female counterparts and shared the same trends.
As seen by the data young people tend to smoke less over the last years but obesity and unhealthy diets were increasing. However, the end of military drafting did not change this kind of behaviour, neither in a negative nor a positive way. The significant result we received from parts of the Diff-in-Diff regressions and robustness checks were contradicted by other robustness checks, where the results were also insignificant and or by the existing measurement error. The fear of politicians and the general population, that people might become lazy and undisciplined in terms of their health were not confirmed, the strong increase of fat people seems to be a more general trend of the German society. Furthermore, the consumption of cigarettes does not seem to be affected by the reform itself, but by a more general trend of smoking as a drug becoming obsolete in the general public.

Still as there is a zero effect result, we cannot identify the casual effect why both trends occur, either smoking bans in commercials, increased taxes or forbidden smoking in public are the reasons for lowered smoking rates. The same applies for the BMI trend, as the military reforms had no significant effect on the BMI of young people, maybe other things such as lowered overall prices for food, higher availability and less mobility due to the digital revolution led to a higher rate of obesity among the population.

6. Outlook

Looking at the final results and the conclusion of the output, further investigations regarding the individual effects of a mandatory military reform might be a good idea. Especially the drinking habits (which have been dropped here, because of bad data quality) could be further investigated. To perform such a task, own quality data might be a first step, that features also interviews with post reform groups and their likely hood of being drafted, if the drafting would still be there.
Another part of the picture might be better controls, especially regarding serial correlation and overall instruments to avoid collision with other reforms, which is further discussed by Bertrand et al. (2002).

However, the main problem of my thesis, the measurement error (all men are included in the pre-cut-off) needs several adjustments, so it could be fixed for future investigations. For a starting point, I tried to select for after-cut-off groups in a later stage of my work, to get rid of an all-included men group (including those who were not drafted and civil workers), but as the interviews do not provide an after-cut-off screening, selecting for screened men results in selection bias. As mentioned above, own data with the collection of screening results would provide the opportunity to compare pre-cut off selected men with after cut-off selected men. But those measurements would be quite expensive as a trained doctor would be needed to gather the information of each interviewed individual, to put them in the different T groups.

Another idea to address the measurement issue would be a strong instrument such as psychological data, which I discussed in the literature section. The problem is, that the data provided in the SOEP Data Set does not match with my gathered observation or is not fully available, as the German Military did not use random draft lotteries, also the instrument would be much less valid

as in other studies with random drafts.
As we were able to observe, the confidence intervals of smoking had been rather narrow and smoking itself is stable among people over long periods of time, like drinking behaviour, it might be better to skip the BMI to avoid measurement errors and only collect smoking and drinking behaviour for own interviews.

Finally concluding, my paper provides an overall look on current trends for young people and their health outcomes, however due to measurement errors in the data and by design, the result of the Diff-in-Diff yields no valid result and therefore the reform needs further investigation, if one likes to know more about its effects.

A. Bibliography

References

Angrist, J. D. (1990). Lifetime earnings and the vietnam era draft lottery: Evidence from social security administrative records. *American Economic Review*, 80(3):271–300.

Angrist, J. D., Chen, S. H., and Song, J. (2011). Long-term consequences of vietnam-era conscription: New estimates using social security data. *American Economic Review*, 101(3):334–38.

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *The Review of Economics and Statistics*, 60(1):47–57.

Autor, D. H., Duggan, M. G., and Lyle, D. S. (2011). Battle scars? the puzzling decline in employment and rise in disability receipt among vietnam era veterans. *American Economic Review*, 101(3):339–44.

Bauer, T. K., Bender, S., Paloyo, A. R., and Schmidt, C. M. (2009). Evaluating the labor - market effects of compulsory military service - a regression-discontinuity approach. (141).

Bauer, T. K., Bender, S., Paloyo, A. R., and Schmidt, C. M. (2014). Do guns displace books? the impact of compulsory military service on educational attainment. *Economics Letters*, 124(3):513 – 515.

Bedard, K. and Deschênes, O. (2006). The long-term impact of military service on health: Evidence from world war ii and korean war veterans. *The American Economic Review*, 96(1):176–194.

Bertrand, M., Duflo, E., and Mullainathan, S. (2002). How much should we trust differences-in-differences estimates? (8841).

Di Cesare, M. and Bentham, J. (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *The Lancet*, 387(10026):1377–1396.

Eknoyan, G. (2008). Adolphe quetelet (1796–1874)—the average man and indices of obesity. *Nephrology Dialysis Transplantation*, 23(1):47–51.

Galiani, S., Rossi, M. A., and Schargrodsky, E. (2011). Conscription and crime: Evidence from the argentine draft lottery. *American Economic Journal: Applied Economics*, 3(2):119–36.

Hjalmarsson, R. and Lindquist, M. (2016). The Causal Effect of Military Conscription on Crime and the Labor Market. (11110).

Jackson, J. J., Thoemmes, F., Jonkmann, K., Lüdtke, O., and Trautwein, U. (2012). Military training and personality trait development: Does the military make the man, or does the man make the military? *Psychological Science*, 23(3):270–277.

Rohlfs, C. (2010). Does combat exposure make you a more violent or criminal person? evidence from the vietnam draft. *Journal of Human Resources*, 45(2):271–300.

Schneider, T. (2003). Wehr- und zivildienst in deutschland: Wer dient, wer nicht? / basic military and alternative service in germany: Who serves, who doesn't? *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics*, 223(5):603–622.

Schult, J. and Sparfeldt, J. R. (2015). Compulsory military service and personality development. *DIW-Soeppaper*, (751).

Siminski, P. and Ville, S. (2011). Long-run mortality effects of vietnam-era army service: Evidence from australia's conscription lotteries. *American Economic Review*, 101(3):345–49.

Siminski, P., Ville, S., and Paull, A. (2013). Does the Military Train Men to Be Violent Criminals? New Evidence from Australia's Conscription Lotteries. (7152).

B. Tables

Table 1: Bundeswehr reforms timeline

Time Period between a reform	Duration of service
01.04.1957 - 31.03.1962	12 Month
01.04.1962 - 30.06.1962	15 Month
01.07.1962 - 31.12.1972	18 Month
01.01.1973 - 30.09.1990	15 Month
01.10.1990 - 31.12.1995	12 Month
01.01.1996 - 31.12.2001	10 Month
01.01.2002 - 30.11.2010	9 Month
01.12.2010 - 01.07.2011	6 Month

Notes: Every period of time in the table had a specific duration of service for drafted soldiers. Source; `https://www.bundeswehr.de`.

Table 2: Conscripts serving from 2000-2009

Year	Amount of drafted men
2000	144.647
2001	129.441
2002	123.812
2003	102.600
2004	79.850
2005	68.428
2006	71.321
2007	67.823
2008	68.270
2009	68.304

Notes: Drafting decreased rapidly over the last decade.

Table 3: Bundeswehr recruiting process

Technical fitness status	Description of status
T1	Perfectly fit for military service
T2	Fit for military service
T3	Fitness for military is limited (abolished in 2000)
T4	Currently not fit for military service
T5	Permanent unfit for military service

Notes: T3 was excluded after 2000 and T4 only used for people that were in the military and excluded because of medical reasons during service. T1 and T2 were drafted in every reform, T5 wasn't drafted in any reform.

Table 4: Drafting Process between 2000-2008

Year	2000	2002	2004	2006	2008
Fit	86.28 %	83.9 %	79.2 %	61.6 %	54.1 %
Currently unfit	3.68%	3.35 %	2.9 %	9.22 %	29,18 %
Unfit	10.04 %	12.74 %	17.89 %	29.18 %	42.28 %

Notes: Fit are T1 and T2, currently unfit are T3 and T4 and T5 is unfit. As seen in the table, fit decreased while currently unfit and unfit increased.

Table 5: Difference-in-Difference overview

	Treatment group	Control group
Period 1	$\overline{Y_1^T}$	$\overline{Y_1^C}$
Period 2	$\overline{Y_2^T}$	$\overline{Y_2^C}$

Notes: The left side describes the Treatment group male, the right side the control group female. The difference between male and female and the two time periods delivers the average treatment effect.

Table 6: Men in Military and civil service over birth-cohorts in the used data sets

Birthyear	Men in military and civil service
1980	101
1981	85
1982	68
1983	54
1984	48
1985	52
1986	50
1987	30
1988	38
1989	37
1990	26
1991	9
1992	1
1993	0
1994	0

Notes: Before the drafting system was abolished, young men either tried to avoid the drafting or it got less strict and therefore we can observe a fading out effect.

Table 7: Diff-in-Diff regression table for age cohorts 18-19 (Data Set 1)..The birthyear cut-off is at 1990.

	(1) Smoking	(2) Smoking	(3) BMI	(4) BMI
Time	-0.122***	-0.115***	0.645***	0.625***
	[0.0227]	[0.0225]	[0.197]	[0.197]
Male	0.0247	0.0279	1.104***	1.120***
	[0.0182]	[0.0181]	[0.139]	[0.139]
Diff-in-Diff	0.0472	0.0427	-0.281	-0.286
	[0.0331]	[0.0326]	[0.280]	[0.280]
Migration		-0.0295***		0.0864
		[0.0104]		[0.0878]
Region		0.0221		-0.103
		[0.0200]		[0.152]
HHincome		-0.0867***		-0.292***
		[0.0113]		[0.0878]
Constant	0.323***	1.291***	21.61***	24.85***
	[0.0128]	[0.129]	[0.0927]	[0.996]
Observations	3637	3637	3637	3637
R^2	0.011	0.031	0.024	0.027
BIC	4711.0	4659.1	19751.1	19762.9

Standard errors in brackets

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Time is the dummy for "before" and "after" the cut-off. (Here before 1990 (0) and after 1990 (1)); Male is the treatment dummy (male and female, male is 1); Diff-in-Diff is the interaction dummy between both Male and Time; Migration is a 3rd part categorial variable, native, 1st migration and 2nd (born in Germany, parents from abroad); Region is either Eastern (1) or Western Germany (0); HHincome is the logarithm of the household income of each individual referring to their parents (using the original household number)
Smoking and BMI are the outcome variables for each model; everything estimated with OLS.

Table 8: Diff-in-Diff regression table for age cohorts 18-19 (Data set 1). The birthyear cut-off is at 1984.

	(1) Smoking	(2) Smoking	(3) BMI	(4) BMI
Time	-0.105***	-0.0995***	0.761***	0.754***
	[0.0284]	[0.0284]	[0.187]	[0.187]
Male	0.0977***	0.101***	1.016***	1.022***
	[0.0370]	[0.0368]	[0.231]	[0.230]
Diff-in-Diff	-0.0764*	-0.0781*	0.0235	0.0337
	[0.0406]	[0.0403]	[0.269]	[0.269]
Migration		-0.0286***		0.0789
		[0.0104]		[0.0877]
Region		0.0332*		-0.159
		[0.0197]		[0.150]
HHincome		-0.0837***		-0.308***
		[0.0114]		[0.0879]
Constant	0.377***	1.306***	21.16***	24.59***
	[0.0259]	[0.131]	[0.162]	[0.999]
Observations	3637	3637	3637	3637
R^2	0.018	0.038	0.027	0.030
BIC	4684.4	4631.9	19739.7	19750.1

Standard errors in brackets

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Time is the dummy for "before" and "after" the cut-off. (Here before 1984 (0) and after 1984 (1)); Male is the treatment dummy (male and female, male is 1); Diff-in-Diff is the interaction dummy between both Male and Time; Migration is a 3rd part categorial variable, native, 1st migration and 2nd (born in Germany, parents from abroad); Region is either Eastern (1) or Western Germany (0); HHincome is the logarithm of the household income of each individual referring to their parents (using the original household number)

Smoking and BMI are the outcome variables for each model; everything estimated with OLS.

Table 9: Diff-in-Diff regression table for age cohorts 18-19 (Data Set 1). The birthyear cut-off is at 1984.

Birthcohorts from 1985-1990 are excluded

	(1) Smoking	(2) Smoking	(3) BMI	(4) BMI
Time	-0.175***	-0.160***	1.093***	1.070***
	[0.0320]	[0.0322]	[0.237]	[0.238]
Male	0.0977***	0.101***	1.016***	1.023***
	[0.0370]	[0.0369]	[0.231]	[0.231]
Diff-in-Diff	-0.0258	-0.0321	-0.193	-0.190
	[0.0462]	[0.0458]	[0.336]	[0.335]
Migration		-0.0394***		0.0612
		[0.0145]		[0.121]
Region		0.0572*		-0.152
		[0.0327]		[0.235]
HHincome		-0.0773***		-0.292**
		[0.0161]		[0.120]
Constant	0.377***	1.231***	21.16***	24.41***
	[0.0259]	[0.184]	[0.162]	[1.368]
Observations	1650	1650	1650	1650
R^2	0.048	0.071	0.036	0.040
BIC	2113.2	2094.7	8812.3	8828.0

Standard errors in brackets

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Time is the dummy for "before" and "after" the cut-off. (Here before 1984 (0) and after 1990 (1)); Male is the treatment dummy (male and female, male is 1); Diff-in-Diff is the interaction dummy between both Male and Time; Migration is a 3rd part categorial variable, native, 1st migration and 2nd (born in Germany, parents from abroad); Region is either Eastern (1) or Western Germany (0); HHincome is the logarithm of the household income of each individual referring to their parents (using the original household number)
Smoking and BMI are the outcome variables for each model; everything estimated with OLS.

Table 10: 95% confidence interval of all Diff-in-Diff estimators of all regressions of Data Set 1

Regression Table 1	Outcome Variable	Lower Bound	Upper Bound
Diff-in-Diff, no controls	Smoking	-0.0176407	0.1120636
Diff-in-Diff, controls	Smoking	-0.0212035	0.1065918
Diff-in-Diff, no controls	BMI	-0.8307607	0.2684876
Diff-in-Diff, controls	BMI	-0.8346887	0.2620345
Regression Table 2	Outcome Variable	Lower Bound	Upper Bound
Diff-in-Diff, no controls	Smoking*	-0.1559024	0.0031937
Diff-in-Diff, controls	Smoking*	-0.1572409	0.0009674
Diff-in-Diff, no controls	BMI	-0.5042166	0.5511584
Diff-in-Diff, no controls	BMI	-0.4933556	0.5606939
Regression Table 3	Outcome Variable	Lower Bound	Upper Bound
Diff-in-Diff, no controls	Smoking	-0.1164396	0.064821
Diff-in-Diff, controls	Smoking	-0.1219274	0.0577396
Diff-in-Diff, no controls	BMI	-0.8512505	0.4653112
Diff-in-Diff, controls	BMI	-0.8460582	0.4669355

Notes: No controls stands only for the Diff-in-Diff dummy, while controls stands for the Diff-in-Diff estimator, controlled by region, hhincome and migration (Bigger model). Stars note a significant coefficient. Prob>F is 0.0000 for all models.

Table 11: Diff-in-Diff regression table for age cohorts 20-21 (Data Set 2).. The birthyear cut-off is at 1990.

	(1) Smoking	(2) Smoking	(3) BMI	(4) BMI
Time	-0.132***	-0.138***	0.512***	0.419**
	[0.0228]	[0.0228]	[0.196]	[0.192]
Male	0.0345*	0.0399**	1.303***	1.333***
	[0.0179]	[0.0177]	[0.132]	[0.132]
Diff-in-Diff	0.0901***	0.0900***	0.230	0.237
	[0.0336]	[0.0332]	[0.286]	[0.284]
Migration		0.00258		0.0117
		[0.0106]		[0.0824]
Region		0.0527***		-0.0639
		[0.0195]		[0.154]
HHincome		-0.0788***		-0.469***
		[0.0102]		[0.0957]
Constant	0.387***	1.244***	22.04***	27.25***
	[0.0125]	[0.116]	[0.0928]	[1.084]
Observations	4064	4064	4064	4064
R^2	0.011	0.031	0.037	0.045
BIC	5650.8	5595.5	22325.9	22315.7

Standard errors in brackets

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Time is the dummy for "before" and "after" the cut-off. (Here before 1990 (0) and after 1990 (1)); Male is the treatment dummy (male and female, male is 1); Diff-in-Diff is the interaction dummy between both Male and Time; Migration is a 3rd part categorial variable, native, 1st migration and 2nd (born in Germany, parents from abroad); Region is either Eastern (1) or Western Germany (0); HHincome is the logarithm of the household income of each individual referring to their parents (using the original household number)

Smoking and BMI are the outcome variables for each model; everything estimated with OLS.

Table 12: Diff-in-Diff regression table for age cohorts 20-21 (Data Set 2). The birthyear cut-off is at 1984.

	(1) Smoking	(2) Smoking	(3) BMI	(4) BMI
Time	-0.155***	-0.158***	0.827***	0.780***
	[0.0230]	[0.0228]	[0.166]	[0.165]
Male	0.0249	0.0301	1.237***	1.263***
	[0.0276]	[0.0273]	[0.184]	[0.184]
Diff-in-Diff	0.0508	0.0507	0.169	0.178
	[0.0330]	[0.0326]	[0.237]	[0.237]
Migration		0.00360		0.0131
		[0.0107]		[0.0820]
Region		0.0613***		-0.103
		[0.0191]		[0.152]
HHincome		-0.0764***		-0.478***
		[0.0102]		[0.0954]
Constant	0.456***	1.286***	21.61***	26.94***
	[0.0193]	[0.117]	[0.130]	[1.083]
Observations	4064	4064	4064	4064
R^2	0.020	0.039	0.044	0.053
BIC	5616.5	5559.9	22295.7	22283.7

Standard errors in brackets

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Time is the dummy for "before" and "after" the cut-off. (Here before 1984 (0) and after 1984 (1)); Male is the treatment dummy (male and female, male is 1); Diff-in-Diff is the interaction dummy between both Male and Time; Migration is a 3rd part categorial variable, native, 1st migration and 2nd (born in Germany, parents from abroad); Region is either Eastern (1) or Western Germany (0); HHincome is the logarithm of the household income of each individual referring to their parents (using the original household number)

Smoking and BMI are the outcome variables for each model; everything estimated with OLS.

Table 13: Diff-in-Diff regression table for age cohorts 20-21 (Data Set 2). The birthyear cut-off is at 1984.
Birthcohorts from 1985-1990 are excluded.

	(1) Smoking	(2) Smoking	(3) BMI	(4) BMI
Time	-0.202***	-0.200***	0.936***	0.815***
	[0.0271]	[0.0272]	[0.217]	[0.213]
Male	0.0249	0.0323	1.237***	1.265***
	[0.0276]	[0.0273]	[0.184]	[0.185]
Diff-in-Diff	0.0997**	0.0976**	0.296	0.310
	[0.0397]	[0.0391]	[0.313]	[0.311]
Migration		-0.0106		0.0234
		[0.0136]		[0.104]
Region		0.106***		-0.140
		[0.0269]		[0.210]
HHincome		-0.0803***		-0.529***
		[0.0129]		[0.124]
Constant	0.456***	1.321***	21.61***	27.51***
	[0.0193]	[0.147]	[0.130]	[1.410]
Observations	2356	2356	2356	2356
R^2	0.031	0.060	0.053	0.065
BIC	3282.2	3235.4	12868.4	12863.3

Standard errors in brackets
* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Time is the dummy for "before" and "after" the cut-off. (Here before 1984 (0) and after 1990 (1)); Male is the treatment dummy (male and female, male is 1); Diff-in-Diff is the interaction dummy between both Male and Time; Migration is a 3rd part categorial variable, native, 1st migration and 2nd (born in Germany, parents from abroad); Region is either Eastern (1) or Western Germany (0); HHincome is the logarithm of the household income of each individual referring to their parents (using the original household number)
Smoking and BMI are the outcome variables for each model; everything estimated with OLS.

Table 14: 95% confidence interval of all Diff-in-Diff estimators of all regressions of Data Set 2

Regression Table 1	Outcome Variable	Lower Bound	Upper Bound
Diff-in-Diff, no controls	Smoking***	0.0241121	0.1560373
Diff-in-Diff, controls	Smoking***	0.0249021	0.1550156
Diff-in-Diff, no controls	BMI	-0.3309268	0.7911838
Diff-in-Diff, controls	BMI	-0.31878	0.7932129
Regression Table 2	Outcome Variable	Lower Bound	Upper Bound
Diff-in-Diff, no controls	Smoking	-0.0138348	0.115373
Diff-in-Diff, controls	Smoking	-0.0132324	0.1147084
Diff-in-Diff, no controls	BMI	-0.2958165	0.633065
Diff-in-Diff, no controls	BMI	-0.2855573	0.6421361
Regression Table 3	Outcome Variable	Lower Bound	Upper Bound
Diff-in-Diff, no controls	Smoking**	-0.0379154	0.2373265
Diff-in-Diff, controls	Smoking**	0.0209298	0.1743091
Diff-in-Diff, no controls	BMI	-0.3182223	0.9107026
Diff-in-Diff, controls	BMI	-0.2995799	0.9204242

Notes: No controls stands only for the Diff-in-Diff dummy, while controls stands for the Diff-in-Diff estimator, controlled by region, hhincome and migration (Bigger model). Stars note a significant coefficient. Prob>F is 0.0000 for all models.

Table 15: Diff-in-Diff regression table for age cohorts 22-23 (Data Set 3). The birthyear cut-off is at 1984.

	(1) Smoking	(2) Smoking	(3) BMI	(4) BMI
Time	-0.0721***	-0.0722***	0.830***	0.827***
	[0.0247]	[0.0245]	[0.202]	[0.201]
Male	0.0347	0.0410	1.464***	1.495***
	[0.0251]	[0.0250]	[0.181]	[0.182]
Diff-in-Diff	0.0196	0.0173	-0.0916	-0.0942
	[0.0354]	[0.0352]	[0.280]	[0.279]
Region		0.0549***		0.0112
		[0.0208]		[0.166]
HHincome		-0.0648***		-0.369***
		[0.0125]		[0.0991]
Constant	0.428***	1.129***	22.09***	26.16***
	[0.0174]	[0.141]	[0.129]	[1.114]
Observations	3080	3080	3080	3080
R^2	0.006	0.019	0.043	0.048
BIC	4392.6	4366.6	17108.3	17109.1

Standard errors in brackets

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Time is the dummy for "before" and "after" the cut-off. (Here before 1984 (0) and after 1984 (1)); Male is the treatment dummy (male and female, male is 1); Diff-in-Diff is the interaction dummy between both Male and Time; Migration is a 3rd part categorial variable, native, 1st migration and 2nd (born in Germany, parents from abroad); Region is either Eastern (1) or Western Germany (0); HHincome is the logarithm of the household income of each individual referring to their parents (using the original household number)
Smoking and BMI are the outcome variables for each model; everything estimated with OLS.

Table 16: 95% confidence interval of all Diff-in-Diff estimators of all regressions of Data Set 2

Last Regression Table	Outcome Variable	Lower Bound	Upper Bound
Diff-in-Diff, no controls	Smoking	-0.0498089	0.0890965
Diff-in-Diff, controls	Smoking	-0.0516978	0.0863149
Diff-in-Diff, no controls	BMI	-0.6401597	0.4569702
Diff-in-Diff, controls	BMI	-0.6418644	0.4534382

Notes: No controls stands only for the Diff-in-Diff dummy, while controls stands for the Diff-in-Diff estimator, controlled by region and hhincome (Bigger model). Stars note a significant coefficient. Prob>F is 0.0000 for all models

C. Figures

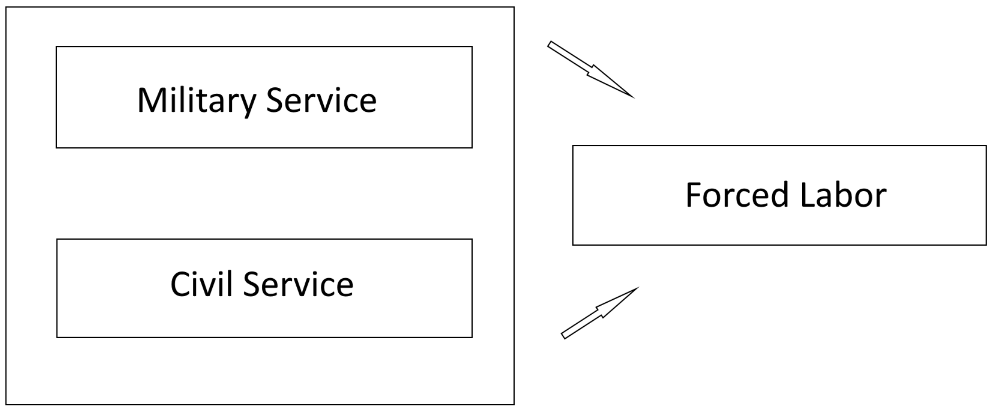

Figure 1: Selection Process, delivering two different selection groups

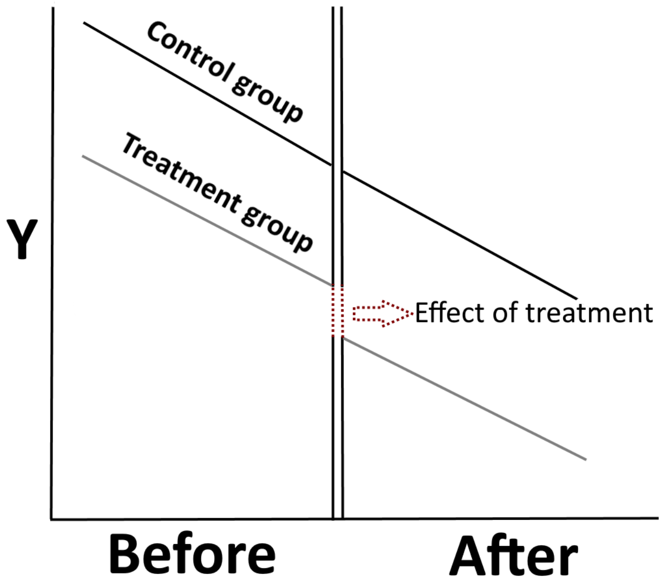

Figure 2: Standard linear parallel trend for treatment and control group

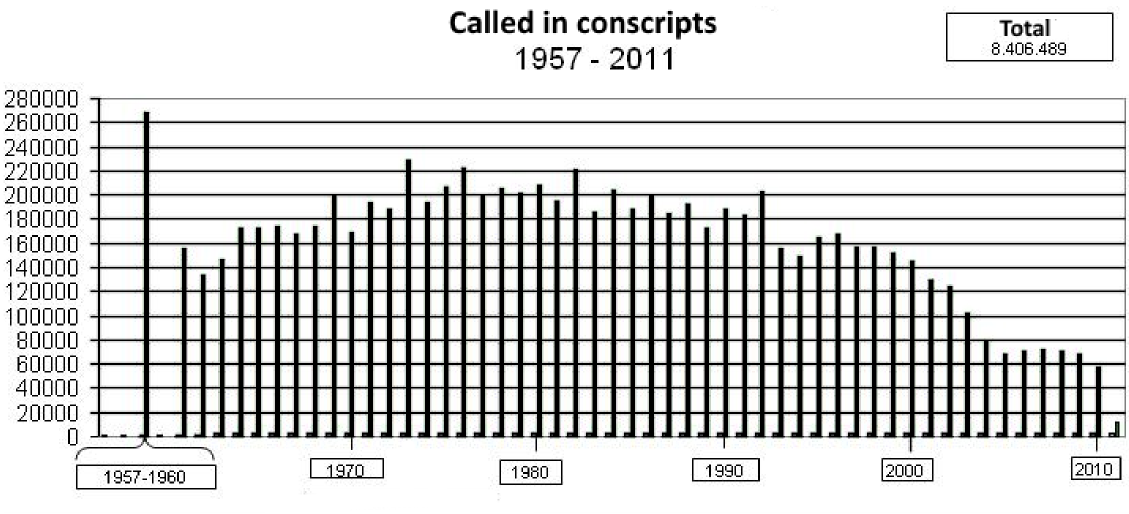

Figure 3: Called in conscripts per year, source; `https://de.wikipedia.org/wiki/Musterung`

Figure 4: Correlation between BMI and sex, men are 1 and female are 0

Figure 5: Correlation between smoking and sex, men are 1 and female are 0

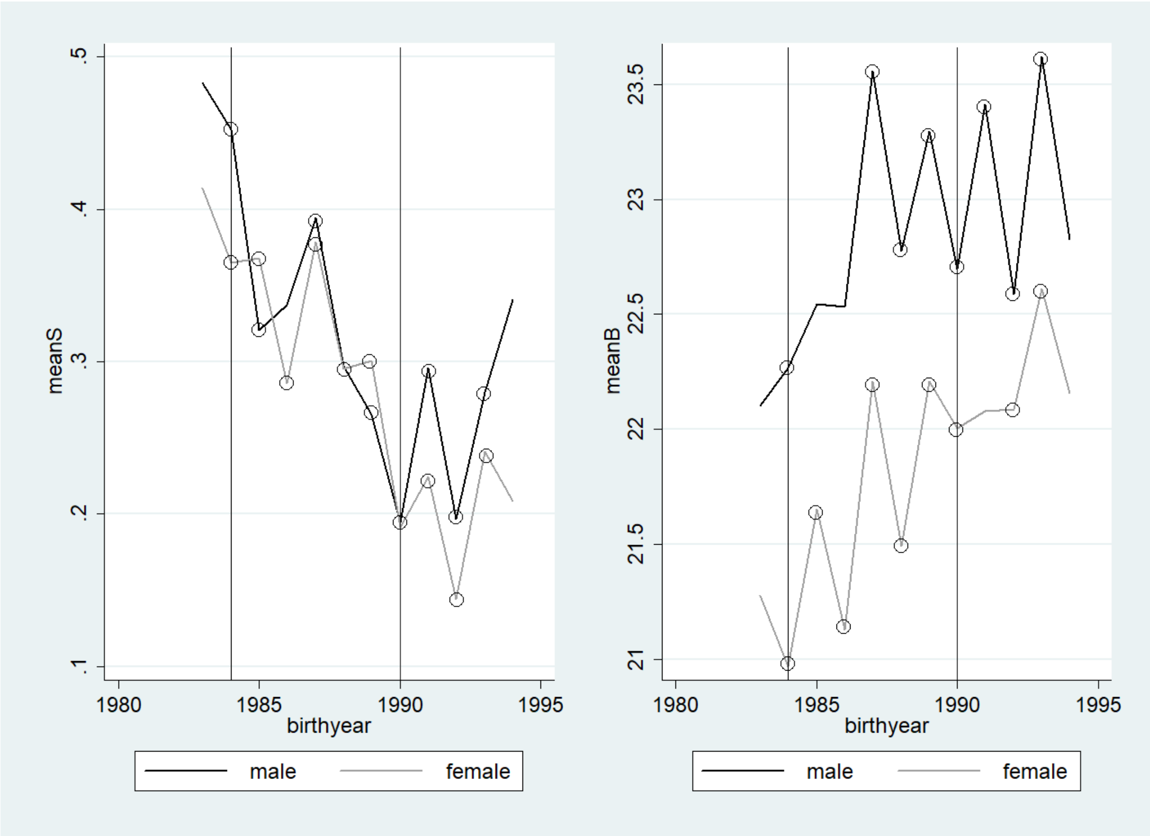

Figure 6: Parallel trend for both birth-year cut-offs (1984 and 1990) and outcomes. Data set 1.
Remember: Military reform 2011, share of smoker = meanS , average BMI = meanB

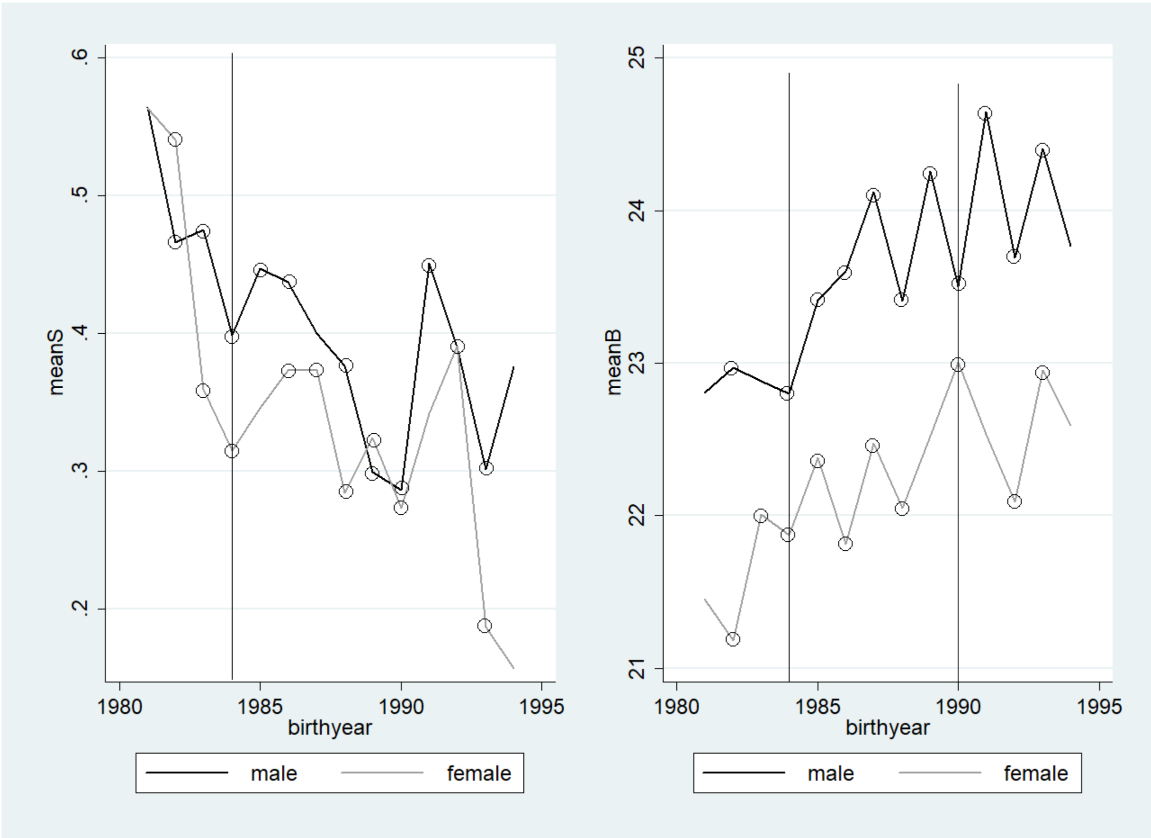

Figure 7: Parallel trend for both birth-year cut-offs (1984 and 1990) and outcomes. Data set 2.
Remember: Military reform 2011, share of smoker = meanS, average BMI = meanB

Figure 8: Parallel trend for both birth-year cut-offs (1984 and 1990) and outcomes. Data set 3.
Remember: Military reform 2011, share of smoker = meanS, average BMI = meanB