Aalto University
School of Business

The Bottle and the Damage Done: Alcohol Supply and Related Crime

Estimating the effects of Keskiolutlaki on subsequent crime rates

Economics

Master's thesis

Juha Luukkonen

2015

Department of Economics
Aalto University
School of Business

Aalto University School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Juha Luukkonen

Title of thesis The Bottle and the Damage Done: Alcohol Supply and Related Crime – Estimating the effects of Keskiolutlaki on subsequent crime rates

Degree M.Sc. in economics and business administration

Degree programme Economics

Thesis advisor(s) Matti Sarvimäki

Year of approval 2015 **Number of pages** 105 **Language** English

Abstract

The aim of this study is to find how municipal crime levels changed in Finland in relation to the municipal alcohol availability during the years 1960-1975. The main focus is in the legislative change known as Keskiolutlaki (enforced in 1969), which liberated the retail sales of medium-strength beer across the land and ended the prohibition-like conditions that existed in the rural municipalities. The study of this subtle relationship between alcohol and crime is done with regression analysis. Furthermore, tools such as fixed effects, differences-in-differences and Poisson regression are used in distinguishing this relationship. Before the regressions, I review the legislative change in a historical perspective, its implications to alcohol consumption and previous studies on how alcohol availability affects consequent crime rates.

The main dependent variables of this study are overall crime, traffic crime, other crime, drunk arrests (which means drunk and disorderly conduct), batteries and 'fatal offences and attempts at these'. The main independent variables reflect the municipal-level alcohol supply in the consequent years. The crime levels are population adjusted (or exposed to the population in Poisson regressions). In order to further distinguish between how alcohol supply affected different municipalities' crime levels after the legislative change, a quasi-experimental, differences-in-differences -based research setup is used. The treatment and control groups are formed according to the placement of Alko stores, which the most distinct form of alcohol supply in Finland

The spatial availability of alcohol supply can be seen to significantly influence arrests regarding more precise study: More time-invariant variables especially regarding socioeconomic and-demographic municipal differences and more accurate estimates on how the alcohol supply and the municipal populations were geographically distributed could lead to more accurate results.

Keywords alcohol, crime, Keskiolutlaki, Finland, regression, Poisson, fixed effects, diff-in-diff

Table of Contents

1 Introduction 1

1.1 Background 1

1.2 The importance of the subject 3

1.3 The approach and unfortunate limitations 4

2 Research question 7

3 Historical background: Kekkonen era Finnish alcohol politics 9

3.1 Keskiolutlaki (462/1968)..9

3.2 The aftermath 11

3.3 Alcohol supply location decisions and endogeneity 16

4 Previous studies on alcohol availability and crime 21

5 Empirical approach 23

5.1 Acquisition of data 23

5.2 Variables defined & descriptive statistics 25

5.2.1 Municipal statistics 25

5.2.2 The distances of liquor stores 30

5.2.3 Preliminary observations of graphical interpretation 32

5.3 Regressions used and acknowledged limitations 37

5.3.1 Fixed effects regression 38

5.3.2 Poisson regression 39

5.3.3 Differences-in-differences regression 42

5.4 Defining the experimental design 45

5.4.1 The trouble of defining a solid experimental design 47

5.4.2 Evidence of preliminary significance and other observations 52

6 Empirics 1960-1975 56

6.1 Findings 56

6.2 Proposals for further studies 65

7 Concluding remarks 67

8 References 68

9 Appendices 72

List of tables

Table 1. The main legislative changes introduced by the Keskiolutlaki 10
Table 2. The license types and their respective amounts in 1963 and 1970 18
Table 3. Descriptive statistics of different crime types (years 1960 and 1975) 26
Table 4. Descriptive statistics of distances to Alko (years 1960 and 1975) 27
Table 5. Descriptive statistics of license dummies (years 1960 and 1975) 28
Table 6. Descriptive statistics of population densities (years 1960 and 1975) 29
Table 7. OLS results for overall crime 56
Table 8. OLS results for traffic crime 58
Table 9. OLS results for drunk arrests 59
Table 10. Group setup OLS results for drunk arrests 60
Table 11. Poisson results for batteries (IRR) 62
Table 12. Group setup Poisson results for batteries 63
Table 13. Poisson results for fatal offences and attempts at these 64

List of figures

Figure 1. Alcohol consumption per capita aged 15+ (pure ethanol) 11
Figure 2. Alcohol consumption by alcohol types per capita aged 15+ (pure ethanol) 12
Figure 3. Alcohol consumption by purchase places per capita aged 15+ (pure ethanol) 13
Figure 4. Alko stores and licensed premises in relation to average municipal crime 14
Figure 5. Medium-strength beer serving outlets in relation to average municipal crime 15
Figure 6. Average distance to the nearest Alko store in relation to crime per capita 32
Figure 7. Scatter plot of municipal distance to the nearest Alko and crime per capita 33
Figure 8. Scatter plot of municipal population density and crime per capita 34
Figure 9. Scatter plot of municipal distance to the nearest Alko and batteries 35
Figure 10. Scatter plot of municipal population density and batteries 36
Figure 11. Scatter plot of municipal population density and distance to the nearest Alko 37
Figure 12. Causal effects in the differences-in-differences model 43
Figure 13. Cumulative distr. of the average municipal population densities (1960-1975) 48
Figure 14. Interpreting the group trends for the Diff-in-diff regressions 52
Figure 15. Favorable vs. unfavorable trend changes regarding the Diff-in-diff 54

1 Introduction

1.1 Background

Before 1969 buying liquor was hard on the Finnish countryside as the sales were limited to cities and municipalities with almost city like rights (kauppala in Finnish). Things changed with the introduction of *Keskiolutlaki (462/1968,* translated to Act on Medium-Strength Beer*)*. It partially substituted the *Laki väkijuomista (45/1932, translated to Act on Hard Liquor)*, which had been active since *Kieltolaki (158/1922, the Finnish Prohibition)* was brought to an end in 1932. As one can predict by its name, the new law's center point was expanding and legalizing the sales of medium-strength beer in retail stores and beer-serving licensed premises nationwide, but in addition the age limits for purchasing alcohol were toned down and licensed premises[1] and state monopoly liquor stores could be established in the countryside.

Out of the three main reforms stated earlier, two are quite similar when considering the geographical impact they had in consumption: The lowered age limits and the medium-strength beer's march across the land with retail stores and beer joints as its acting envoys. Location-wise, the medium-strength beer sales network can be seen as a reasonably efficient market. The age limits did not spatially discriminate either, because they are nationally universal.

The third reform, concerning the former rural-urban divide on the alcohol supply, on the other hand is geographically quite asymmetric (yet there is endogeneity related): The Finnish alcohol monopoly's, Alko Inc.'s spreading in the countryside can be seen as an asymmetric alcohol supply shock in terms with the Alko stores' more limited geographic coverage. Alko Inc. controls the retail monopoly for alcoholic beverages exceeding 4.7 % alcohol by volume (ABV) to this day. The monopoly did not and still does not distribute in the most efficient possible manner, with its guidelines somewhere in between being a responsible distributor and a profit maker. Considering the fact that more than half of the Finnish population was

[1] Note that there is a distinction between licensed premises that serve more potent alcohol (restaurants and bars) and premises that were established onwards from 1969 and could only serve medium-strength beer.

living in the rural municipalities in 1968, the new legislative changes made alcohol more available to roughly 2.4 million persons[2]. The rural limitations for the licensed premises were also shifted with the new legislation, which makes them an interesting part of this reform as well.

Altogether these changes had an imminent effect on the annual Finnish alcohol consumption, which rose 46 % converted into absolute alcohol in just one year (Oy Alko ab, 1970). While it is clear that the changes brought by the new law altered the Finnish consumption statistics, it is also interesting to see what kind of social implications the then-new legislation had.

The aim of this study is to determine whether the legislative changes in different elements of Finnish alcohol supply affected municipal-level crime in Finland. In order to study this relation there are two main observed variables: The municipal distance to the closest municipality with an Alko store and a dummy variable expressing licensed premises in the municipal area. Simultaneously the other elements of the legislative change are attempted to control with a dummy variable.

The justification for the main explanatory variables goes as follows: The Alko stores and the licensed premises can be seen as asymmetrical supply shocks, while the other changes introduced by Keskiolutlaki can be seen as a rather symmetrical supply shock. With this in mind, I am going to estimate on how the changes in distances to the nearest Alko stores and in the existence of municipal licensed premises correlated in the municipal crime statistics.

This framework, more historical background and the statistical approach will be discussed more thoroughly in their respective chapters.

[2] Population statistics 1968: Rural population 2404400 (51%) vs. city population 2294000 (49%) (Central Statistical Office, 1970).

1.2 The importance of the subject

As the legislative change is the biggest alcohol reform since the Finnish prohibition was ended in 1932, it serves as a huge social experiment of how toned down alcohol control affects alcohol use and alcohol-induced problems in the society.

Of course the people living in the countryside were able to get a hold of alcohol from the cities before the legislative change, or even distil it in the wilderness behind some spruce, but it got a whole less laborious in 1969. One interesting aspect is that, as peculiar as it may sound, in some regions the rather controlled alcohol outlets might even reduce crime opposed to uncontrolled use of alcohol.

At first one might think that the relative effect of the new Alko stores or the role of the licensed premises on the alcohol induced crimes might be rather marginal or even nonexistent, diminished by the overall legislative changes. However, "*Alcohol Availability, Prenatal Conditions, and Long-term Economic Outcomes*", a study by J Peter Nilsson (2014), acts as inspiration in terms of the scale of effects. Nilsson has studied the effects of a rather brief prenatal exposure to loosened alcohol supply in terms of strong beer in retail stores. The experiment took part in two regions of Sweden for a time window of 8.5 months and according to it, the prenatal children that were influenced by this sales experiment have lower earnings, cognitive and non-cognitive abilities and generally have it worse than their surrounding (non-prenatal while the experiment took place) cohorts. Thus even relatively small sounding changes in availability of alcohol can make a significant difference in the outcomes, and considering that Keskiolutlaki leveled out huge regional differences in alcohol supply makes this a fruitful setting research-wise.

1.3 The approach and unfortunate limitations

In order to estimate whether the proximity of Alko stores and the licensed premises had an effect on (the alcohol related) crime, regression analyses are performed. The crime statistics chosen for this study are overall crime, ‘murders, manslaughters & attempts at both of these’, batteries, traffic crimes, other crimes and drunk arrests (which expresses drunk and disorderly conduct), which, on a side note, is not a crime itself, just a reason for an arrest. For crimes that involve violence (‘murders, manslaughters and attempts at both of these’ and batteries), Poisson regression is chosen as the primary tool, as violence in a municipal scale is after all a relatively rare event and this approach can estimate the differences influenced by the legislative change on these relatively low statistics.

In addition to the statistics that involve municipal criminal activity or alcohol availability, factors such as municipal population density, different dummy variables regarding legislative changes and a time trend variable are used in the regressions. The fixed effects regressions attempt to control the unobserved municipal differences that are time-consistent.

The statistics are comprised of criminality known to the police statistics from 1960-1975 (Central Statistical Office, 1960-1971) (Central Statistical Office of Finland, 1972-1975), the year books of Alko Inc. for the respective years (Oy Alkoholiliike ab, 1960-1969) (Oy Alko ab, 1970-1975) and the municipal areas and distances are calculated from municipal division maps from the years 1960 and 2015 (Maanmittaushallitus, 1960) (National Land Survey, Esri Finland , 2015). The acquisition of these statistics is thoroughly explained in chapter 5.

Naturally there are also limitations for this study. Below I have listed some issues of importance:

It must be noted that while the proximities of alcohol serving outlets do not tell anything about individual consumption of alcohol, rather the variables associated with the supply can be used as a proxy for the opportunity to consume alcohol (or as changes in proximities as changes in opportunities to consume alcohol). Statistics-wise these consumption opportunities are controlled for the alcohol over 4.7 % ABV, which was limited to the licensed premises and Alko stores before and after the legislative change. However for the medium-strength beer serving outlet statistics are not included due to the fact that I could not find them on a consistent statistical form. Instead of being announced on a municipal level, they were largely announced on a sales territorial level, which did not function that well with the other statistics. However there is a workaround for this in the form of a dummy variable as the medium-strength beer sales network's launch can be seen as rather all-compassing nation-wide.

As for the limitations regarding the crime statistics, they are largely announced on an aggregated level due to the fact that they had to be manually compiled and because they were initially announced on a rather general level. This leads to a fact that certain notable crime types, such as property crimes are pooled in the category of '*other crimes'* and as the drunk driving statistics known to the police would be an interesting addition to the dependent variables, they are included in the traffic crimes. While this absence of more accurate statistics is partly a choice between multiplying the workload and diminished statistical power, I firmly believe that this kind of setup is enough to identify plausible relationships with crime and alcohol supply in Finland. That is unless the time span of six years after the legislative change is too short, which leads to another issue: If there is a mechanism between crime and alcohol supply, how fast does it show in the crime rates? Is the effect of alcohol supply lagged?

Also *endogeneity bias*[3] and *omitted variable bias*[4] have to be considered. Which came first, alcohol or crime, and are these factors intertwined in some unobserved factors. Is there some selection, in which individuals that committed crimes had gathered in areas with more concentrated alcohol supply or did the alcohol supply concentrate in such areas for some policy reasons (or did the alcohol supply concentrate in more stable areas in which less crime occurs)? Are there some unobserved external factors that might explain crime rates and alcohol availability or are there some unobserved factors that affect how crime was reported or perceived? For example the major urbanization that took place especially during this period might be correlated with both of these. More accurate municipal statistics regarding employment, age & gender distributions and the amount of minorities such as Swedo-Finns in

---

[3] Endogeneity bias refers to an explanatory variable being correlated with the error term thus causing bias in the regression results. This might be due an omitted variable, reverse causality (e.g. does a private school improve its student's grades or are the students that good to begin with) or due to a measurement error.

[4] Omitted variable bias refers to the absence of a relevant explanatory variable, which causes bias in the regression results. In order to cause bias, the omitted variable has to be correlated with the dependent variable at least one explanatory variable. Thus omitted variable bias is a plausible source of endogeneity bias.

the municipal area[5] could be useful in identifying differences in both alcohol availability and crime rates.

Indeed, one factor to be considered is the efficiency/perceptiveness of the police. There definitely were and still are regional differences on the perceptive capabilities of the police force. Personally I think this might correlate quite well with the municipal area and the population density (compare drunk driving in Helsinki vs Inari), although certain crime types might be less affected by these kind of differences in perceptiveness, e.g. non-complainant offences such as offences against one’s life.

There are efforts in order to control these issues: The unobserved municipal factors can be dealt in two categories; I) the time-consistent factors and II) the time-inconsistent factors. The time-consistent omitted variables that might have explanatory power in explaining municipal-level crime differences are tackled with fixed effects –approach, which takes care of time-invariant municipal levels of crime in the form of a dummy variable for each municipality. The time-inconsistent factors, such as the size of the municipal car fleet or maybe the municipal police resources on the other hand might prove a bit problematic, as the only way to control these is to insert some additional variables to the models. Last but not least it must be noted that there is at least minor endogeneity related to the location decisions of alcohol supply and this endogeneity will be dealt in its respective chapter. Other issues are further discussed throughout the thesis or ultimately in suggestions for further studies.

[5] A personal observation while compiling the statistics: Swedo-Finnish municipalities generally seemed to have smaller crime rates.

2 Research question

How did the asymmetric permanent alcohol supply shocks, in the forms of Alko stores and licensed premises, affect municipal crime rates before and after Keskiolutlaki?

The main interest of the research question consists of studying how the increased alcohol supply shows in the population adjusted municipal crime statistics especially in the rural Finland (different crime rates as dependent variables). The focus lies in the asymmetrically distributed Alko stores and licensed premises, while a secondary interest is how the medium-strength beer, which swept across the land, affected the crime rates. The tools used in identifying this plausible relationship are fixed effects OLS and Poisson regression analyses. The Poisson regression is used for the fatal violent offences & attempts at these (murders and manslaughters) and for batteries, as they are relatively rare events in the municipal scale, which can cause major biases in linear regressions.

The main explanatory variables used in estimating this relationship with crime and alcohol supply are as follows: Annually estimated municipal distance to the nearest municipality with an Alko store and the presence of licensed premises in the municipal area, which is taken into consideration in the form of a dummy variable (gains the value one if there are any licensed premises in the municipal area). Additionally there is a dummy variable for controlling other changes of the increased alcohol supply introduced by Keskiolutlaki (such as lowered age limits, and the introduction of places serving only medium strength beer). This dummy is used as an interaction variable with the distances to Alko stores and the dummies expressing licensed premises in the municipal area in order to distinguish the effect of the legislative change on them.

The second issue of interest is how the rural municipalities reacted to the alcohol supply, as there might be differences with the cities and the rural municipalities on how the alcohol induced crime comes into existence. The prohibition that existed in the rural municipalities before the legislative change was implemented, as there was concern that the uncivilized rural folk could not handle alcohol and that the countryside would be ridden with alcohol indecent ailments. While this might be valid, this also raises concern regarding the statistical inference: The decisions on alcohol licenses and Alko stores were ultimately made by the alcohol monopoly’s supervisory board and thus they were not randomly assigned to the municipalities. This may cause some problems related to the regressions above and thus the endogeneity caused by the decision making process is further scrutinized with a differences-in-differences based research setup: A quasi-experimental setup is considered in estimating the impact of the locations in which the alcohol supply was directed after the legislative change. The municipalities will be divided into different dummy categories in the following manner: Rural municipalities which received Alko stores after the legislative change will act as the treatment group, while the other rural municipalities without Alko stores and cities which had Alko stores before the legislative change will be used as two separate control groups in estimating the plausible differences how the crime levels in these municipalities reacted to alcohol supply. More of the specifics and justification of this setup will be dealt with later on.

In addition to the variables mentioned, other variables considered are municipal population densities, a time trend and dummy variables expressing changes in speed limit legislation in certain regressions. The time-invariant municipal differences in crime rates are tackled with the fixed effects regressions, which attempt to isolate the time-invariant municipal-level differences related to different crime types. Finally, Åland and archipelago municipalities are omitted from the regressions and the statistics due to being geographically too different.

3 Historical background: Kekkonen era Finnish alcohol politics

Alcohol has been a continuing source of debate and in the Finnish politics since the days of early independence and the most recent parliamentary elections (2015) did not deviate from this continuum. Despite the fact that public scrutiny rises its head in semi-regular intervals (or due to it), there have been only three major legislative overhauls regarding alcohol availability in the Finnish history: The introduction of prohibition in 1919 and its ending in 1932, the liberation of the sales of medium-strength beer and the sales in the rural municipalities in 1969, and the changes in Alko Inc.'s monopoly position in 1995 brought by the European Union. Of these three, I will focus solely on the changes introduced with the Keskiolutlaki. Instead of going that much into the legislative jargon, I am going to go through the incentives behind and major changes introduced by Keskiolutlaki on the supply of alcohol, the supply side's impact on the demand and the further implications of the increased demand. Additionally, I take a brief overview on how the overall criminality changed during 1960-1975 in respect to the alcohol supply network in order to give justification on the subject.

3.1 Keskiolutlaki (462/1968)

The general opinion on the alcohol political environment had been altered in the rather long run and by the end of the 60s, after a long mental tug-o-war between the liberals and the temperance movement, the main parties reached an understanding and the parliamentary vote favoring the new law resulted in a 137-33 victory (Häikiö, 2007).

Historian Martti Häikiö goes through the background factors of the new law in his 2007 book "*Alkon historia*" ("The History of Alko"), The principle section of the new law describes the changed attitudes quite well: As the principle section of the previous law was centered on the overall minimization of alcohol consumption, now it was centered on the minimizing the damages of alcohol, not the consumption itself. The damage minimization process was meant to be done in the form of redirecting consumption to milder alcoholic drinks, mostly to medium-strength beer, which gives the law its name. (Häikiö, 2007)

Häikiö also writes that it was also seen by the then MP, future director general of Alko Inc., Pekka Kuusi that the monopoly had failed in changing the consumption habits of the Finnish people (to favor milder alcohol, as consumption of hard liquor was seen as the biggest problem) and that the state's opportunities for changing the "*sensitive and intimate* *relationship between man and alcohol*" were limited. Individual freedom and responsibility were emphasized.

As for the almost prohibition-like conditions in the rural municipalities, according to Häikiö, the main incentive for the liquor store-free countryside had been the 19th century -oriented view of protection the degenerate peasants from the dangers of alcohol, which was not thought to be a valid concern anymore. As a matter of fact, he states that Alko Inc. was one of the major driving forces for the equal treatment of the rural areas, although all the parties agreed that cautionary measures were in order. Pekka Kuusi, wrote in his preparation memo that the store program must be carefully measured in order to secure a steady and gradual alcohol political development despite the joint effect of medium-strength beer and the new supply channels (Häikiö, 2007).

Below are tabulated the major changes which led to fierce improvements on the supply of alcohol:

	Väkiviinalaki (1932–1968)	**Keskiolutlaki (1969–1995)**
Age limits	Strict age limit of 21 years for all alcohol purchases.	Purchase limit of 18 years for alcohol under 22 % ABV. No restrictions for ages 20 and over.
Supply outside the Alko network	No retail store sales. Restaurant sales limited to cities and city-like municipalities.	Retail store sales for alcohol under 4.7 % ABV. More liberal legislation for Cafes and restaurants (petrol stations added in 1970). Rural limitations removed.
Alko network coverage	Limited to cities and city-like municipalities.	Expanded to the rural municipalities.
Restrictions in demand	Viinakortti (liquor card), a tool for regulating purchases applied to different alcohol types (limitations were frequently changed throughout its existence).	Purchase restrictions abolished for all vines and beers, Viinakortti itself abolished in 1971, although some restrictions apply until 1986.

Table 1. The main legislative changes introduced by the Keskiolutlaki
(Häikiö, 2007)

Estimations on the consumption changes by the new legislation were made by Alko Inc. and its directorate. MP Pekka Kuusi estimated that the consumption would rise somewhere between 10 and 15 % on the first year, and that the use of hard liquor would eventually diminish to favor beer and wines. Director general K.A. Fagerholm concurred that the increased sales would focus on beer and wines. (Häikiö, 2007) According to Häikiö, there were naysayers also: Researcher Klaus Mäkelä estimated to the parliamentary alcohol committee in 1966, that the changes will probably not shift consumption away from the hard liquor, as was pursued, but he also agreed to the general opinion that the patronage regarding alcohol control must be ended.

3.2 The aftermath

The consumption was not expected to rise in the way it did after Keskiolutlaki entered into force. As a peculiar fact, Alko Inc. chose to use logarithmic scale on its consumption related graphs in order to make them look less steep (Häikiö, 2007). Illustrated below is the consumption of 100 % alcohol per capita ages 15 and over:

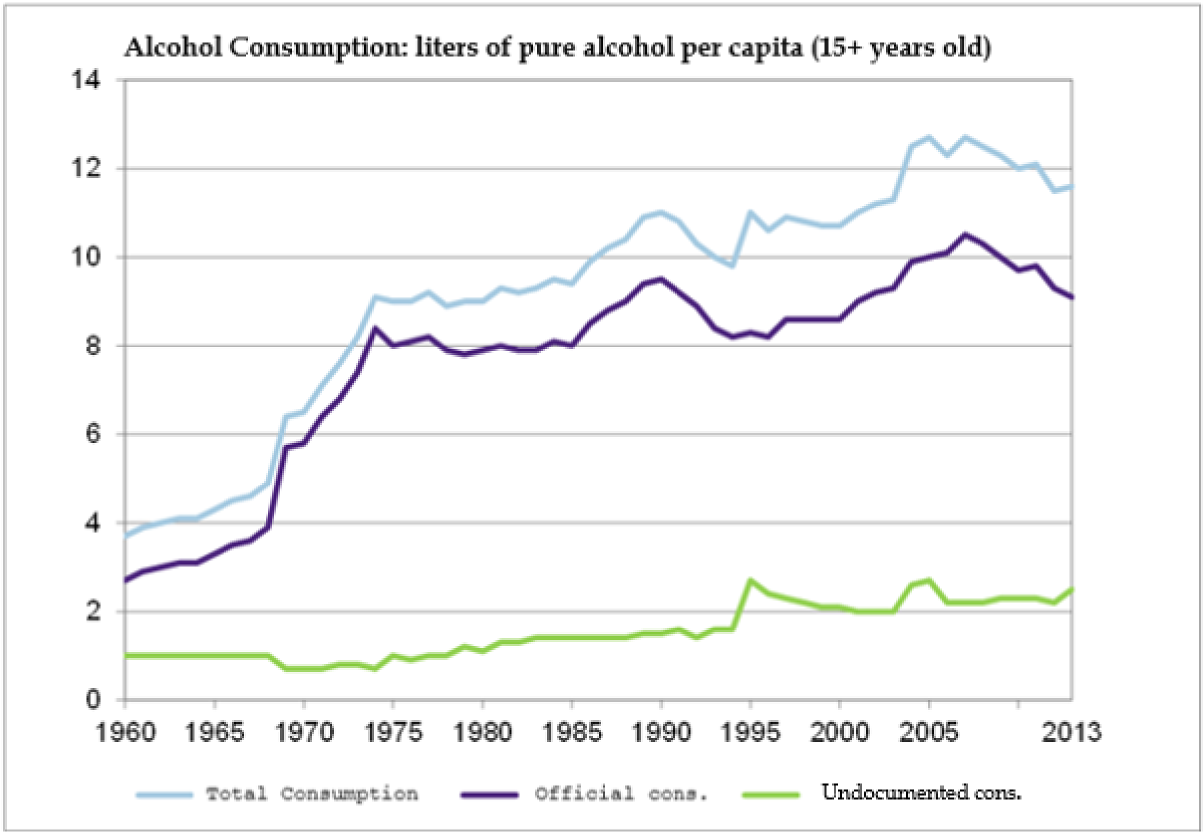

Figure 1. Alcohol consumption per capita aged 15+ (pure ethanol)
(National Institute for Health and Welfare, 2014)

As we see from the figure above, the consumption rose in two batches, first it took an instant charge upwards and secondly over the course of five or so next years it grew a bit more steadily before reaching a steady state of sort for a decade. The immediate impact was that altogether the annual Finnish alcohol consumption rose by 46 % in just one year (Oy Alko ab, 1970). The undocumented and thus estimated consumption consists of alcohol exported to Finland by travelers, smuggling of alcohol, legal and illegal home production and the alcohol consumed by Finns outside the Finnish borders. If we take a closer look at the consumption and unfold it by alcohol types, in Figure 2 below, a pattern can be seen:

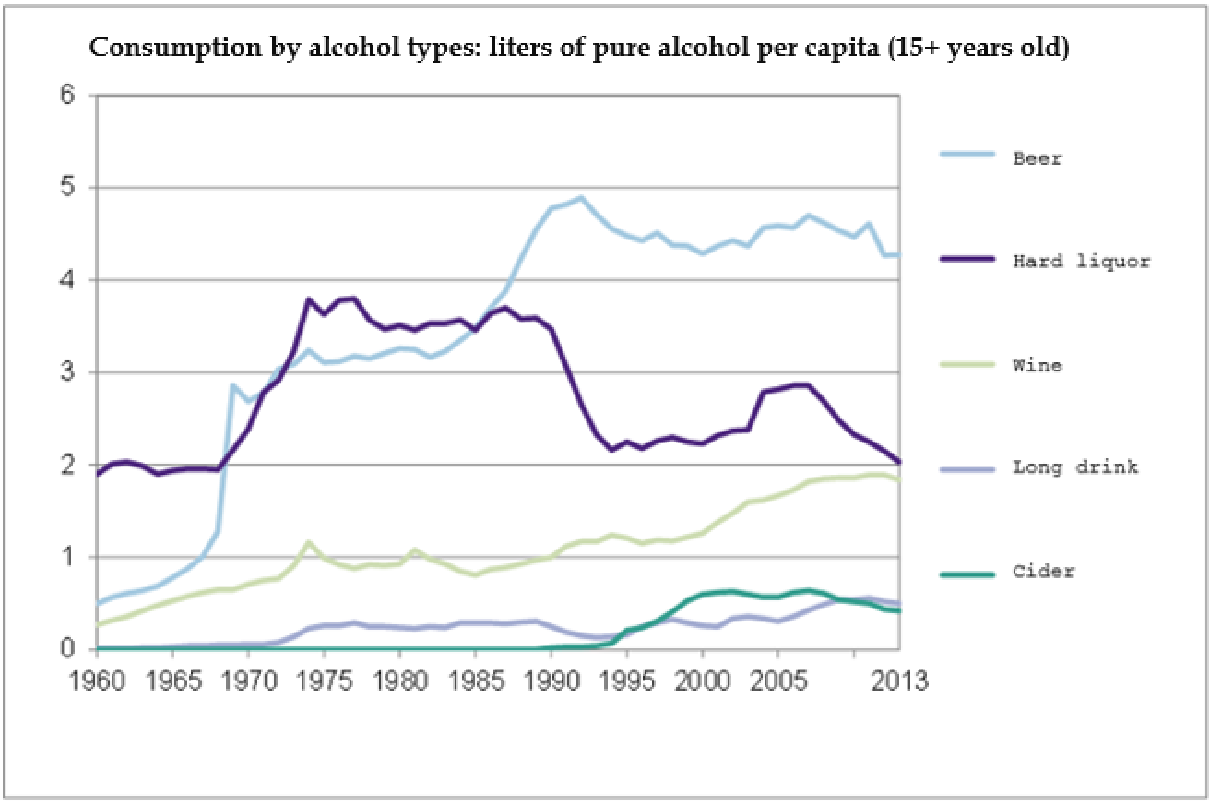

Figure 2. Alcohol consumption by alcohol types per capita aged 15+ (pure ethanol)
(National Institute for Health and Welfare, 2014)

The first impression is that the law was partially a success considering its target to substitute hard liquor with milder alcohol: At first beer did surpass hard liquor as the most popular beverage. Straightforward interpretation however tells us that there was no actual substitution: The fact that the overall consumption of both beer and hard liquor rocketed, dilutes this initial success and the fact that it took hard liquor a few years to surpass beer again dilutes this even more. Although, it must be noted that while the overall consumption rose, we should take into account how the consumption changed between persons that previously had very limited access to alcohol opposed to persons that had more access to alcohol.

Back to the straightforward interpretation; indeed, beer consumption rose immediately on a whole new level and continued to grow more steadily in the coming years, while the hard liquor consumption took roughly five years to settle into its steady state (until the next shock). This is probably due to the fact that while the medium-strength beer was made roughly available to the whole of Finland, the Alko network started carefully and asymmetrically (actually the beer-network surpassed almost whole of Finland, as some rural municipalities initially lacked licensed restaurants/cafes and thus petrol stations were allowed to sell beer with certain precautions starting from 1970 (Häikiö, 2007)).

The next interesting issue is where the consumed alcohol was purchased: From retail stores (including Alko stores and groceries) or from licensed premises (restaurants, bars, cafes and eventually those petrol stations):

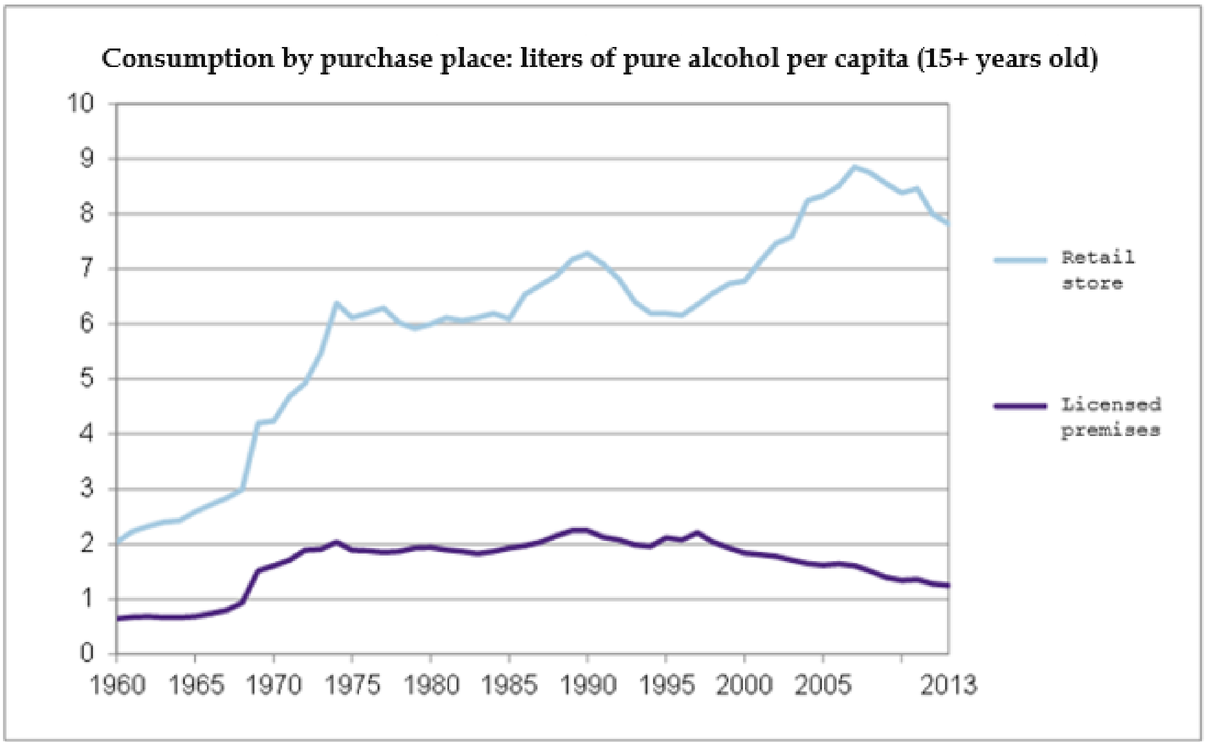

Figure 3. Alcohol consumption by purchase places per capita aged 15+ (pure ethanol)
(National Institute for Health and Welfare, 2014)

It is fairly evident that the consumption boost favored retail sales more, while the licensed premises also received a minor nudge upwards. While a mighty part of the retail sales boost was undoubtedly caused by the licensed retail stores selling only medium-strength beer, Figure 2 indicates that a big part of retail sales boost must come from the Alko sales.

So the supply and the consumption both rose, but in terms of the research question, how does this all go with crime?

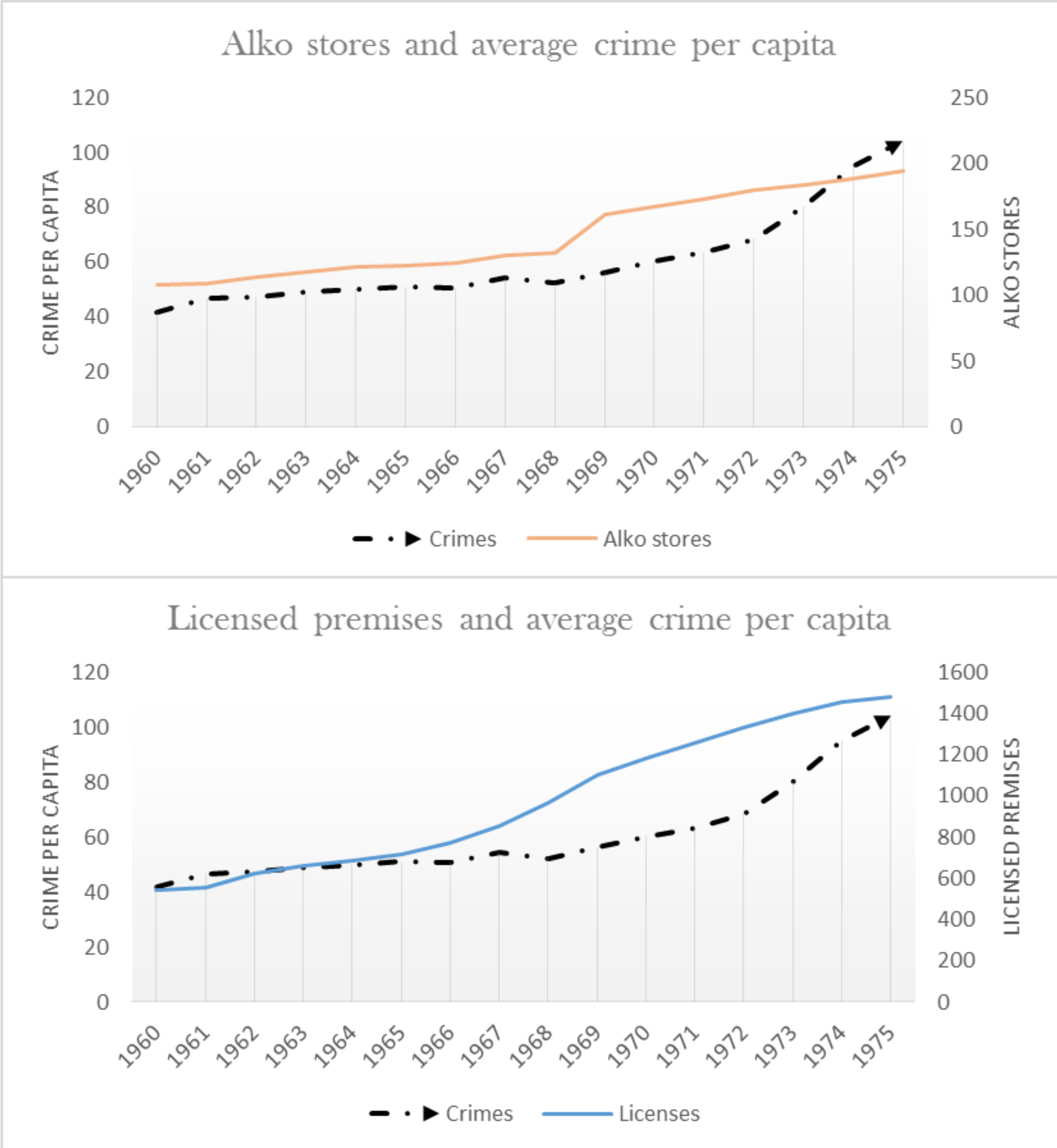

Figure 4. Alko stores and licensed premises in relation to average municipal crime

Illustrated above are crimes known to the police[6] in relation to the amount of licensed premises and in relation to the amount of Alko stores. The licensed premises that could serve

---

[6] The crimes known to the police that are announced as per capita are taken from the municipal records and announced per capita (crimes/1000 persons) and are calculated annually in the following manner:

$$\sum_{i=1}^{n}\frac{1}{n}\left(\frac{\text{overall crime in municipality}_{1}\text{ on year }a}{\text{Population in municipality}_{1}\text{ on year }a}+\cdots+\frac{\text{overall crime in municipality}_{n}\text{ on year }a}{\text{Population in municipality}_{n}\text{ on year }a}\right)$$

only medium-strength beer that were introduced in 1969 are excluded from the above (see figure 5).

What can be interpreted, is that there might be a lagging positive trend in the relation of crime and the alcohol supply network. We have to bear in mind that not all the crimes are potentially influenced by alcohol, and thus this serves as a very preliminary sign that there are plausible results to be found. Another fact to consider regarding the increased crime is that it might also be due to the population changes. After WWII the Finnish society went through baby booming, and this is likely to reflect in the crime rates. To take this into account, some adjustment should be made to the municipal age distributions regarding the citizens that are "capable of crime" or criminally liable but, alas, I could not get my hands on statistics on such municipal age distributions. To analyze this further, below is presented the average municipal crime per capita in relation to the places serving only medium-strength beer (both retail stores and licensed premises).

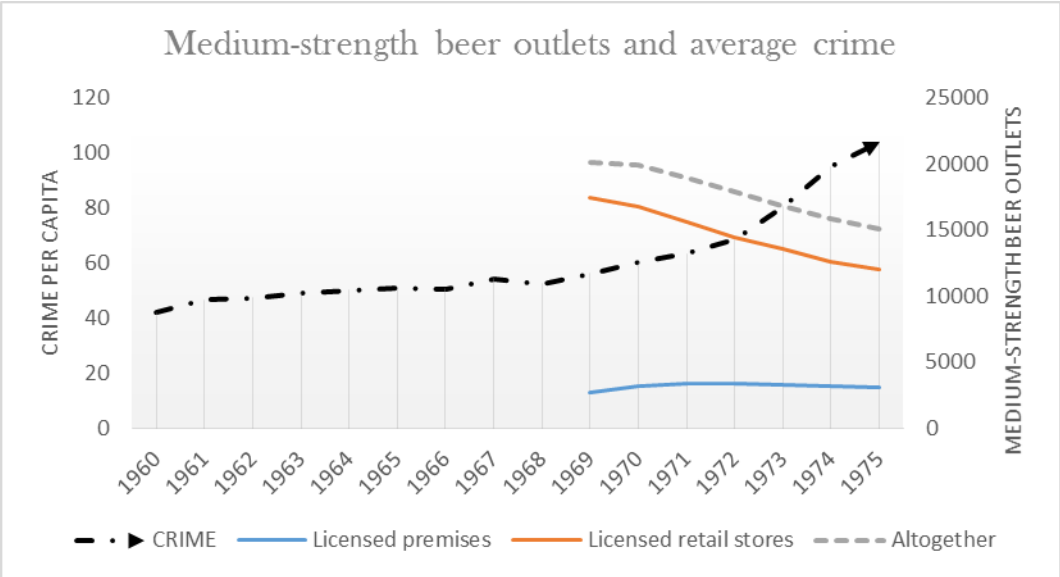

Figure 5. Medium-strength beer serving outlets in relation to average municipal crime

What is noteworthy, is that the amount of retail stores selling medium-strength beer starts off relatively high and comes down in just a few years, as for the licensed premises selling beer, the number is quite steady. Possible reasons for this might be urbanization (less need for stores) or the immediate saturation of the market that occurred in 1969. When comparing to the other graphs, there seems to be no immediate positive relationship with the amount of medium-strength beer outlets and crime, rather a negative one.

What have we learned here? Well, overall the alcohol supply, alcohol consumption and the overall crime levels rose throughout the years, but there is little we can say about the causality yet. Additionally we cannot rule out, that if alcohol supply has an effect on the municipal crime levels, it might have at least a partially lagging effect, e.g. due to the effects of longer misuse, but unfortunately the statistics used in this study are limited only for the initial six years after the legislative change.

3.3 Alcohol supply location decisions and endogeneity

The location decisions related to the alcohol supply were controlled by the alcohol monopoly. The key question is ‘did such guidelines exist that affected how the alcohol supply was located?’ That is, were alcohol serving outlets located in such fashion that they would potentially cause as little harm as possible pre-and post Keskiolutlaki and if such location decisions were made, how successful were they? Is there an omitted factor which both determined how alcohol sales were directed and at the same time dampened criminal activity? The question regarding endogeneity is important, as failure to identify such relationships with the locations and crime, could cause bias in the regressions and false interpretations. Speculation aside, I have gathered some evidence that the location decisions of alcohol serving outlets were not randomized, and thus certain assumptions might not apply for the regressions. Here I will gather information on how licenses were granted and Alko stores were established pre- and post-1969.

Before 1969 the alcohol monopoly had the right to sell alcohol in cities and city-like municipalities (*Väkijuomalaki 9.2.1932/45, § 28*). After Keskiolutlaki came into force the permitting of alcohol sales in any form was a decision of the municipal (city) council. After the permit had been given by the municipal council, the supervisory board of Alkoholiliike (the then Alko Inc.) would grant the licenses or determine whether an Alko store could be established in the municipality (*Alkoholilaki 26.7.1968/459, § 29*).

After 1969, some municipal councils did not allow the sales of alcohol in their region, but they were a minority comprised of small municipalities, and thus their existence is acknowledged, but it is neglected in the regressions[7]. The municipal councils also had the right to abolish the granted rights for sales, and the first municipalities did so in 1974. However these municipalities were also very small in general and these negative decisions were mostly revoked by the coming of 1980s due to economic reasons. (Turunen, 2002) (Peltonen;Kilpiö;& Kuusi , 2006)

Pre-1969, according to the law, the alcohol monopoly would organize the retail sales and the licensed premises. Retail sales it did organize all by itself, but it could however grant temporary licenses to establish premises to outside parties, such as companies or private persons. These outside licenses were in force for three years in cities and for one year in rural areas. In the year 1963, out of the 623 operating licensed premises, 594 fell in to the temporary licenses -category, making the alcohol monopoly more like an organizer of licenses instead of an organizer of the premises as stated in the law. The law also stated that the alcohol sales were to be organized in such a manner that the consumption of liquor should be reduced to an absolute minimum and that drunkenness and its baleful affects are prevented. (Lampela, 1964)

Opposed to the Alko stores, licensed premises could be established in the countryside with certain presuppositions: Hotels and transportation devices, such as cruise ships, that were deemed important enough for travelers, could serve alcohol. (*Väkijuomalaki 9.2.1932/45, § 30*)

The letter of the law, nor any source of mine does not indicate that pre-1969 there were any other limiting decrees for alcohol supply location decisions, other than the fact that alcohol serving outlets could *mostly* be established in cities only. In addition, Häikiö (2007) states that the first Alko stores were established pretty much wherever there was available space, and from between the lines, I interpret that the location elements were not in the limelight in the process of attempting to 'reduce the consumption of liquor to the absolute minimum and preventing drunkenness and its baleful affects', not at least in the cities, where alcohol could be served. Rather other measures were used in saving the citizens: Häikiö indicates that buying alcohol was made, well if not as awkward as possible, then quite awkward. In Alko

[7] Municipalities of Ähtävä and Luoto (Peltonen;Kilpiö;& Kuusi , 2006)

stores, there were limitations in retail quantities and the service was made awkward and troublesome with no visible bottles (except for the bottoms of the flasks), no recommendations and almost totalitarian surveillance of the customers. There were even detectives tasked with finding possible alcohol abusers. The licensed premises were also picky on their customers and the consumption was supervised with almost ritual-like behavior.

As for the license decisions made by the supervisory board of the alcohol monopoly pre-1969, there is only evidence that the board favored more classy restaurants over common joints. Below we can see the division of the alcohol serving rights granted by different licenses before and after Keskiolutlaki entering into force.

	Väkiviinalaki (1932–1968)	Amount in 1963
A-rights (including club rights)	All drinks (club rights: all drinks, alcohol serving limited to the members of the club; augmented club rights: serving also to the club guests)	458
B1- & B2-rights	B1, maximum of 25 % ABV B2, maximum of 17 % ABV	127
C-rights	Malt drinks	38
	Keskiolutlaki (1969–1995)	**Amount in 1970**
A-rights	All drinks	930
B- rights	Maximum of 25 % ABV	171

Table 2. The license types and their respective amounts in 1963 and 1970
(Lampela, 1964) (Häikiö, 2007)

As we can see from the above, the existing premises were mostly limited to the A-rights category both before and after 1969. After the legislative change and by the end of 1969, most of the former B- and C-rights premises were converted one notch up, with the C-rights completely abolished (Häikiö, 2007). In addition to their drink selection, the licensed premises were also categorized in standards of service, which equals price categories. The profit margin from the sales of alcohol was controlled by the law, and the fancier the place, the more profit the law enabled to take (Lampela, 1964). While I could find no proper statistics of the price division, Häikiö (2007) states that the amount of cheaper A-rights premises went up after the legislative change, thus addressing the concerns that Lampela stated in the '*1963 yearbook of Oy Alkoholiliike ab*,' stating that the pricier and fancier joints were over-represented to the places that the common preferred or afforded, which he saw as intentional strategy.

As there were no clues of intentional locational limiting of alcohol sales in the cities pre-1969, post-1969 tells a different story. As stated ample times before, the limitations to alcohol serving outlets were torn down, yet there were new guidelines to consider.

The then-new law states that the retail sales of alcohol could be allowed in municipalities of which the municipal council had agreed upon to. In addition to the prior legislation, the new law took a stance how the location decisions should be handled: The retail and licensed premises sales had to be limited to such population or business centers that the supervisory board of the alcohol monopoly deemed necessary considering the municipal population, the distance to the closest alcohol serving retail store or licensed premise and other factors. (*Alkoholilaki 26.7.1968/459, § 29, 39, 40*)

On a side note, the licenses and retail decisions seemed to be processed separately if it does not stick out from the wording, the process just was similar. Unfortunately, I did not find the exact guidelines for this location approval process, the closest such thing comes from Häikiö (2007), who states that Alko Inc. used volume and distance factors in placing the new Alko stores to the rural municipalities, and that the aggregate factors settled the store locations. As a conclusion, the approach regarding the alcohol supply location decisions was changed quite a lot with the legislative change. Previously the guidelines were on a rather macro-level, with the aim to exclude the rural people from alcohol (and to some extent the city commoners too). After the legislative change, the guidelines went into a more micro direction with the population and distance instructions.

What can we say about the endogeneity related to these location decisions? Certainly the location process was not random, that much can be stated. Will this affect the tools used and introduced in the next chapter or the integrity of the results? Perhaps, but it must be noted that on a municipal level, the main aims after the legislative change seemed to be: I) keep the alcohol supply not clustered and II) satisfy the local demand beyond certain threshold. This is not randomized, nor is it pure competition, but I have found no signs that there were other external factors that affected the location decisions, such as unemployment rates or demographic factors. Alcohol supply seemed to be maintained on a rather steady density, with municipalities with lower populations maybe hindered.

All in all, this is not a hopeless setup for an empirical study. The endogeneity of these location decisions will be further discussed in chapter 5.4.1.

4 Previous studies on alcohol availability and crime

The previous chapter focused on loose graphical analysis of the amounts of alcohol outlets in relation to overall crime per capita. However there are ample studies that researched the subtle relation on alcohol availability and crime (especially violent crimes) and I will skim through a select few of them.

Richard A. Scribner, David P. MacKinnon and James H. Dwyer have “*The Risk of Assaultive Violence and Alcohol Availability in Los Angeles County*” (1995). As implied by the article’s name, the rates of assaultive violence were studied in 74 cities in Los Angeles with alcohol outlet density, ethnicity structures, economic structures and age structures as their main explanatory variables. Their findings demonstrate a geographic association assaultive violence and the density of alcohol outlets with all the 74 cities in the Los Angeles County (both with off-sale and on-sale outlets). They however stress that the study is not able to control cross-boundary purchases. In addition to outlet density, the unemployment rates, the rate of males aged 20-29 or 40-44, and the rate of certain minorities in the city were strongly related to the violence rates. Scribner &al. also notify that cross-sectional relations with alcohol outlet density and other alcohol-related outcomes have been detected, which include civil offenses, alcohol related mortality and alcohol involved motor vehicle crashes.

A more recent study from the United States however shows a bit different results. Paul J. Gruenewald &al. have contributed to the subject with the article: “*Ecological Models of Alcohol Outlets and Violent Assaults: Crime Potentials and Geospatial Analysis*” (2005), which broadens the scope from Los Angeles to the whole state of California. The findings of the article state that the population and place characteristics contribute more to the assault rates than the alcohol outlets do. The assault rates are greater in densely populated, poor minority urban areas and in the areas adjacent to these. The rates were also significantly related to off-sale outlets in all regions, but only to on-sale outlets in the unstable poor minority-inhabited areas. This can be interpreted as the alcohol outlets having more profound effect on violent assaults with certain prevailing sociodemographic groups.

The effect of alcohol outlets on violence is further supported with more studies in different states (Texas, New Jersey), which found supporting results regarding sociodemographic factors and alcohol outlet density (see L. Zhu &al. (2004) & Speer &al. (1998)). There is also an article from Finland “*Väkivaltarikokset ja alkoholi*” by Martti Lehti and Reino Sirén

(2008) which states that majority of batteries and fatal offences are done under the influence of alcohol. The article focuses on the more recent years, and states that the amount of batteries done under influence took a steep and permanent rise during the years 1994-1996, while the amount of exported alcohol (and thus "undocumented consumption") rose with the liberated legislation. Also in recent years approximately 80% of all fatal offences and attempts at these are done under the influence of alcohol. Robberies are also considered: Approximately half of the robberies are done under the influence of alcohol.

The focus has been on violence, but how do other crimes do in comparison? The fact is that most studies are concentrated on violence, but at least one expands the scale to property crimes and overall crimes. Kabena Gyimah-Brempong (2001) shifts the focus to the rust belt, more precisely to the city of Detroit and finds that in addition to violent crimes, the total crime and property crime rates also have a significant positive relationship with alcohol availability, while controlling and sociodemographic and -economic factors.

To reflect these findings on the scope of this thesis and on the scope of Finland in the turn of 70s: There are no major ethnic minorities not at least in the same scope as in the U.S., so I do not think that this poses a threat to my thesis. Instead, municipal-level sociodemographic and –economic distributions, such the unemployment figures and the age and gender cohorts, sure could have come handy. Instead of going ahead of schedule on contemplating what kind of statistics I do not possess, let us move onwards to what I have actually gathered with the introduction to the empirical approach.

5 Empirical approach

5.1 Acquisition of data

The data used in this study was acquired from two major sources: I) *the criminality known to the police statistics from 1960 to 1975* (Central Statistical Office, 1960-1971) (Central Statistical Office of Finland, 1972-1975) and II) *the year books of Oy Alko Ab 1960-1975* (Oy Alkoholiliike ab, 1960-1969) (Oy Alko ab, 1970-1975). The statistics are limited to the years 1960-1975 due to direct personal incapability to process more data (tenosynovitis) and due to the fact that a longer data set would have required much more control measures in order for the results to be consistent (municipal amalgamations, legislative changes just to name a few difficulties). Without further introductions; the data gathered from criminality known to the police consists of the following for the consequent years[8]:

Overall crimes in the municipality in a given year

Murders, manslaughters and attempts at both of these in a given year (pooled together pre 1970 in the statistics books, so in order to have consistency the statistics are pooled together post 1970)

Assaults (minor and major) in a given year (pooled together pre 1970 in the statistics books, so in order to have consistency the statistics are pooled together post 1970)

Traffic crimes in a given year (post 1970 drunk driving was announced separately in the statistics books, but in order to have consistency, the statistics are pooled with all traffic crimes)

---

[8] I'd like to extend my gratitude to three persons that helped me in the grueling processing of the data: Marja Luukkonen, Anna Pinomaa and Pasi Saukkonen. I could not find the data presented below in any digital form, and my attempts with optical character recognition went in vain.

Drunk arrests in a given year (not listed in the overall crimes in the municipality as not a crime itself)

Population in the municipal area in a given year

What is noteworthy, is that even though the crimes against property were announced in the statistics books, I did not gather them due to the sheer amount of labor it would have taken (however there is a workaround, see the variables used in the regressions in the next subchapter and the proposals for further studies as well). Also as stated earlier, the crime types were initially announced in a rather general level (e.g. overall traffic crimes), so the more precise numbers introduced later had to be pooled according to the initial form. Anyway, from the year books of Alko Inc. I have gathered the following data for the consequent years:

Amount of Alko stores in the municipality in a given year

Amount of licensed premises in the municipality in a given year

From 1969 onwards, statistics existed for licensed places serving only medium-strength beer and for licensed retail stores serving only medium-strength beer. However, I chose not to include them for two reasons:

I) the statistics were not comparable as they were first announced in sales territories (≠municipality, rather some cluster of municipalities which I found no clear definitions for) and II) the outlets serving medium-strength beer were pretty much everywhere (see figure 5 on page 15 about the initial surfacing of roughly 20 000 places providing medium-strength beer in 1969).

There are bigger asymmetries related to the Alko stores and the other licensed premises to be exploited statistics-wise and as these are municipally controlled in the statistics in use, I will introduce a dummy variable that tries to capture the other essential changes introduced by Keskiolutlaki. I acknowledge that the omitting of medium-strength beer supply network is a liability in terms of the statistical power and might even be a source for omitted variable bias, but the dummy variable will help to combat this (more of this in the chapter 5.2).

Finally, it must be admitted that there are limitations especially with the criminality known to the police data: The municipalities are clustered according to the hundreds (county divisions, kihlakunta in Finnish) from 1960 to 1970, and thus I had to pool the later data according to these hundreds in order to have consistency. Later municipal amalgamations and changes in the statistical bookkeeping are also taken into consideration with the pooling process. The process is done according to the information regarding municipal amalgamations provided by the Association of Finnish Local and Regional Authorities (2015, Kuntaliitto).

5.2 Variables defined & descriptive statistics

The data compiled by me is available upon request. It is by no means easily approachable, but a separate record is kept of all the municipal amalgamations and corresponding adjustments of the data (both due to real life and statistical bookkeeping).

5.2.1 Municipal statistics

Below are listed the variables based on the municipal statistics and their respective descriptive statistics. In order to give an overview how the statistics changed, the descriptive statistics are listed as municipal averages ($n = 269$) for the years 1960 and 1975. Bear in mind, that the municipality X might refer to a single municipality or a municipal cluster that had to be pooled according to the hundreds introduced in the '60s crime statistics.

The crime statistics are measured per municipal capita (1000) for the sake of comparability.

Explained variables:

Overall crime rates in municipality X/Municipal population (1000): Four different variables for the following annual crime rates: Overall crimes, fatal violent offences and attempts at these, batteries (major & minor) and traffic crimes. The fatal violent offences consist of murders and manslaughters.

Other crimes in municipality X/Municipal Population (1000): Represents the annual crimes that are left after deducing violent crimes and traffic crimes from the overall crimes. This is the workaround for property crimes mentioned earlier, although it contains everything from conspiracies against the state to animal cruelty.

Drunk arrests in municipality X/Municipal Population (1000): The annual rate of drunk arrests (not a crime itself!).

Overall crime/pop.	Mean	St.dev.	Max	Min	Median	Kurtosis	Skewness
1960	34.88	18.74	118.29	10.17	28.92	2.85	1.52
1975	97.15	42.75	353.14	31.97	89.75	7.18	1.92
Traffic crime/pop.							
1965*	27.04	12.13	74.27	7.68	24.34	1.60	1.18
1975	62.86	34.28	293.71	17.21	56.73	13.74	2.87
Drunk arrest/pop.							
1960	15.62	19.85	96.66	0.00	7.00	4.24	2.17
1975	33.49	33.04	152.15	0.00	21.73	2.27	1.64
Other Crime/pop.							
1965*	16.05	8.43	65.86	4.26	13.61	5.29	1.83
1975	32.24	18.94	141.86	3.11	27.99	7.53	2.05
Battery/pop.							
1960	0.78	0.61	4.37	0.00	0.67	7.20	2.12
1975	2.01	1.30	7.07	0.00	1.75	1.54	1.19
Fatal off. etc./pop							
1960	0.04	0.08	0.78	0.00	0.00	29.88	4.33
1975	0.05	0.09	0.52	0.00	0.00	8.10	2.61

Table 3. Descriptive statistics of different crime types (years 1960 and 1975)

*The year 1965 is chosen due to difficulties with statistical bookkeeping (see appendix D)

The descriptive statistics for the above crime rates are quite reasonable. I will not delve deeply into the more general statistical indicators as they speak for themselves better than I can. As stated before, the crime figures rise quite a lot during the years, and as for overall crimes, the biggest cause for this are the traffic crimes, which are on the rise mostly due to the enforcement of speed limit legislation. The medians of different crime types tend to be smaller than the means, which explains the skewness figures. Overall, if we move from more general definitions to more precise crime types, the kurtosis tends to grow (except for batteries and drunk arrests). All the distributions of the crime variables are positively skewed, which means that they are asymmetrical and have a longer right tail. If we speculate that the population adjusted municipal crime level is somehow linked with the population densities, this is somewhat understandable as the vast majority of the municipal population densities are concentrated quite near each other, and are focused on less dense municipalities, while the denser municipalities, which are relatively few in numbers, might show as the right tail outliers in the crime levels (see the descriptive statistics of municipal population densities table. 6). Of course this presumes that the municipal population density is a proxy of the municipal crime level, so I will keep this purely at a speculative level.

As for the kurtosis figures, all the distributions are leptokurtic, which means that their peak is higher than a normally distributed variable would be. The greater the number, the higher the peak near the mean and the greater the declination around the peak. Exceptionally high peak is encountered with fatal offences and attempts at these at 1960, which is understandable as vast majority of counts is zero-valued[9]. As we take a look at the exceptionally high skewness as well, the use of the Poisson regression can be seen justified (more of this in its own chapter). These numbers come down when we reach 1975, thus the distribution is more symmetrical.

To conclude the distribution changes of different crimes, drunk arrests, batteries and fatal offences per population become more symmetrically distributed and flatter along the years, while the overall crimes', traffic crimes' and other crimes' distributions become increasingly asymmetrical and peaked.

Explanatory variables:

Distance to Alko store (km): The avian distance from the center of the municipality X to the center of the nearest municipality with an Alko store. Estimated annually and thoroughly explained in the next subchapter.

Distance to Alko	Mean	St.dev.	Max	Min	Median	Kurtosis	Skewness
1960	30.30	32.16	269.20	0.14	24.02	19.41	3.46
1975	17.25	14.44	112.57	0.11	13.54	7.67	1.95

Table 4. Descriptive statistics of distances to Alko (years 1960 and 1975)

As seen above, the average distances to Alko drop from 31.30 to 17.25 kilometers during 1960-1975, and the median distance and standard deviation also drop (see figure 6 on page 32

[9] The 'fatal offences and attempts at these' are population adjusted even though their median is zero. While this might inflate figures in smaller municipalities, this is done in order to give a better estimation of the overall aggregated figures. Comparing cities with rural municipalities is more correct with population-adjustment than without it. Ultimately this problem is tackled with the Poisson regression.

on how the average distance varied throughout the years). The kurtosis figure comes also down and thus the distribution "flattens" throughout the years. The distribution also becomes less skewed with the introduction of more Alko stores. This makes sense, as with more Alko stores, the extreme distances become smaller and the distribution becomes more symmetric.

Licensed premises dummy (1, 0): Indicates that there is at least one licensed premise in municipal area in the given year. Not to be confused with the licensed premises selling only medium-strength beer.

License dummy	Mean	St.dev.	Max	Min	Median	Kurtosis	Skewness
1960	0.30	0.46	1.00	0.00	0.00	-1.25	0.87
1975	0.92	0.27	1.00	0.00	1.00	7.48	-3.07

Table 5. Descriptive statistics of license dummies (years 1960 and 1975)

First and foremost, it must be noted why the distances of licensed premises are not calculated, while the distances to Alko stores are. This is due to the fact that with the licensed premises, the consumption of alcohol happened in the municipal area while the alcohol bought from the Alko stores could be consumed elsewhere. This supposedly limits the alcohol induced crime to the municipalities in which the premises existed. Additionally implementing two different distances would prove laborious.

For the descriptive statistics, the licensed premises become a common sight in the municipalities throughout the years, which can be seen from the change in the means and the changes in skewness as the distribution changes the orientation of its long tail.

Keskiolutlaki (1, 0): A dummy indicating Keskiolutlaki's entry into force. Gains the value of 1 onwards from 1969. Captures the additional alcohol supply (places serving and selling only medium-strength beer) and loosened control (age limits) that came with Keskiolutlaki and is used as an interaction term with the distances of Alko stores and licensed premises dummies in order to capture the effects of additional alcohol supply and legislative changes on these.

Distance to Alko store X Keskiolutlaki: Interaction variable in which the Alko store distances are paired with Keskiolutlaki entering into force, which is used as a treatment effect in order to distinguish the effect of Keskiolutlaki on Alko stores over time.

Licensed premises dummy* X *Keskiolutlaki: Interaction variable in which licensed premises dummies are paired with Keskiolutlaki entering into force, which is used as a treatment effect in order to distinguish the effect of Keskiolutlaki on licensed premises over time.

Population density in municipality X (population/km^2): The annual population or municipality X per the municipal area (square kilometers). Used to estimate the effect of the municipal population density in crimes. This can be seen as a proxy for the perceptive capabilities of the police force with a somewhat underlying assumption that the perceptiveness of the police force is reduced in less densely populated municipalities (and vice versa). The municipal area is calculated with ArcMap component of the ArcGIS software by Environmental Systems Research Institute and the municipal populations are taken from the 'Criminality Known to Police' –statistics.

Population density	Mean	St.dev.	Max	Min	Median	Kurtosis	Skewness
1960	65.21	213.93	2868.50	0.33	13.70	112.44	9.31
1975	86.11	274.98	3182.49	0.36	11.92	65.73	7.03

Table 6. Descriptive statistics of population densities (years 1960 and 1975)

The average population densities rise throughout the years rise as does the standard deviation, while the median becomes smaller. This can be seen as a product of the heavy urbanization that took place during that time. The distribution also becomes flatter and less skewed to the left than before, which means that the distribution of the population density became more symmetrical along the years. All, in all, the distributions for both years are highly peaked and concentrated to the smaller values.

Time trend (1-16): Linear indicator for capturing time trends for different regressions for the time span of 16 years.

Speed 19XX (1, 0): Dummy variables indicating speed limits entering into force. These variables are used in the regressions that explain traffic crimes and overall crimes. Four different variables indicating changes in the speed limits applicable to the whole country, gaining the value of 1 onwards from their respective years of entering into force. These are on a rather general level, and there were also more minor regional experimentation, but were chosen not to be included. No other major traffic legislations entered into force as far as I know. In addition, it has to be acknowledged that the size of the vehicle fleet was on the rise, which most likely also correlated with the amount of traffic crimes.

Speed 1968 (1, 0): General speed limits of 110 km/h during summer time and 90 km/h during winter. (Salusjärvi, 1990)

Speed 1972 (1, 0): General speed limit of 80 km/h for young drivers. Changes in vehicle-based speed limits. (Salusjärvi, 1990)

Speed 1973 (1, 0): Speed limits shown in road signs for the main roads. The general speed limit is 80 km/h unless a road sign suggests _a lower_ limit. (Salusjärvi, 1990)

Speed 1974 (1, 0): Speed limits shown in road signs for the main roads. The general speed limit is 80 km/h unless a road sign suggested other speed limits. (Salusjärvi, 1990)

One has to notice that most of the distributions are quite skewed and contain rather extreme outliers, and thus for them logarithmic transformation is used in the regressions (more about this on Appendix F).

5.2.2 The distances of liquor stores

To give a more precise view on how the alcohol supply affected crime and to control plausible spillover effects, the distances of Alko stores are taken into consideration. The change of distance to the nearest Alko store after the legislative reform itself does not tell that much of individual alcohol consumption, but can be used as a proxy to model consumption possibilities of alcohol over 4.7 % ABV. Alko Inc. kindly provided me with the annual information of individual Alko stores as well as the information on the licensed premises on a municipal level (108 Alko stores in 1960, 194 Alko stores in 1975).

The municipal distances to the nearest Alko stores are computed with the ArcGIS software in the following manner (See appendices A and B for a bit more thorough explanation and images of the process):

The central point of each municipality is used to calculate the shortest Euclidian distance to the geographic center of the nearest municipality with an Alko store. The corresponding coordinate of the nearest Alko would be the center of the municipality due to sheer simplification and due to the fact that obtaining the coordinates for the Alko stores would be time consuming as heck, and secondly, if there is anything significant to be found it is found with simpler measures.

If the municipality of residence had an Alko store in its region the estimated distance would be zero as the geographical points used in the estimation would be overlapped. In order to combat this, basic geometry is used to calculate the distance from the municipal centroid as a circle radius of the municipal area which was estimated with the ArcGIS software. This radius is then used to approximate the distance in the following fashion:

$$Distance\ to\ Alko = 0{,}5 * \frac{approximated\ radius}{amount\ of\ Alko\ stores\ in\ the\ municipality}$$

If there are multiple Alko stores in the municipal area, the radius is diminished even more in an arbitrary fashion (divided by the number of Alko stores). I do acknowledge that this is a work-around, but I think that it reflects reality better than a zero-distance estimates as the Alko stores were likely established near clusters of people instead of clusters pinewood, moose and other forest critters (even though the somewhat conflicting purpose of Alko Inc. is still to make a profit while limiting alcohol consumption).

There were also certain practical problems related to this distance estimation. Since obtaining an older map of the municipal division containing the coordinates necessary for the distance estimation in fully digital form is nigh impossible, a newer version from 2015 (National Land Survey, Esri Finland , 2015) is used as a base map. The newer map is then aligned with a municipal division map of the year 1960 (Maanmittaushallitus, 1960) in a process called *georeferencing* with the ArcMap. The municipal borders are then drawn as polygons according to the situation in 1975. As stated earlier, certain municipalities have vanished from the face of the Earth since the ‘60s (or rather have been incorporated in other municipalities) and also the statistics had to be clustered according the reported statistics from the ‘60s. These polygons are drawn according to these facts in mind and are ultimately used in calculating the distances to the Alko stores. The calculated distances for each consequent year provided by ArcMap are then joined with the other statistics in order to perform the regressions.

Below is illustrated the fruit of the distance estimations: How the annual average municipal distance to the nearest Alko store has changed over time and how the average municipal crime per capita has changed.

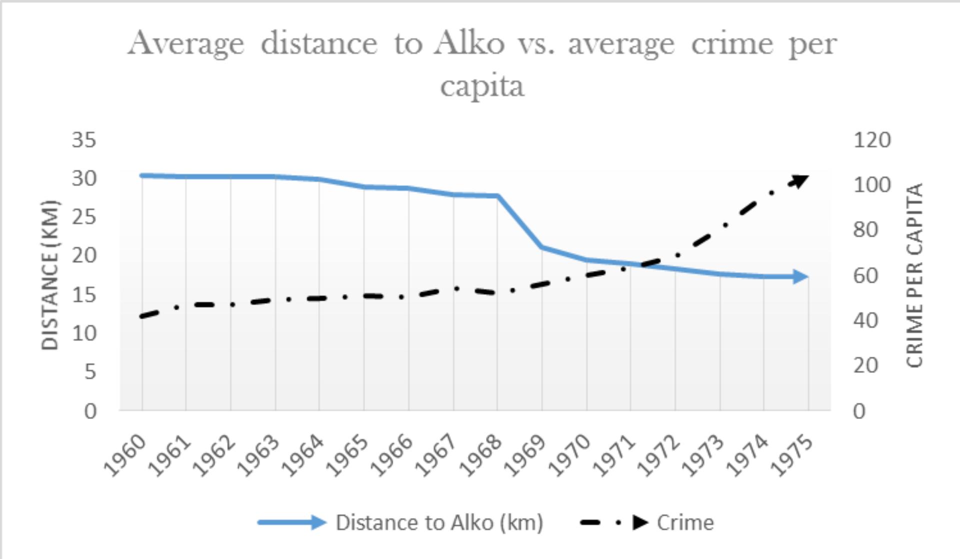

Figure 6. Average distance to the nearest Alko store in relation to crime per capita

As we can see from the graph, the estimated average distance to the nearest Alko store fell from 30.29 kilometers to 17.25 kilometers in the course of years, while the average municipal crime per capita shifted upwards. Next we are going to delve deeper into this relation among others.

5.2.3 Preliminary observations of graphical interpretation

Previously we saw how the amount of alcohol outlets rose together with crime, but could not determine any relations with the two factors. In order to take this graphical analysis further, let us see how the distances to the nearest Alko store went together with crime.

Below is illustrated a scatter plot regarding municipal population adjusted overall crime in relation to given municipality's distance to its nearest Alko store. The outliers have been dropped out to enhance readability (thus there might be some peculiar single observations in the graphs). From below we can see how the relation of average municipal crime and distance to Alko has shifted along the years 1960, 1969 and 1975.

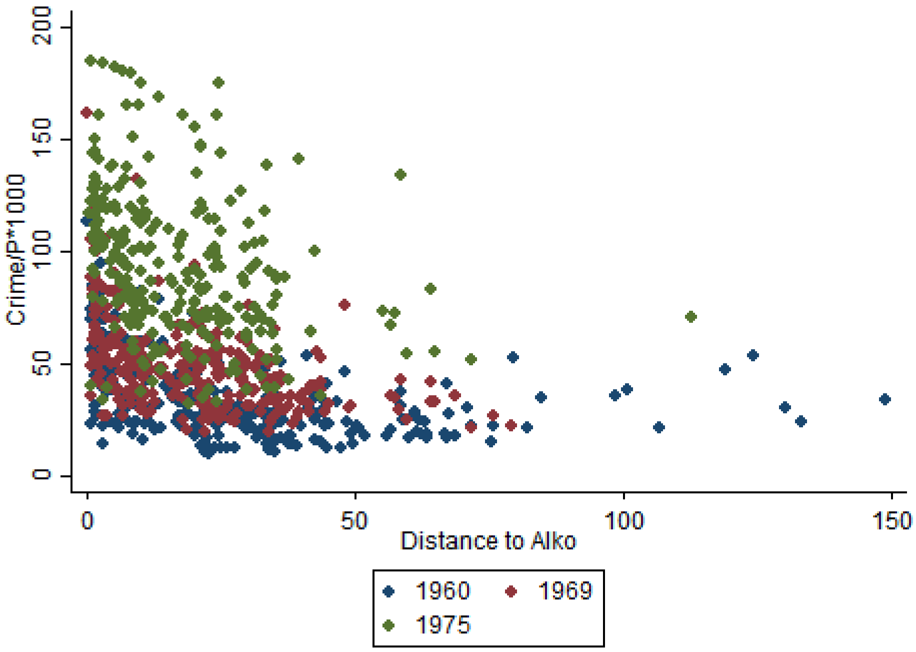

Figure 7. Scatter plot of municipal distance to the nearest Alko and crime per capita (Observations from 1960, 1969 and 1975)

The distance to Alko shows signs that there might exist a relation with the proximity of liquor stores and crime. Horizontally, the points are more dispersed in 1960, while vertically they are quite flat. This changes during the years, and while there seems to be no drastic change in Alko distances from 1969 to 1975, the average municipal crime levels shift up and become more dispersed. Figuratively speaking, the scatter plot goes through a vise and expands in the vertical dimension.

Next up is a graph concerning the relation of municipal crime and population density. In order to avoid the scatter plots looking like pixelated porridge, I have chosen to remove the observations concerning the year 1969 from the scatter plots hereafter. The reader has to trust my word and can place the missing observations somewhere in between the observations of 1960 and 1975.

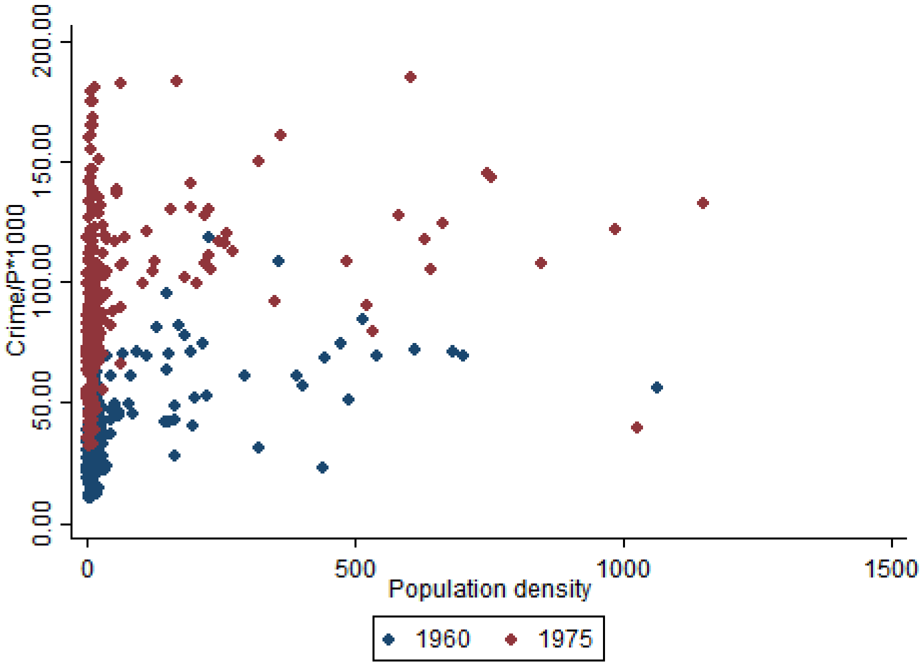

Figure 8. Scatter plot of municipal population density and crime per capita (Observations from 1960 and 1975)

While it is evident that the crime levels have shifted upwards along the years, there are no clear signs that population density affects them. If we take a look on the left side of the figure, which has clusters of low-density municipalities, the crime-levels are highly varied. It must also be noted that while the municipalities with higher densities are quite dispersed as well, they tend to be located on the upper scale of the scatter plot. As it must be noted that overall crimes might not be the best choice for interpreting the role of alcohol induced crimes, next we move away from general-level crime to more specific and severe cases of crime: Fatal offenses and attempts at these in relation to distances to Alko stores and municipal population densities.

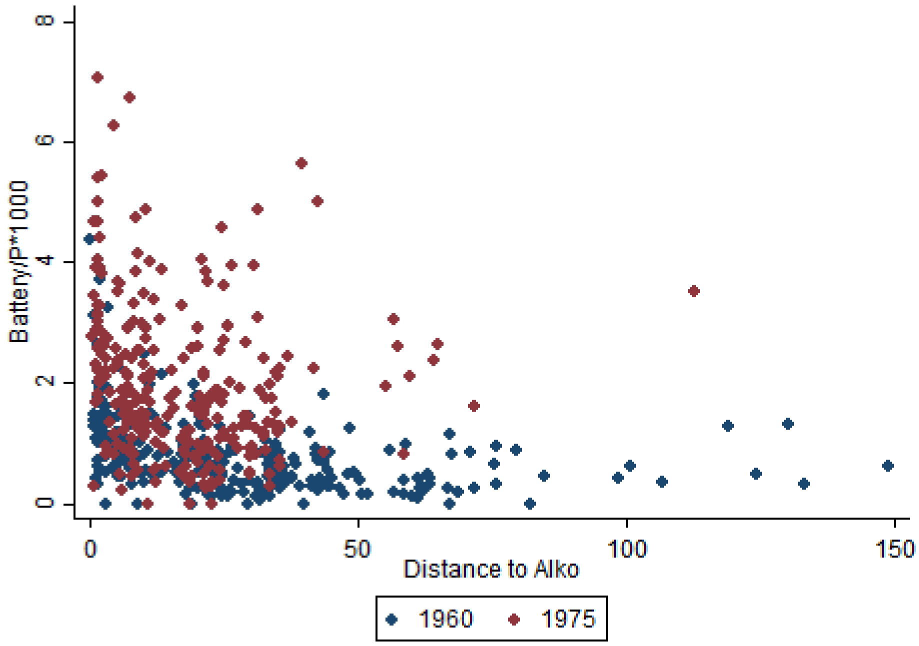

Figure 9. Scatter plot of municipal distance to the nearest Alko and batteries (Observations from 1960 and 1975)

Same kind of observations can be made with batteries as with overall crime. While the crime counts are smaller, the "intertemporal through-the-vise-transformation" of the observations also applies here. The observations are more horizontally dispersed and vertically more flat in 1960, while there seems to be more extreme levels of batteries in 1975 when the distances to the nearest Alko store have gone down.

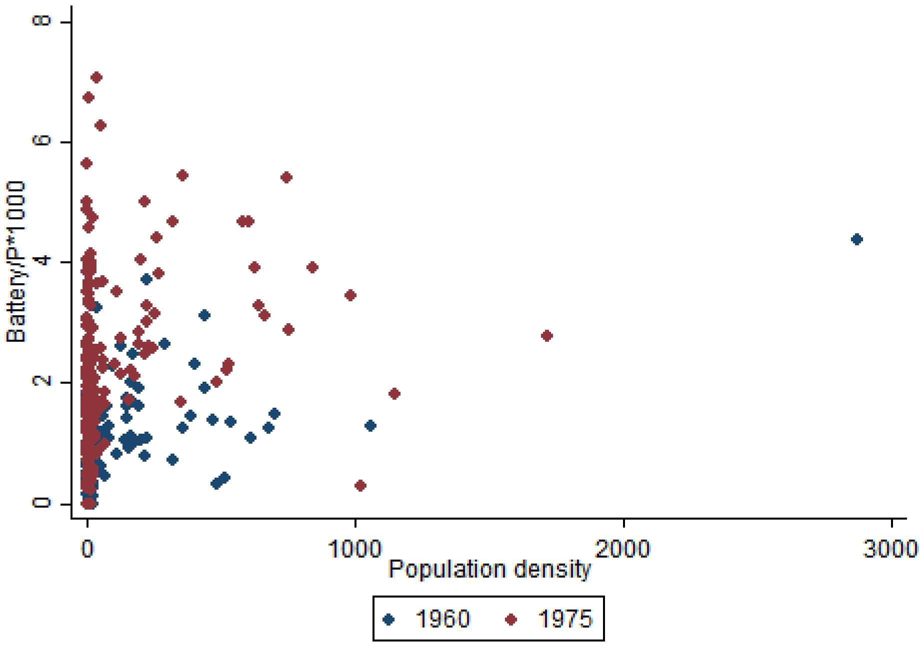

Figure 10. Scatter plot of municipal population density and batteries (Observations from 1960 and 1975)

Yet again, there are more extreme crime levels associated with 1975 and with more dense populations, but there are no clear trends among the clusters. The battery-levels seem to shift up during the years, but there is no immediately clear trend among the population densities, an observation on which I will return shortly.

Finally, we analyze the relation between the distance to Alko store and municipal population density. Below is illustrated the average municipal distance to Alko in relation with population density. As stated before, I found evidence that there was a guideline in which the Alko stores were placed to satisfy certain demand. The graph below indicates that there was a clear threshold for population densities beyond certain point to be quite near an Alko store. Of course this is somewhat self-evident, as municipalities with lower population densities have bigger areas and thus bigger distances and so on, but this serves as an indicator that the process of assigning new Alko stores most certainly was not randomized in terms of population density.

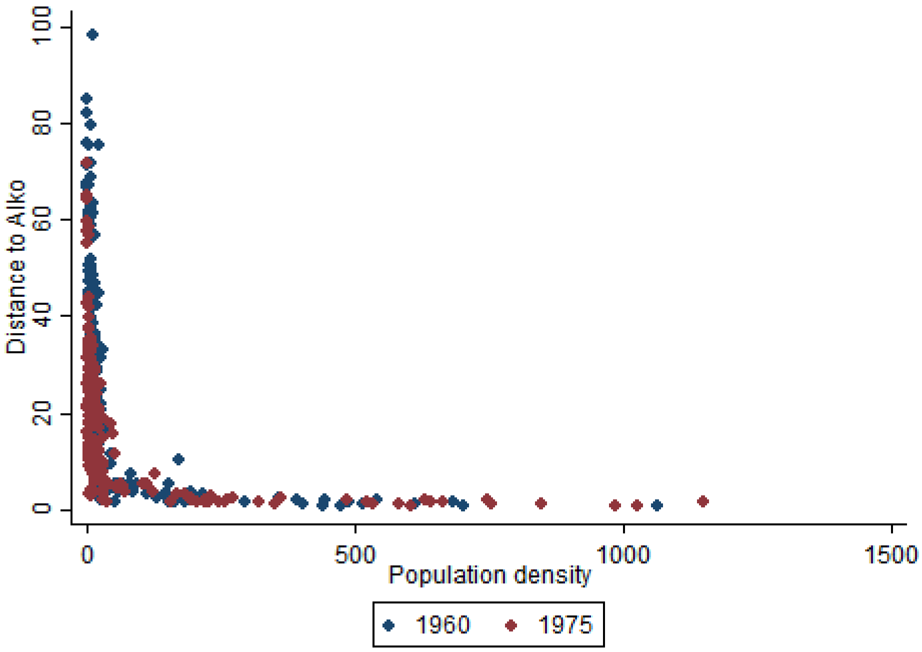

Figure 11. Scatter plot of municipal population density and distance to the nearest Alko (Observations from 1960 and 1975)

Coming back to the relations with crime and population density: Based on the loose graphical interpretation, it seems that the distance to Alko is a better indicator of crime than population density is, while population density is a quite good indicator of the distance of Alko after a certain point. With these in mind, we are going to define a quasi-experimental setup in order to better understand how crime, the distance to Alko and the municipalities go hand in hand.

5.3 Regression used and acknowledged limitations

The regressions will be performed in Stata and are based on fixed effects –approach. For crimes that are plentiful, ordinary least squares linear regression is chosen as the tool, and for less abundant crimes, Poisson regression is used. After the regression tools overview, a more specific approach will be considered in the form of differences-in-differences regression in order to try to address the differences of the cities and the rural municipalities in terms of how the alcohol supply caused crime and to scrutinize whether different levels of alcohol supply, mostly in the form of Alko stores, affected crime. First I will introduce these regression tools in the scope of this thesis.

5.3.1 Fixed effects regression

As stated before, there are loads of omitted variables on the municipal level such as ethnic breakdown, employment figures, demographic structures etc., which might have an explanatory role in the criminal statistics. In order to have credible results, we need to take these unobserved factors somehow into consideration. This is done with the fixed effects regression.

In order to save time and space, I will not go to the specifics of the OLS, instead I will jump directly to a generalized equation of a two-period fixed effects model. The following is modelled after "*Introductory Econometrics – a Modern Approach*" by Jeffrey M. Wooldridge (2013):

A two period linear regression is considered:

$$y_{it} = \beta_0 + \delta_0 d2_t + \beta_1 x_{it} + a_i + u_{it} \tag{1}$$

i denotes the unit that the statistics are announced (municipalities in this thesis). $t = 1{,}2$ denotes the time period. $d2_t$ is a dummy variable indicating year 2. The intercept for t=1 is β_0 and the intercept for t = 2 is $\beta_0 + \beta_1$. y_{it} is the dependent variable, x_{it} is some independent variable and u_{it} is the time-varying or idiosyncratic error that represents the unobserved factors that change in time.

The main component for the fixed effects regression is a_i, which captures all the unobserved time-constant factors that affect y_{it}. To be specific, the regressions will include a dummy indicating every unit that the statistics are announced in. The underlying assumption is that these units differ from one another in some unobserved factors and these dummies will try to control and isolate the effect of each municipality on crime. As the name of this approach states, these unobserved variables that affect different crime rates in different municipalities are seen as constant over time and thus different municipalities have fixed effects on crime rates. In addition to the fact that a_i should be fixed in time, another underlying assumption is that the explanatory variable x_{it} should not be correlated with u_{it} nor with a_i (Wooldridge, 2013).

Regarding the 16 year time period, it is highly unlikely that these factors stay absolutely constant, but I reckon they will be constant enough. Although, should the regression fail due to explanatory variables being correlated with u_{it}, additional time-inconsistent variables should be introduced to the fray. Examples of such variables might be municipal police resources and municipal car fleets. As for the explanatory variables being correlated with the fixed municipality variables: There is some endogeneity related to the location decisions of the alcohol supply network as it is not randomized. Thus a_i is most likely correlated with x_{it}, but this is overlooked for now.

5.3.2 Poisson regression

All the crime statistics used can be seen as count data, which is fine for OLS when it comes to big numbers, but if the incidence to population ratio gets smaller and there are more and more zero-valued counts for certain crime statistics, we have some issues to address. This is due to the fact that linear regressions make wrong kind of assumptions on how the count data looks like on when the observations are few, their mean is small and thus the distribution is likely skewed (Grace-Martin, 2010).

As D. Wayne Osgood states in article "*Poisson-based Regression Analysis of Aggregate Crime Rates*" (2000), OLS is not a very good regression for modelling relatively rare crimes for two reasons: First because the precision of the estimated crime rate depends on the population size, variation in population sizes across the aggregate units will lead to violating the assumption of homogeneity on the error variance (i.e. its variance is not constant among observations), which lead to larger errors of prediction on smaller populations. Secondly, normal or even symmetrical error distribution of crime rates cannot be assumed when crime counts are small or even zero. For small populations, a crime rate of zero is even expected on certain crimes and thus OLS can generate a bias that can even lead to negative predictions which are theoretically impossible (Crawley, 2007).

These issues can be addressed with the Poisson regression, which itself is a maximum likelihood –based regression, which uses Poisson distribution to observe any number of discrete events given an underlying rate of events. The occurrence rate of these events is approximately Poisson-distributed, its expected value dependent on explanatory variables. A characteristic of the Poisson distribution is that its mean equals its variance. (Smith, 2015)

The regression is simply modelled as follows:

$$\ln(\lambda_i) = \sum_{k=0}^{k} \beta_k x_{ik} \tag{1}$$

$$\mathrm{P}(Y_i = y_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \tag{2}$$

Equation (1) is a regression equation relating the natural logarithm of the mean and expected number of events for case i, λ_i, to the sum of the products of each explanatory variable x_{ik}, which is multiplied be a regression coefficient β_k. The constant, β_0 , is multiplied by 1 for each case. Equation (2) shows the probability of y_i the observed outcome, which follows the Poisson distribution. Thus the expected distribution of the incidence counts and corresponding distribution of the regression residuals depend on the predicted mean count in (1). The regression coefficients reflect proportional differences in incidence rates due to the logarithmic transformation. (Osgood, 2000)

Poisson-based models are built on the assumption that the underlying data takes the form of nonnegative integer counts of events, such as crime rates. Osgood states that population adjusted Poisson regression presents a more precise estimation than the mere incidence counts. In his article, he pits municipal *crime counts* against *population adjusted crime rates*: As the populations increase, the range of likely crime rates decrease, while yet at the same time the range of likely crime counts increases. The population adjustment can be modelled in traditional Poisson in the following manner:

$$\ln\left(\frac{\lambda_i}{n_i}\right) = \sum_{k=1}^{k} \beta_k x_{ik} \tag{3}$$

$$\ln(\lambda_i) = \ln(n_i) + \sum_{k=0}^{k} \beta_k x_{ik} \tag{4}$$

n_i is the population size for the unit, λ_i is the expected amount of incidents per a time unit, x_{ik} represents the explanatory variables[10]. The population adjusted crime rates adjust the amount of opportunity that an event has, as with an increase in the population, the possibility of crime can be seen to increase (Grace-Martin, 2012).

[10] With Stata the exposure is taken into account with the *exposure(varname)* –command.

Osgood says, that Poisson regression standardized for the municipal population sizes acknowledges the greater precisions of rates based on larger populations, and thus addresses the problem of the heterogeneity of variance. In other words, Poisson does not assume homogeneity of variance, instead it assumes that residual variance is expected to be a function of the number of incidents measured. Additionally Osgood states that the zero-values presented with low crime rates do not pose a threat to the Poisson regression, while they might cause bias in the OLS. Instead of directly predicting the probability of an incident, the model computes the probability of the observed count of incidents based on the predicted value for the mean count.

The interpretation of these results is most easily done by calculating the incidence rate ratios (IRR), which means, the relative incidence rate of crime as alcohol supply increases in the scope of this thesis. This ratio implies the change in the odds if all the other independent variables are held constant, except for one. If E_j is the exposure, the expected number of events C_j will be:

$$C_j = E_j e^{\beta_0+\beta_1 x_{1,j}+\cdots+\beta_k x_{k,j}} \tag{5}$$

$$= e^{\ln(E_j)+\beta_0+\beta_1 x_{1,j}+\cdots+\beta_k x_{k,j}} \tag{6}$$

Thus if the dummy-variable x_i takes the value of 1 in the first observation and 0 in the other, the IRR for a one-unit change in x is:

$$\frac{e^{\ln(E)+\beta_1 x_1+\cdots+\beta_i(x_i+1)+\beta_k x_k}}{e^{\ln(E)+\beta_1 x_1+\cdots+\beta_i x_i+\beta_k x_k}} = e^{\beta_i} \tag{7}$$

(StataCorp, 2015)

As there is an excessive number of zeroes in the municipal violence statistics (both in offences & attempts against life and batteries), a valid concern would be why are zero-inflated models not in use? This is due to the fact that I do believe that these are “true zeroes”, i.e. there is no external factor outside that causes any distortion or downward bias in the municipal statistics. These offences just are so rare that there is an ample amount of zeroes in the statistics.

Another factor that has to be considered, is the so called overdispersion of the observed variance, which means that the variance is greater than the mean, which should be equal according to the underlying assumption of Poisson distribution. This could lead to the model not being appropriate. However, the model is not as weak as it sounds: The presumptions can be relaxed with the robust standard errors the estimator of the variance-covariance matrix does not assume mean equalizing the variance, nor it does not require homoscedasticity of the variance (Gould, 2011) (see appendix E for more information about robust standard errors).

The fixed effects are also considered with the Poisson regression, which are used in the same manner as in OLS regression to control the time-invariant effects of the municipalities.

5.3.3 Differences-in-differences regression

The Alko stores are not randomly assigned to different municipalities and thus this is not a natural experiment. Some form of control is needed in containing these plausible differences that arise with the endogenous direction of the alcohol supply.

Differences-in-differences (DID) is a tool to estimate treatment effects when comparing the pre- and post-treatment differences in the outcome of a treatment and control group and thus is a suitable tool to estimate plausible differences in alcohol-induced crime. First I will go through the theoretical framework in this chapter and then in the next subchapter I will try to form and justify a quasi-experimental setup considering the endogenous nature of the alcohol supply and the limitations of the regression models.

Back to DID, which is quite well illustrated in graphical form:

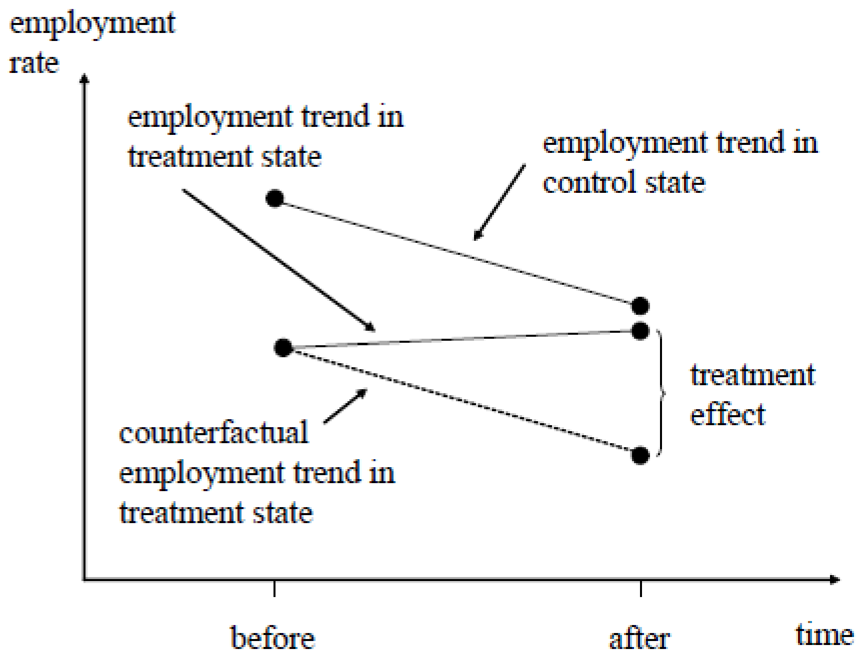

Figure 12. Causal effects in the differences-in-differences model
(Angrist & Pischke, 2008)

In the figure above, there are two states, of which one experiences treatment while the other stays the same (the control state). The main idea of the figure goes as follows: 1) Difference between the groups' pre-treatment trends is their normal difference, 2) difference in the post-treatment trends is their normal difference and causal difference and 3) difference of these differences is the causal effect of the treatment. Thus, as the model's name states, the causal effect is interpreted as the difference of the post-treatment and pre-treatments trend differences as everything else is held constant. The counterfactual trend refers to the trend before the treatment. The key identifying assumption is that the difference of the trends would be the same in both states without the treatment (so called parallel trends assumption). The treatment induces a deviation to this pre-treatment trend, which can be interpreted as its causal effect. Although the treatment and control trends can differ from each other, their difference is captured by the state fixed effects. (Angrist & Pischke, 2008)

To put this into more mathematical context, Joshua Angrist & Jörn-Steffen Pischke use the following example regarding minimum wage legislation's effect on fast food restaurant employment in neighboring states in their book an '*Mostly harmless Econometrics: An Empiricist's Companion*' (2008). The setting is simple: New Jersey encounters a minimum wage lift and Pennsylvania is used as the control state.

To model this more formally, take y_{1ist} and y_{0ist}, which imply:

y_{0ist} = fast food employment at restaurant i at period t if there is low state minimum wage

y_{1ist} = fast food employment at restaurant I at period t if the is high state minimum wage

s denotes state: New Jersey (NJ) and Pennsylvania (PA). t denotes time, of which there are two points: February (Feb) and November (Nov). February is seen as the pre-period, before the minimum-wage increase, while November is seen as the post-period. These are both potential outcomes and in practice we only get to see one or the other, e.g. the latter is seen in New Jersey in November. The change of employment in New Jersey is compared to the change of employment in PA before and after the minimum wage increase in NJ. The minimum wage in PA is the same as in NJ before the legislative change and stays the same in both periods.

Following additive mechanism is used to model the Diff-in-diff approach

$$E(Y_{0ist} \mid s, t) = \gamma_s + \lambda_t \tag{1}$$

Basically this equation implies that in the absence of a wage change, employment is determined by the sum of a time-invariant state effect and a year effect that is common across the states. To model wage change into this we introduce a dummy D_{st}, which indicates a high-minimum wage state. We assume that the employment difference between high and low minimum wage states is constant:

$$E(Y_{1ist} - Y_{0ist} | s, t) = \beta \tag{2}$$

And if we combine (1) and (2) we get

$$Y_{ist} = \gamma_s + \lambda_t + \beta D_{st} + \varepsilon_{ist} \tag{3}$$

Where $E(\varepsilon_{ist} \mid s, t) = 0$. If we introduce the states:

$$E(Y_{ist} | s = PA, t = Nov) - E(Y_{ist} | s = PA, t = Feb) \tag{4}$$

$$= \gamma_{PA} + \lambda_{Nov} - (\gamma_{PA} + \lambda_{Feb}) = \lambda_{Nov} - \lambda_{Feb}$$

And for New Jersey, as the $D_{NJ,Nov}$ gains the value 1:

$$E(Y_{ist}|s = NJ, t = Nov) - E(Y_{ist}|s = NJ, t = Feb) = \lambda_{Nov} - \lambda_{Feb} + \beta \qquad (5)$$

Thus the population difference-in-differences is:

$$[(Y_{ist}|s = PA, t = Nov) - E(Y_{ist}|s = PA, t = Feb)] \qquad (6)$$

$$- E(Y_{ist}|s = NJ, t = Nov) - E(Y_{ist}|s = NJ, t = Feb) = \beta$$

And thus β can be interpreted as the causal effect of the change. (Angrist & Pischke, 2008)

Naturally there are limitations to the model. Coming back to the parallel trends assumption mentioned earlier: There has to be an identifiable trend before the treatment, as otherwise the counter-factual is null and void. If we cannot properly detect the trends before the treatment, then we can mistake them for a causal effect even though the treatment would have no effect or alternatively we can identify purely random abbreviations in the employment as a trend. Thus more observations points are needed in order to identify how the control and treatment groups' trend together in order to prove that there indeed is an identifiable change in the trends.

Another critical issue is appointed by Marianne Bertrand, Ester Duflo and Sendhil Mullainathan in their critically acclaimed article "*How Much Should We Trust Differences-In-Differences Estimates?*" (2004). According to them most papers that employ diff-in-diff estimations use many years of data and focus on serially correlated outcomes, yet ignore the inconsistent standard errors. I try to acknowledge this issue by identifying plausible correlation and taking clustered standard errors (see appendix G for more of this).

5.4 Defining the experimental design

Let's pretend that there exists such a mechanism in which the proximity of alcohol supply induces crime. In addition to how the supply induces crime, we are interested in how the crime caused by the supply was affected by its location. Although the municipal time-invariant differences in crime are controlled with the fixed effects approach, the alcohol supply locations themselves might affect these crime rates in a way that does not show in the municipality fixed crime rates.

To put it more precisely: Two plausible sources for such differences arise: I) are the rural municipalities and cities different in terms of how crime is induced and additionally II) are the rural municipalities in which Alko stores were located and/or licensed premises were allowed different from the rural municipalities that received no Alko stores nor licensed premises in terms of how they induce crime? The latter concern is valid, if there is such endogeneity related to the location decisions that directed the alcohol supply in such municipalities that less crime arose. That is, is there an omitted variable(s) that affects how crime is induced and is this omitted variable a determinant of how the alcohol supply was directed? Even if I could not find any signs of such policy measures, it does not mean that they did not exist. To justify these thoughts in a historical perspective, as stated before, a factor to consider is that the effects of the alcohol supply might have been profound in the countryside, as before the legislative change, the prohibition-like conditions in the countryside were somewhat justified by stating that the rural folk could not handle liquor.

A rather silly example is in order to clarify what I mean with the location differences' effect on alcohol supply: Consider two tennis courts, one of them a clay court and the other one a grass court. Both of them are hit with a standard construction stage 3 tennis ball, which bounces differently from the surface of these courts. Same kind of thought process can be applied to the alcohol serving outlets and the municipalities in which they are introduced: An Alko store bounces (or might bounce) differently if it hits a rural municipality rather than a city. If the alcohol supply was directed in such a manner, this effect hits the statistical inference with a tennis racket.

If the municipalities in which alcohol supply was directed were somehow different from each other, we can try to identify their effect on crime with some kind of an experimental setup involving dummy variables that interact with other variables indicating alcohol supply. With this kind of setup we should be able to control how the 'choice of court affected the bounce'. If the interaction variables have different coefficients that are statistically significant, we can see that there are differences on how they respond alcohol supply.

The following sub-chapters focus on in forming this experimental design and in analyzing preliminary significance of the chosen design.

5.4.1 The trouble of defining a solid experimental design

As stated before on chapter 3.3, the Alko stores were not randomly assigned and thus there needs to be some form of design that takes care of the related endogeneity. After declaring that there is a plausible problem with endogeneity related to the location decisions of alcohol outlets and therefore plausible differences on how different municipalities reacted to this supply in terms of crime, the next problem is how to measure or define the "difference" of these municipalities? First and foremost, it must be stated that the ultimate quasi-experimental setup that is used in the regressions can be seen as somewhat crude, and thus I will go through how and why it came to shape. Intuition played a major part in defining it.

The first intuitive approach was to form a treatment group and two control groups out of the municipalities according how the distance to the nearest Alko store changed during the time period. However this approach was not chosen due practical difficulties explained shortly hereafter. The division would have gone as follows (preliminary threshold distances in brackets):

Treatment group: Municipalities of which distance to Alko drastically dropped during the time period (from >30 kilometers to <10 kilometers)

Rejected

Control group 1: Municipalities of which distance to Alko stayed relatively long during the time period (>30 kilometers)

Control group 2: Municipalities of which distance to Alko stayed relatively short during the time period (<10 kilometers)

The problem with this setup was that no matter how I tried to calibrate the setup, I could not form a sufficient treatment group, with the best amount of treatment municipalities I could generate limiting to ~10 with highly non-satisfactory distance boundaries. The same did not apply for the control groups which would have both consisted of over 50 municipalities. Ultimately I had to accept that the data just does not bend for this kind of experimental division and a new approach had to be found.

The second attempt on experimental design was inspired by Figure 11 (p. 37) indicating the scatter plot of municipal population densities and distances to the nearest Alko store: What if population density could be used as a defining factor on estimating how municipalities react to crime. After all, it is clearly linked with the distances to the nearest Alko store. Practical problems also came in the way with this kind of line of thought. These problems are pretty well summed up with the cumulative municipal population density figure below. The population densities in the above graph are expressed as municipal averages over 1960 to 1975 ($n = 269$).

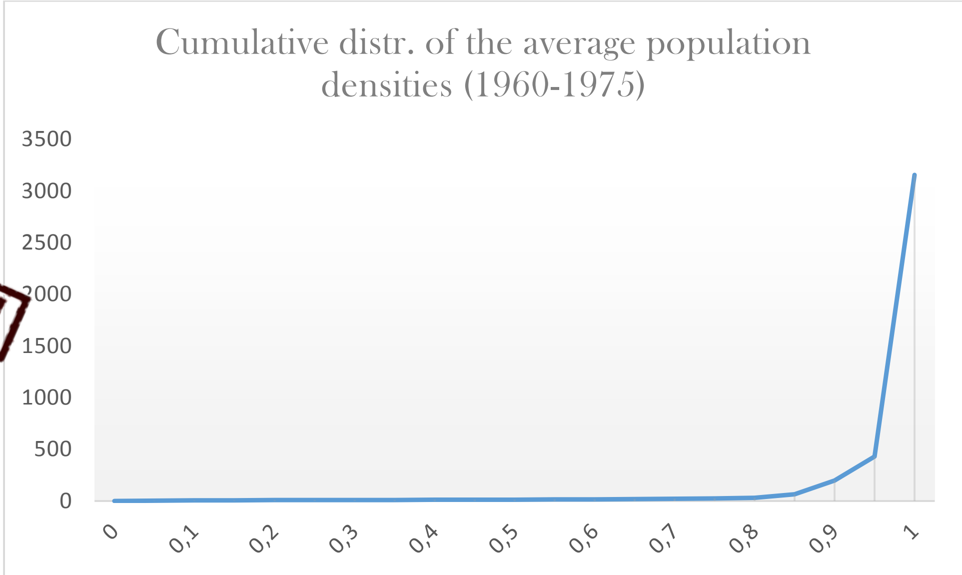

Figure 13. Cumulative distr. of the average municipal population densities (1960-1975)

The low-density municipalities are overrepresented as for the municipalities with higher population densities are highly dispersed. The question is where to draw lines, when the majority of the densities are quite near each other and then highly dispersed when it comes to greater values. While it is evident that forming a setup according to the population densities is difficult, the solution I found for the experimental design lies within the essence of the whole Keskiolutlaki, as it gives an easy way to divide the municipalities itself with the defining factors being whether the municipality had an Alko store and when it received it. That is in a sense the same way that Alko Inc. defined municipalities eligible for Alko stores, without seeing the original list of criteria.

Treatment group: “Rural” municipalities that received an Alko store after the legislative change (by 1975). Takes the alias “Alko pre-69”

Control group I: “Urban” municipalities that already had an Alko store before the legislative change (by 1968). Takes the alias “Alko post-69”

Control group II: “Rural” municipalities that had not received an Alko store by 1975. Rest of the countryside. Takes the alias “No Alko”

So what do we wish to see with this setup? Essentially three things:

I) Pre-1969 the crime trends of control groups and treatment groups are quite similar, while there may be some fixed differences regarding the crime levels (that the fixed effects -approach hopefully tackles).

II) The crime trend of the treatment group drifts away from control group II post-1969 as there should be “more alcohol available” in the treatment group.

III) The crime trend of the treatment group goes in the same direction with control group I post-1969 as their alcohol supply network should be similar.

Considering the overhauling nature of the legislative change, it is evident that if crime is affected by alcohol supply, then all of the groups encounter changes in their trends (as medium-strength beer hit the streets almost everywhere). The presumed result I) is due to the fact, that basically the treatment group should have similar alcohol availability in their municipal areas as the control group I (at least at some point after the legislative change). The presumed result II) is due to the fact, that the control group II should lack the Alko stores.

This is not your average diff-in-diff, and honestly I do not even know whether it should be called that way, but anyway it has influenced my thought process. The treatment group is chosen as it is, as essentially the municipalities in it achieve the same level of alcohol supply as the control group 1. Basically, the treatment happens in different points of time and different ways (beer joints introduced by Keskiolutlaki, Alko stores, licensed premises). It must be also considered that there is always a distance to Alko, and thus all the observations are undergoing "different levels of treatment" in all points of time[11]. A thorough overview on how the regression coefficients should be interpreted will be presented in appendix H and in the findings chapter.

So in a nutshell, basically there are two things that are achieved with this kind of setup

I) Without seeing the exact guidelines that made the municipalities eligible for Alko stores before and after the legislative change, we have now have an approximation of them.

II) The three groups sum up the different levels of change to alcohol supply that occurred across the municipalities. As the Alko stores can be seen as the most exclusive source of alcohol, it can be used as a determent factor on whether the alcohol supply affected crime levels.

With this kind of setup, the amount of treatment and control municipalities is thus satisfactory (with minimum of 50 municipalities per group). This is a rather basic way to approach the problem, but we should be able to solve whether the crime trends in the municipalities that received Alko stores post-1969 were different from the crime trends in the municipalities that had Alko stores pre-1969 and from the municipalities that did not receive any Alko stores by 1975. Back to the tennis court analogy, now we have a crude measure whether the choice of court affected how the balls bounced back. If there are no differences and at the same time we can identify an overall positive correlation on alcohol supply and crime, then we can suspect that the Alko stores were deliberately placed in municipalities in which they would cause least harm. More about the specifics of this setup can be found on the next subchapter and on appendix D.

There are problems with this setup also: We cannot detect whether a municipality that did not receive an Alko store pre-1975, would not have been eligible to receive one. As such, the line drawn between the municipalities that received Alko post-legislative change and those that did not receive one, is a bit hazy. One way to further distinguish between these municipalities would be to take the list of Alko stores established "in the future" and use them per se as the

[11] To give a simplified example: Imagine a quasi-experiment that involves non-smokers and smokers and two periods, in which after the first some non-smokers begin smoking. All the non-smokers are exposed to different levels of passive smoking at both periods, which resonates in the results as spillover effects.

treatment group. This however is also a bit hazy, as the criteria used by the Alko Inc. in order to place the Alko stores and the municipal conditions that would have made them eligible for the stores could have changed along the years. Thus, this is chosen not to be included in the experimental design.

The descriptive statistics for these groups reveal interesting things (see the whole table in appendix C). The treatment municipalities' average distance to Alko drops significantly from 50.80 kilometers to just under 12 kilometers in the course of years, and while their population density is significantly lower than control group I's, their net-migration (or birth rate) can be interpreted as positive. While there are few licensed premises in these municipalities in 1960, there is one in almost everyone by 1975.

As for the control groups: Group I (which consists of municipalities that had an Alko store before the legislative change) has considerable positive population growth while group II - municipalities tend to lose their inhabitants (no Alko store by 1975). The average distance to Alko stays quite low for group I (from roughly 9 kilometers to 5 kilometers), while for group II it stays considerably high (over 25 kilometers for the whole time). There are licensed premises in almost every municipality in group I to start with in 1960 and at least one in every municipality by 1975. For the control group II, there are almost no licensed premises in their area in 1960 and 85 % coverage by 1975. Overall, this seems like an interesting division for the municipalities and in sense achieves quite well the spirit that was tried to achieve with the preliminary compositions introduced earlier.

The next chapter will focus on the legitimacy of the differences-in-differences approach on different crime types. If there are no noticeable trend changes among the experimental groups, then there is no justification for the differences-in-differences design and it will be left out for certain crime types. Instead an "ordinary" fixed effects regression will be performed to measure the effect of the alcohol supply to the crime levels.

5.4.2 Evidence of preliminary significance and other observations

The parallel trends assumption needs to hold in order for the differences-in-differences regression to be of any use. All the figures and interpretation regarding the trends for the treatment and control groups are listed in appendix D, while some of them are presented here in order to give clarification on what we wish to see. We start off with the average overall crime rates with and without the traffic rates.

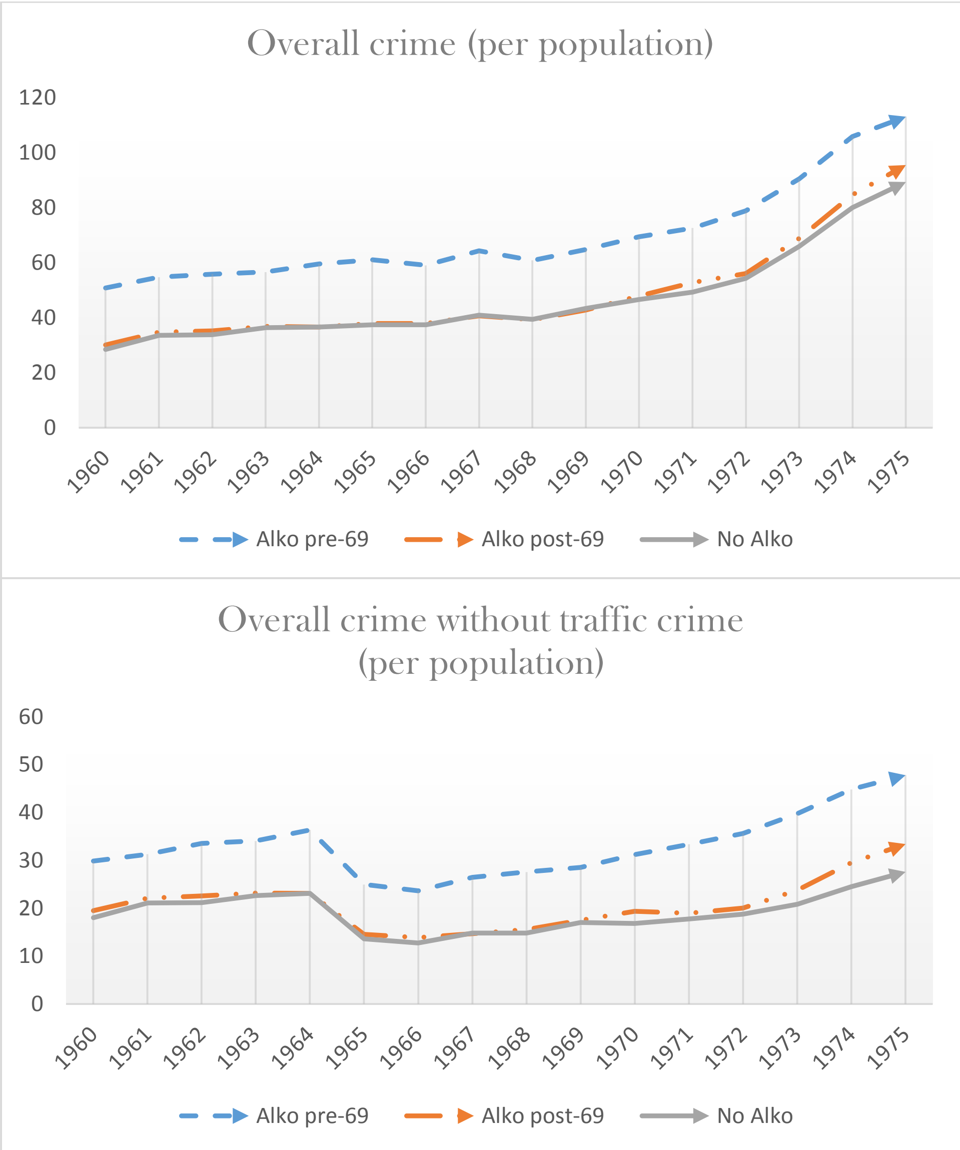

Figure 14. Interpreting the group trends for the Diff-in-diff regressions

I will cite the previous chapter on what we wanted to see with this setup:

I) *Pre-1969 the crime trends of control groups and treatment groups are quite similar, while there may be some fixed differences regarding the crime levels.*

II) *The crime trend of the treatment group drifts away from control group II post-1969 as there should be "more alcohol available" in the treatment group.*

III) *The crime trend of the treatment group goes in the same direction with control group I post-1969 as their alcohol supply network should be similar*

First observation: The trends of the different groups' pre-1969 seem to follow each other in quite a slavish manner (condition I is satisfied). There are fixed differences regarding the groups, but that much was expected.

Second observation: There is a minor divergence between the treatment group and the control group II as the treatment group shifts away 1969. This change of trends can better be seen from the lower figure, which has traffic crimes deducted from the overall average crime rates. This is something that we wish to see after the treatment period (II is satisfied).

Third observation: There is no evident difference between the trends of the treatment and control group I post-1969 (III can be interpreted as satisfied). The changes in the trends can be seen as preliminarily significant and thus the DID-regression is justified. Based on the presumptions that are assumed with alcohol-supply and crime, it is also natural that 'Alko-pre-69' -curve shifts after the legislative change as there is also a boost in these municipalities' alcohol supply.

An additional observation concerns the drop in crime rates mid-60s in the lower figure. The level of average overall crime rates excluding the traffic crimes comes down quite a bit in 1965. This is something that I have missed while I was processing the statistics, and is likely related to statistical bookkeeping changes: Some crime type previously not listed as a traffic crime started being listed as one (see figures concerning traffic crimes and other crimes in appendix D). While this is somewhat irritating, the nuisance is brushed aside by limiting the observations for the traffic crimes and other crimes to the years 1965-1975 in the regressions.

To continue with the graphical interpretation, next we have an example on exactly what kind of trend changes we wish to see and exactly what kind of trends we wish to avoid in figure 15, which compares the trend changes of drunk arrests to trend changes in fatal offences and attempts at these.

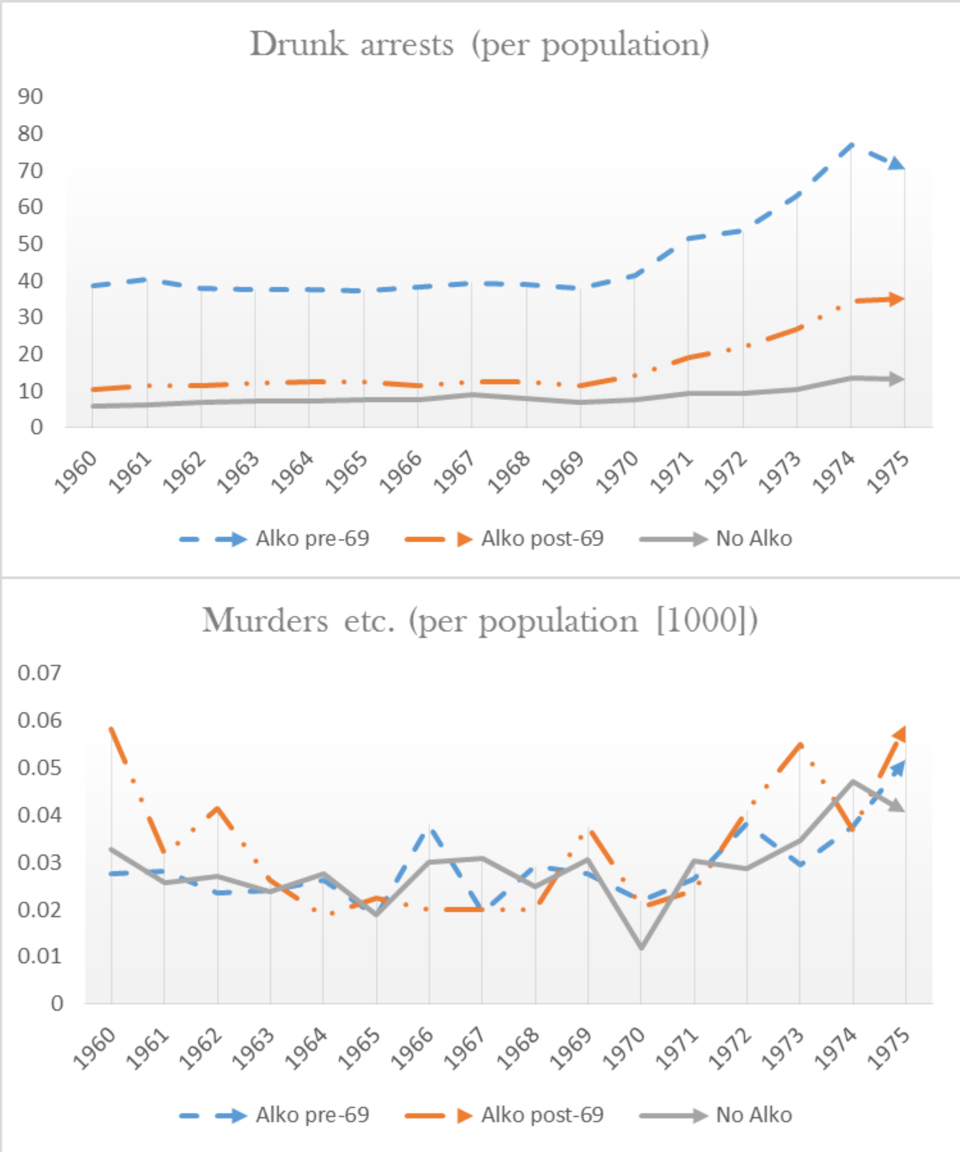

Figure 15. Favorable vs. unfavorable trend changes regarding the Diff-in-diff

There are more profound trend changes with the population adjusted drunk arrests (arrests regarding drunk and disorderly conduct) than with the overall crime rates. As for the murders, well, they seem more like a random walk process without a clear drift, which is something that we definitely do not wish to see. To conclude: After eyeing the figures containing the trends of different subsets, basing the evidence on the graphical interpretation presented in appendix D, the differences-in-differences based approach is taken into fruition with the following crime rates: Overall crime, other crime, batteries and drunk arrests.

There was no graphical evidence that the groups encountered any trend changes regarding traffic crimes or fatal offences (and attempts at these) that were presumably caused by changes in alcohol supply. Thus for them, regressions without the differences-in-differences approach are performed and listed in the appendices, but these regressions will be more like a curiosity. See the appendix D for a more thorough thought process for each crime type.

Besides the graphical interpretation, there are no preliminary significance tests performed for the groups. While a nonparametric test (Wilcoxon signed-rank test for instance) might help with the paired observations and the evidently present skewness, ultimately the highly plausible autocorrelation, which violates the assumption of independence, became the demise of preliminary significance testing. (Northwestern University Feinberg School of Medicine, 1997). The autocorrelation will be considered in the regressions however.

6 Empirics: 1960-1975

Few forewords before the grand finale: As the data contains a lot of approximations and arbitrary transformation (although I do hope I have given reasons convincing enough for these measures), it must be noted that whatever results there may be found, they must be approached with mild to moderate caution. Instead of interpreting the findings as written in stone, I will use a very conservative manner in deciphering them. Significance and magnitude are more likely reported instead of taking the numbers as godsend.

In order to fend off heteroscedasticity, skewed distributions and serial correlation, logarithmic transformation of variables, robust standard errors and clustering of standard errors are used in the upcoming regressions. Appendices E, F and G focus on the specifics of these. Appendix H gives more insight on how the coefficients of the interaction variables should be interpreted.

6.1 Findings

The regressions will be presented in two categories: First overall regressions without the experimental group design and then with the design if there was anything significant to be found. Unfortunately, the quasi-experimental group setup mostly did not yield any significant results, and is mostly excluded from here. The most interesting findings will be reported here in a rather efficient manner, while the more insignificant findings can be found in appendix i among all the regression results.

Let us start with the overall crimes. The preliminary trend figure (in appendix D) suggested that there would be a minimal shift in the trend of the treatment group (rural municipalities that received Alko stores), when comparing with the control groups (cities and really rural municipalities). The group-setup regression results however are not presented here, as they are largely not significant.

Overall crime

	Pre	Post
Distance (ln)	-0.039 (ns)	-0.0038 (ns)
License (1,0)	0.054 (*)	-0.062 (ns)
Keskiolutlaki (1,0)	-	0.129 (*)

Overall crime (%)

	Pre	Post
Distance (ln)	-3.94	-0.38
License (1,0)	5.55	-6.012
Keskiolutlaki (1,0)	-	13.76

Table 7. OLS results for overall crime

The percentages listed in the right hand side table are transformed from the logarithms. The distance is used as a direct percentage value (e.g. for the period before Keskiolutlaki, one percent more in the distance to Alko equals -~0,039% drop in crime rates, which is the distance elasticity of crime), while the dummies' coefficients are transformed to percentages according to: $100 * (\exp(\beta_i) - 1)$, which means that a unit change equals that much percentage change in the respective crime rate (e.g. a licensed premises.

Back to the results, overall crime can be seen influenced with the proximity of alcohol supply. Keskiolutlaki alone increases the crime significantly (estimated ~13.76 % increase). However, the only significant factor besides the legislative change contributing to crime is the presence of licensed premises in the municipal area before the legislative change. After the legislative change this effect is revoked with a larger negative coefficient (which is insignificant). In a way this makes sense, as the crime associated with licensed premises before the legislative change is substituted or diminished with the flood of places where one can buy medium-strength beer. As for the roles of the distances to Alko, well they can be seen as understandable, although they are not statistically significant: the base elasticity is -0,039 % change in crime for every added percent in distance to Alko. The effect after the legislative change is more marginal, but goes in the same direction. However it is nowhere near of being significant. In addition to this, if we use the group setting in the regression, all the coefficients expressing alcohol supply become insignificant, which makes the findings rather inconsistent in terms of the first regression. This is however understandable: the traffic crime trends introduced in figure D.2 in appendix D show us that the trends go in different directions for the groups and thus it is likely that the traffic crimes mess this group setting up.

To get over with the variables with a more supporting nature, speed limit dummies were used with the overall crimes and traffic crimes, and are mostly highly significant. Especially the later ones can be interpreted to majorly increase the consequent crime rates, while the first legislative changes actually seem to reduce crime. This might be due the fact that the legislation became increasingly strict along the years: Initially the introduction of speed limits actually reduced reckless driving and thus crime counts, but after a certain threshold citizens no longer deemed it that binding (some form of mass cognitive bias). Or it might be that the enforcement of the speed limits could have been less harsh during a transition period and after such a period there would be full control. However as fascinating this is, the traffic crimes were not chosen for the diff-in-diff group setup (figure D.2 revisited), but an overall regression is presented for them shortly.

Getting back to the supporting variables, population density has a negative effect on the overall crime and this effect is statistically significant in almost all of the regressions. For example one percent increase in population density decreases the overall crime rates from the geometric mean by ~-0.39 % (and the same kind of effect applies for all the regressions performed, only fatal offences and attempts at these had insignificant result). Suppose that this does not make sense that there is more crime per capita in low-density rural municipalities than in cities. At first I was a little baffled by this negative effect, however I have a proposed explanation to this: The information contained by the population density variable does not take into consideration whether the population is clustered (instead it presumes maximal dispersion), so in a sense the Alko store proximity (and other variables) probably catch the plausible effect of population density on crime better than the population densities generated by me do (see figure 11 on page 37). To put this into a more practical context: Take a municipality with a huge area. The population density variable subjects the population to this whole area, while the population might be clustered in a much smaller area, making this a rather inefficient predictor of the true density. As the Alko stores were gathered in clustered areas by definition, they probably capture these differences better than my estimations do.

The linear time trend also shows high significance: The crime levels rise in a linear fashion it seems. The time trend is also significant for all the regressions.

As for the traffic crimes mentioned earlier with the speed limit dummies: First impressions are that alcohol supply has a negative impact on traffic crimes (significance aside):

Traffic crime	Pre	Post
Distance (ln)	-0.052 (*)	0.0482 (**)
License (1,0)	0.0247 (ns)	-0.036 (ns.)
Keskiolutlaki	-	-0.439 (ns.)

Traffic crime (%)	Pre	Post
Distance (ln)	-0.052	0.0482
License (1,0)	2.501	-3.536
Keskiolutlaki	-	-35.532

Table 8. OLS results for traffic crime

The distance to Alko stores seem to decrease traffic crimes pre-legislative change, and after the legislative change this effect is revoked (both of these are significant). This makes sense, as the role of the Alko stores probably diminishes with the flood of alcohol supply. The presence of licensed premises increases traffic crimes only after the legislative change (both of these coefficients are insignificant however), and the introduction of Keskiolutlaki actually seems to lower the traffic crime rates. This can be seen as rather inconsistent, and intuitively if alcohol supply has an effect on the overall traffic crime rates, it has a rather marginal effect. The experimental group setup was chosen to be excluded from here for the sake of similar reasoning: according to the group trends, alcohol supply was not seen to predict traffic crimes, rather the other way around. My intuitive explanation to this is that the presence of alcohol supply correlates negatively rather well with the sizes of the municipal car fleets which correlate directly with the traffic crimes (see appendix D for a more thorough speculation of this).

Also the regressions explaining the so called 'other crimes' are excluded from here. Alcohol supply seems to have highly insignificant trends regarding the 'other crimes' and is deemed unworthy of gaining space here. However, now we get into more interesting crime rates, or to be precise with drunk arrests, let us call them incidence rates:

Drunk arrests	Pre	Post
Distance (ln)	-0.253 (***)	-0.119 (***)
License (1,0)	-0.079 (ns)	0.36 (***)
Keskiolutlaki (1,0)	-	-0.244 (ns)

Drunk arrests (%)	Pre	Post
Distance (ln)	-0.253	-0.119
License (1,0)	-7.66	43.3
Keskiolutlaki (1,0)	-	-21.7

Table 9. OLS results for drunk arrests

Drunk arrests can be seen highly significant in terms of the how they react to alcohol supply. The distance to Alko stores significantly decreases the amount of drunk arrests both pre- and post-legislative change (thus the presence of Alko stores increases them) and the presence of licensed premises also increases drunk arrests significantly after the legislative change. To be precise, especially the licensed premises have a huge impact after the legislative change. The spreading of the medium-strength beer supply network however, does not seem to correlate positively with drunk arrests (although the effect is insignificant, and thus not much can be said of it).

The role of the licensed premises revisited from chapter 3.3: Pre-legislative change the licensed premises were mostly fancier restaurants, which was seen as an intended strategy to reduce alcohol consumption as stated by Lampela (1964). After the legislative change, the amount of more common bars exploded, which made the licensed premises more accessible to the common man. The logic behind this finding goes as follows: in fancier establishments there are more binding social and financial constraints on getting intoxicated than there are in more sleazy bars. There is also another way to ponder this relation with drunk arrests and licensed premises: The licensed premises were rather stationary points, and it was probably easier for the police to pick intoxicated persons by just patrolling near them at certain time points or by a notification of the staff of the given premise. Thus by getting drunk in a licensed establishment one would expose herself to the public eye. Same does not apply for the liquor bought from a retail store, which could be consumed to the point of being drunk without necessarily causing public disorder. The difference between the licensed premises offering more broad alcohol supply (broad by ABV) and the premises offering only medium-strength beer is rather self-explanatory and is provided by the following equation (not to be taken too seriously in medical terms):

$$Perceived\ state\ of\ drunkness = (liters\ of\ alcohol\ consumed) * (ABV\ of\ the\ alcohol\ consumed) * (personal\ capability\ to\ handle\ alcohol)$$

The major difference being the limited ABV of the medium-strength beer. Furthermore, if we analyze the quasi-experimental group setup, there are some differences. The analysis is quite tedious as one can see from below:

Drunk arrests (%)

	Pre	Post
Distance (ln)	-0.040 (ns)	-0.239 (**)
License (1,0)	-12.71 (*)	57.46 (***)
Keskiolutlaki	-	11.51 (ns)

Group specific effects

Distance (ln)

	Pre	Post
Treat	-0.162 (ns)	-0.0078 (ns)
Ctrl 1	-0.106 (ns)	0.257 (**)

License (1,0)

	Pre	Post
Treat	3.35 (ns)	-36.05 (***)
Ctrl 1	39.09 (ns)	-

Keskiolutlaki (1,0)

	Pre	Post
Treat	-	-33.17 (ns)
Ctrl 1	-	-47.43 (*)

Table 10. Group setup OLS results for drunk arrests

These are directly converted as percentages in a similar fashion as in the right-hand side of the tables presented before. Control group II is used as the reference group (the municipalities which did not receive Alko stores by 1975). Thus the figures, in the upper row of tables speak for it (but these figures have to be taken into consideration with the other groups also).

As we can see, there are similar findings when taking a look on the upper left-hand side table as with the regression without the group setup. The licensed premises seem to have a big impact on the drunk arrests after the legislative change at least in the treatment and Control II municipalities. Nothing should be said of the cities (Control 1), as the figure pre-1969 is insignificant, yet its magnitude is quite large. Also the figure post-1969 is omitted for Control 1 due to collinearity (cities that received Alko stores before the legislative change each had licensed premises in their area after the legislative change). The distance to the Alko stores also seems to decrease drunk arrests after the legislative change (except for the cities, which have a significant positive effect for the distance. However this is not as one-sided as it first seems to be: As the effect is positive for every percent more in the distance and as there exists an approximated distance to Alko for each city that housed one as well, the presence of Alko stores can be interpreted to increase drunk arrests in the cities as well!). Not much can be said of the medium-strength beer network as the values are mostly insignificant, except for that we cannot rule out that the medium-strength beer might have less explanatory power than the places offering more potent alcohol. One more factor to be considered with the licensed premises is that there might be a pooling effect: do the licensed premises gather folk from surrounding municipalities? Do the premises gather persons from surrounding municipalities that are more likely to drink loads? The crime levels that are seen might express that they are more pooled than before, not necessarily that there exists more drunk arrests.

Moving on to more rare counts of crime. Now the logarithmic transformation for the independent variables is lifted: The skewed distributions pose no threat to Poisson regression as the approach is not trying to fit a symmetrical, normal model to the data and thus account for the high counts with high variance estimates (Nussbaum;Elsadat;& Khago, 2008) .

The Poisson regression results are presented mostly as incidence rate ratios introduced in chapter 5.3.2. If the interaction variables seem bring out something of interest, the incidence rate ratios are calculated by taking the underlying coefficients into consideration. An example is in order: There is an underlying continuous variable (distance to Alko), and we are interested in how a dummy interacts with it (Keskiolutlaki). Given the initial continuous variable coefficient, the coefficient of the interaction can be interpreted as to change to this coefficient, when the dummy variable is in effect). What is problematic, is that this effect and the significance levels associated with it, change with the levels of the continuous variable associated with the interaction variable (dummy-dummy –interactions are way simpler). However, I choose to limit the interpretation of the interaction coefficients as more general-level estimators on how a given dummy variable (legislative change or a group dummy) changes the effect of different types of alcohol supply.

Take an arbitrary example of Poisson regression coefficients: $\beta_1 x + \beta_2 D + \beta_3(x * D)$ where $\beta_1 = 0.1$, $\beta_2 = 0.2$ and $\beta_3 = -0{,}05$. To calculate the effect of the dummy (D) on a continuous variable x, we need the coefficients to have form the incidence rate ratio: $\exp(\beta_1 * x + \beta_3 * x)$. If the dummy gains the value 0, the marginal effect of x is $\exp(\beta_1) = \exp(0.1) = \sim 1.10$, which equals 10% increase per one unit of x. If the dummy gains the value 1, the marginal effect of x is: $\exp(\beta_1 + \beta_3) = \exp(0{,}1 - 0.05) = \sim 1.05$ which equals only a 5% increase per a unit of x. If we check how the difference goes with two units: $\exp(\beta_1) = \exp(2 * 0.1) = \sim 1{,}22$ versus $\exp(\beta_1 + \beta_3) = \exp(2 * 0{,}1 - 2 * 0.05) = \sim 1.105$, we can see that the gap widens (Maarten L. Buis, 2010). Instead of tabling and interpreting all the plausible outcomes and standard errors that are associated with different levels of the continuous variable, I am going to focus on the marginal effect.

So instead of calculating different coefficients for different distances to Alko and putting them into a table (and thus the different significance levels), I have chosen to just interpret the results as changes given by one unit. The reasoning behind this goes as follows: the setting is too inaccurate to begin with and the end product would have little information value. Without further interruptions, the Poisson regression results for batteries:

Battery (IRR)

	Pre	Post
Distance (ln)	0.9948 (***)	0.9985 (ns)
License (1,0)	1.066 (ns)	1.1247 (ns)
Keskiolutlaki	-	1.3768 (***)

Table 11. Poisson results for batteries (IRR)

Pre-legislative change, the distance to Alko store decreases crime by -0.0052% with one kilometer added and the effect is significant. After the legislative change this effect is insignificant. The effect of the licensed premises is insignificant on both sides of the legislative change. The introduction of medium-strength beer network highly increases the amount of batteries (~37.7%, highly significant). Alcohol supply seems to bring up the brawling side of the Finns.

Bringing the experimental group setup here yields some interesting results in terms of the Treatment and Control II groups:

Battery

	Pre	Post
Distance (ln)	-0.049 (**)	0.0027 (ns)
License (1,0)	0.2496 (ns)	0.1246 (ns)
Keskiolutlaki	-	0.2011 (*)

Group setting

	Distance (ln)			License (1,0)			Keskiolutlaki (1,0)	
	Pre	Post		Pre	Post		Pre	Post
Treat	-.0011 (ns)	-0.056 (ns)	Treat	0.2756 (*)	-0.4515 (**)	Treat	-	0.4125 (**)
Ctrl 1	0.024 (ns)	0.0049 (ns)	Ctrl 1	01755 (ns)	-	Ctrl 1	-	0.1114 (ns)

Table 12. Group setup Poisson results for batteries

The above table contains the "raw figures" from the Poisson regression. The distance to Alko stores can be seen significant for at least the Control II group pre-1969. After the legislative change, not much can be said of its role as the figures are insignificant. More interestingly, now the licensed premises seem to be on a rather significant role in explaining crime in the Treatment-group municipalities: pre-legislative change they increased crime and post-1969 the effect is somewhat revoked, with the medium-strength beer taking its role.

Indeed, as for the medium-strength beer network, it seemed to have a most prominent effect on the municipal batteries, with significant levels for both Treatment and Control II groups Turned as incidence rate ratios, treatment: $\exp(0.2011 + 0.4125) = 1.847$, which means a considerable ~84.7 % increase in batteries after the legislative change and for Control II: $\exp(0.2011) = 1.2227$, a consequent ~22.3% increase in the batteries (these are dummy-dummy interactions, if we would calculate the incidence rate ratios for the distance to Alko X dummy interactions, say for treatment group pre-1969, it would go as follows: $\exp(-0.049 * 1 - 0.011 * 1) = 0.9417 = \sim 5.82$ % in batteries for one unit change in distance, but as the interaction term is not significant, this is more likely a curiosity). It is also most likely that the batteries were increased in the Control 1 group, as even though the interaction term is insignificant, its magnitude is smaller than the overall effect and it goes in the same direction ($0.0211 > 0.1114$, both positive), which would imply that there was an increase. Intuitively, this would mean that the batteries were more affected by the overall availability of alcohol, not the presence of certain more specific forms of alcohol supply. Additionally, as opposed to the drunk arrests, batteries, (which are a real crime), have a lot smaller chance of staying undetected as they are mostly non-complainant offences. Moving on: After batteries, the next logical step in terms of the severity of crime are fatal offences and attempts at these.

Murder etc.

	Pre	Post
Distance (ln)	-0.0007 (ns)	0.0046 (ns)
License (1,0)	-0.3417(**)	0.5985 (**)
Keskiolutlaki	-	-0.5714 (**)

Table 13. Poisson results for fatal offences and attempts at these

Fatal offences and attempts at these were deemed unfit for the group setup, thus we only focus on the overall effects of alcohol supply. The effect of the distance to Alko seems to be insignificant and quite small, but after that there are some more interesting findings. Before the legislative change, there actually seemed to happen less fatal offences and attempts at these in municipalities where there were licenses premises ($100 * (\exp(-0.3417) - 1) = \sim -29\%$). Post-1969, this effect is reversed, and there is more crime in the municipalities which had licensed premises in their area ($100 * (\exp(-0.3417 + 0.5985) - 1) = \sim 29\%$). The medium-strength beer network however does not seem to increase fatal offences ($100 * (\exp(-0{,}5714) - 1) = \sim -55{,}47\ \%$). The most prominent predictor of fatal offences and attempts at these seems to be the presence of licensed premises after the legislative change, although, it must be stated that the Keskiolutlaki-dummy nearly eliminates this relation and thus it is best not to say anything definitive of the subtle relation between bottle and knife. Instead, I will continue to speculate: as is stated in appendix D, if one wishes to kill someone, it does not look at the time and place, people with a tendency of getting hyper-violent while under the influence are distributed randomly across the population and they probably have their own means of getting alcohol. Also it must be noted that the counts of murder (& attempts) probably are not as influenced by alcohol as opposed to manslaughter (& attempts). If we take the noun "a plan", which can be alternatively expressed as "*an intention or decision about what one is going to do*" (Oxford Dictionaries, 2015). Intuitively crime under the influence of alcohol is not that much planned, rather "spur-of-the-moment" describes it better[12].

[12] Even though according to the contemporary legislation, a manslaughter is treated as a murder if it is e.g. done in a remarkably cruel or brutal manner (Rikoslaki 21.4.1995/ 578, 21:2).

6.2 Proposals for further studies

The whole Keskiolutlaki serves as a huge social experiment on what alcohol availability causes on the societal level and thus is an interesting subject to study. In the scope of this thesis, I have identified three main sources on which to improve.

The accuracy of data:

Aggregated crime levels do not work that well, especially with crimes that are pooled with a broad scope. Regarding certain aggregated crime types, the legislative change might even have a mixed effect on different crime types under the same label: Certain crime counts might rise and others might fall, thus dampening the overall effect (e.g. the crime counts involving illegal distilling probably went down). The basis of the problem is that certain crime types react differently to alcohol supply, and while the effect might go in the same direction, it would make more sense to study the relation of specific crime counts instead of aggregated counts. This would also make the analysis more transparent. Thus my recommendation is to focus on the more specific crime counts rather than on aggregated figures such as 'traffic crimes'.

Another factor to consider with the accuracy of data, is the accuracy of the alcohol supply. There are some rather general-level approximations (dummies), and while the estimated distances to Alko stores might help with catching spill-over effects, they are still quite crude. My solution to this is to take population densities and real coordinates of the Alko stores and conduct spatial analysis with them. The population center of each municipality and the actual coordinates where the Alko store stood would probably yield more consistent results. Of course it is a whole different story, on where one would find such population maps or statistics from the turn of 70s.

Finally, also a longer time period after the legislative change, in order to detect whether alcohol supply has a lagging effect on crime. To conclude with the accuracy of the data: I feel like the statistical significance is a bit drowned in the rather general-level estimations on alcohol supply and in the aggregated figures regarding crime. This is made worse with the pooled crime counts in the municipalities that had to be adjusted according to the hundreds in the initial data.

The omitted variables:

How controlled is the setup? Well, it boils down to the municipal time-variant effects as the fixed effects presumably catches the time-invariant municipal level of crime. As stated in the 'previous studies' -chapter, socioeconomic and –demographic factors are likely to have a huge impact on crime and can even act as envoys for alcohol related crime. Minority-wise Finland can be seen as a rather stable place, but what I have failed to control are the municipal levels of young males (that can be seen criminally as rather active by a multitude of studies). Age and gender figures would have come handy, as would have more socioeconomic statistics such as unemployment figures and median pay, as these kinds of factors had quite a lot of explanatory power in terms of the studies conducted in the U.S. Additionally some factor controlling for the homemade alcohol would also come handy, even though it might be rather constant municipality-wise. I would have greeted statistics of police resources and of municipal car fleet sizes gleefully. So the recipe is: More time-invariant variables and more precisely expressed variables.

The experimental Design:

Personally, I think that my biggest challenge was to create a more efficient experimental setup. Initially the setup in use seemed like a great idea, but when it came to deciphering the results, is became plain tedious. Retrospectively, a simpler and yet more elegant group setting would have probably been less tedious, but I did not manage to develop such a group design (e.g. identifying two highly similar groups of municipalities with the only difference some form of alcohol supply). Although it is not certain that such a group design is easy to attain, as there are a lot of spillover effects and more likely the data just needs more control variables. The fact that the alcohol supply manifested itself at different time points in different municipalities instead of an exact treatment time point additionally complicated this setting.

There is also a fact considering the fixed effects approach. A lot of the alcohol supply stays fixed in the municipalities during the time period. This supply is omitted in the fixed effects regressions as it is constant and thus part of the fixed municipal effect. This leads to loss of statistical power and could be also combated with more variables and more accurate data.

7 Concluding remarks

Alcohol supply can be seen to have affected the following crime rates: Batteries, overall crime, and drunk and disorderly conduct.

The magnitude and significance of the effects varies according to the crime types and the types of alcohol supply. Some evidence is presented that the municipal differences also mattered in how the alcohol supply mediated related crime: The evidence is minor and most likely reflects the need for more time-variant control variables.

There is also some evidence that the types (not only the volume) of alcohol supply (Alko stores, licensed premises, places offering only medium-strength beer) have varying effects on different types of observed crime. I speculate that different types of alcohol supply might correlate differently with the observability of the crime.

As this legislative change is a unique setting in terms of how alcohol supply affected crime rates and social conditions, I suggest the following ways to improve this study: more accurate data and especially more time-variant variables to distinguish between what sort of socioeconomic and -demographic conditions favor crime. More efficient experimental setup is also proposed.

8 References

Angrist, J., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion.* New Jersey: Princeton University Press.

Benoit, K. (2011, March 17). *Linear Regression Models with Logarithmic Transformations.* Retrieved October 21, 2015, from Kenbenoit.net: http://www.kenbenoit.net/courses/ME104/logmodels2.pdf

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *Quarterly Journal of Economics*, 249-275.

Central Statistical Office. (1960-1971). *Official Statistics of Finland 23A, Criminality known to the Police.* Helsinki: Central Statistical Office.

Central Statistical Office. (1970). *Suomen Tilastollinen Vuosikirja 1969.* Helsinki: Central Statistical Office.

Central Statistical Office of Finland. (1972-1975). *Official statistics of Finland. 23 A, Criminality known to the police.* Helsinki: Central Statistical Office of Finland.

Crawley, M. J. (2007). *The R Book.* Chichester, UK: J Wiley & Son.

Drukker, D. (2009, August 25). *RE: st: Stata 11 Random Effects--Std. Errors.* Retrieved from Statalist: the Stata Forum / Statalist archive: http://www.stata.com/statalist/archive/2009-08/msg01228.html

Gould, W. (2011, August 22). *Use poisson rather than regress; tell a friend.* Retrieved from The Stata Blog: http://blog.stata.com/tag/poisson-regression/

Grace-Martin, K. (2010). *Regression Models for Count Data.* Retrieved from The Analysis Factor: http://www.theanalysisfactor.com/regression-models-for-count-data/

Grace-Martin, K. (2012, - -). *The Exposure Variable in Poisson Regression Models*. Retrieved from The Analysis Factor: http://www.theanalysisfactor.com/the-exposure-variable-in-poission-regression-models/

Gruenewald, P., Freisthler, B., Remer, L., LaScala, E., & Treno, A. (2006). Ecological models of alcohol outlets and violent assaults: crime potentials and geospatial analysis. *Addiction, 101*(5), 666-677.

Gyimah-Brempong, K. (2001). Alcohol availability and crime: Evidence from census tract data. *Southern Economic Journal*, *68*(1), 2-21.

Häikiö, M. (2007). *Alkon Historia.* Keuruu: Otava.

Lampela, O. (1964). Alkoholijuomien anniskelu. In O. A. ab, *Yearbook / Oy Alkoholiliike ab 1963* (pp. 3-10). Helsinki: Oy Alkoholiliike ab.

Lehti, M., & Sirén, R. (2008). 5 Väkivaltarikokset ja alkoholi. *OPTL:n tutkimuksia , 238*, 295-304.

Luukkonen, I. a. (2015, October 19). Contemporaries. (J. Luukkonen, Interviewer)

Maanmittaushallitus. (1960). *Tilastokartta Suomi 1960 1:2000000.* Retrieved 4 20, 2015, from Doria: http://www.doria.fi/handle/10024/79346

Maarten L. Buis. (2010). Interpretation of interactions in non-linear models. *The Stata Journal, 10*(2), 305-308.

Miles, J. (2014, April 3). *Quora.com*. Retrieved October 22, 2015, from Why do we use clustering in statistical analysis? Can you give an intuitive explanation or intuitive examples?: https://www.quora.com/Why-do-we-use-clustering-in-statistical-analysis-Can-you-give-an-intuitive-explanation-or-intuitive-examples

National Institute for Health and Welfare. (2014). *Alcoholic Beverage Consumption 2013.* Helsinki: National Institute for Health and Welfare.

National Land Survey, Esri Finland . (2015, 21 01). Municipal Division of Finland.

Northwestern University Feinberg School of Medicine. (1997, March 18). *Do your data violate Wilcoxon paired signed rank test assumptions?* . Retrieved from Prophet Statguide:

http://www.basic.northwestern.edu/statguidefiles/srank_paired_ass_viol.html#Tied%20values

Nussbaum, M., Elsadat, S., & Khago, A. (2008). Best Practices in Analyzing Count Data: Poisson Regression. In J. W. Osborne, *Best Practices in Quantitative Methods* (pp. 306-323). SAGE.

Osgood, D. W. (2000). Poisson-Based Regression Analysis of Aggregate Crime Rates. *Journal of Quantitative Criminology*, 21-43.

Oxford Dictionaries. (2015). *Oxford Dictionaries: British & World English*. Retrieved November 9, 2015, from Plan: http://www.oxforddictionaries.com/definition/english/plan?q=planned#plan__13

Oy Alko ab. (1970). *Yearbook / Oy Alkoholiliike ab.* Helsinki: Oy Alko ab.

Oy Alko ab. (1970-1975). *Yearbook / Oy Alko ab.* Helsinki: Oy Alko ab.

Oy Alkoholiliike ab. (1960-1969). *Yearbook / Oy Alkoholiliike ab.* Helsinki: Oy Alkoholiliike ab.

Peltonen, M., Kilpiö, K., & Kuusi , H. (2006). *Alkoholin vuosisata - suomalaisten alkoholiolojen vaiheita 1900-luvulla.* Helsinki: Suomalaisen kirjallisuuden seura.

Salusjärvi, M. (1990). Kun Suomi sai nopeusrajoitukset. In J. Lievonen, *Tekniikan tekijät* (pp. 10-16). Espoo: Valtion teknillinen tutkimuskeskus.

Scribner, R., MacKinnon, D., & Dwyer, J. (1995). The risk of assaultive violence and alcohol availability in Los Angeles County. *American Journal of Public Health, 85*(3), 335-340.

Smith, M. J. (2015, 5 7). *Statistical Analysis Handbook*. Retrieved from Poisson regression for count data: http://www.statsref.com/HTML/index.html?poisson_regression.html

Speer, P., Gorman, D., Labouvie, E., & Ontkush, M. (1998). Violent Crime and Alcohol Availability: Relationships in an Urban Community. *Journal of Public Health Policy, 19*(3), 303-318.

StataCorp. (2015). *Poisson Regression.* Retrieved May 15, 2015, from Stata 14 Base Reference Manual: http://www.stata.com/manuals13/rpoisson.pdf

Taylor, J. J. (2013, April 22). *Confusing Stats Terms Explained: Heteroscedasticity (Heteroskedasticity)*. Retrieved October 19, 2015, from Stats Make Me Cry: http://www.statsmakemecry.com/smmctheblog/confusing-stats-terms-explained-heteroscedasticity-heteroske.html

The Association of Finnish Local and Regional Authorities. (2015). *Kunnat.net.* Retrieved May 02, 2015, from Kuntajaon muutokset 1922-2014: http://www.kunnat.net/fi/haku/sivut//results.aspx?k=kuntajaon%20muutokset

Tilastollinen Päätoimisto. (1964). Yleinen väestölaskenta 1960: Syntymäpaikka, koulusivistys, siirtoväki ym. *Suomen virallinen tilasto VI C:103*.

Turunen, M. (2002). *Jos täytätte mun lasini : suomalaisen panimo- ja virvoitusjuomateollisuuden vuosisata* . Helsinki: Panimo- ja virvoitusjuomateollisuusliitto.

Wooldridge, J. M. (2013). *Introductory Econometrics - A Modern Approach 5th edition.* Mason, Usa: South-Western.

Zhu, L., Gorman, M., & Horel, S. (2004). Alcohol Outlet Density and Violence: A Geospatial Analysis. *Alcohol & Alcoholism*, *39*(4), 369-375.

9 Appendices

Appendix A: Georeferencing the municipal maps

The picture below is given to give a more clear view on the *georeferencing* process of the municipality maps of Finland from 1960 (Maanmittaushallitus, 1960) and 2015 (National Land Survey, Esri Finland , 2015). The white lines in the zoomed map represent the municipal division of 1960 while the black lines represent the municipal division of 2015. The maps were combined with the ArcMap component of the ArcGIS software (by Environmental systems Research Institute, version 10.2.2). Common points were pinpointed on the maps and the maps were fitted together using third order polynomials according to these points as the map from 1960 was quite distorted. After initial suffering, the end result was quite pleasing.

These maps are further used in modelling the municipalities and municipal clusters into a digitalized form, so that estimations for Alko store distances can be made. The municipal division map of 2015 is used as the base map for the coordinates.

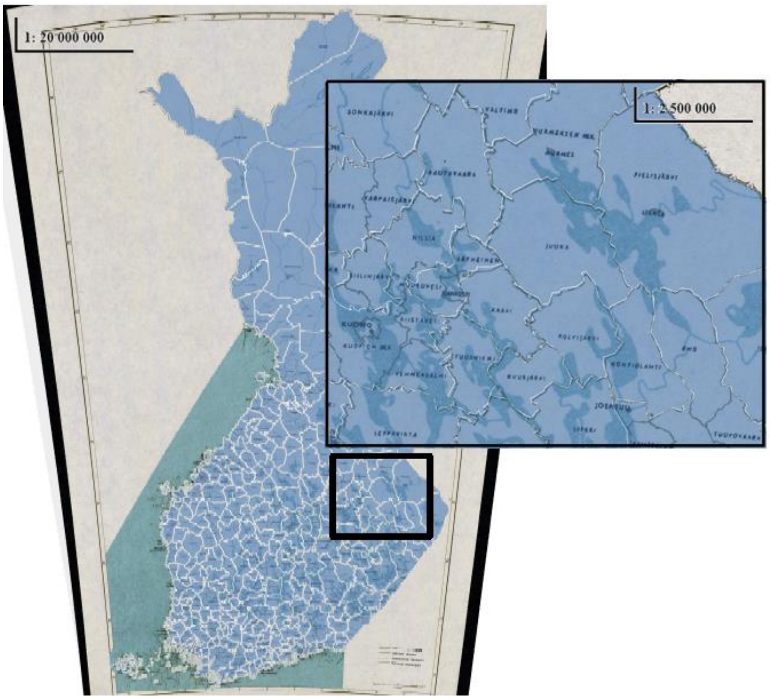

Image A.1. A zoom of the georeferencing process from ArcMap 10.3.1
(National Land Survey, Esri Finland , 2015) (Maanmittaushallitus, 1960)

Appendix B: Estimating the distances

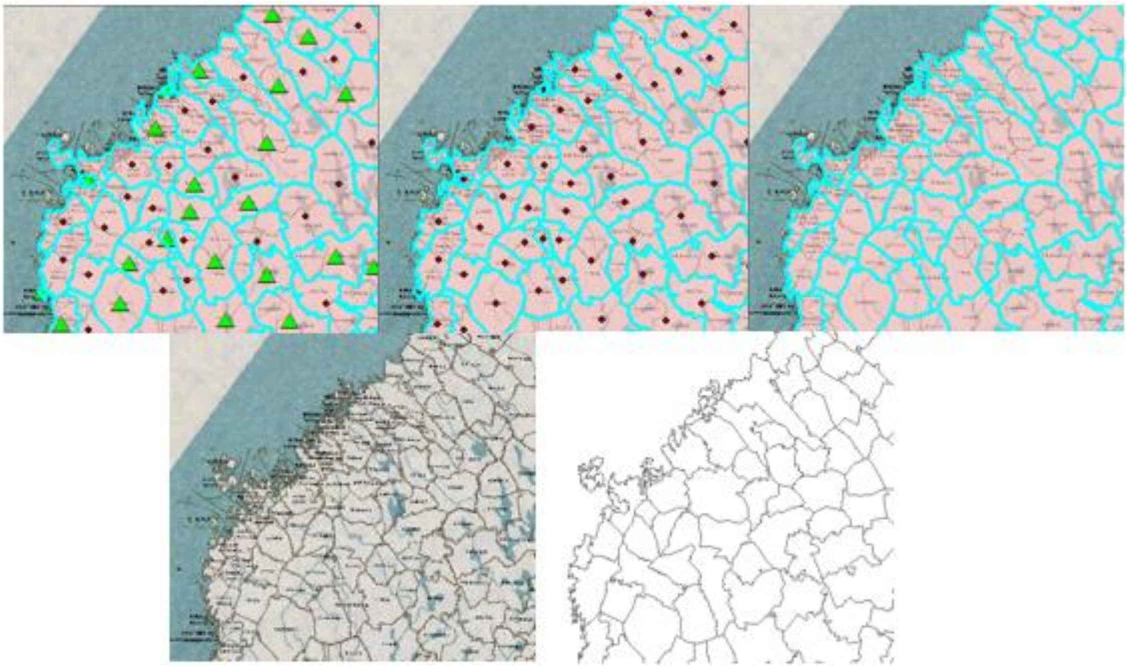

Image B.1. Screenshots of different stages of distance estimation from ArcMap 10.3.1

After the georeferencing process (previous appendix, also bottom two screenshots) was complete, the municipal lines had to be drawn according to the municipal division instated by myself earlier on. These newly drawn county shapes, that appear with pink background and teal borders, are known as polygons in geometry.

The municipal areas and distances are estimated with the help of these polygons. While the areas were quite easily determined, the distances needed additional attention. The distances are annually estimated from each municipality polygon's centroid (red dots on the two top left screenshots) to its nearest municipality polygon's centroid *with* an Alko store (green triangles on the top left screenshot).

There are two things that have to be considered with this approach:

I) In general the distances estimated are quite crude as they are Euclidean distances and ignore all sorts of population distributions.

II) If the given municipality has an Alko store it must be noted that the estimated distance is zero.

The first issue is something that I just have to live with[13]. The second issue however is combated with a work-around in form of circle radius: The municipal area estimated is taken as a circle, and the radius of the circle is calculated with simple geometry. This radius is then used to substitute the zero-estimate of the Alko distance after a subtle transformation:

$$Distance\ to\ Alko = 0{,}5 * \frac{approximated\ radius}{amount\ of\ \ Alko\ stores\ in\ the\ municipality}$$

There are also two reasons for this: The zero distance does not reflect the true distance in any way and, after considerable efforts, I did not figure out how to estimate the average distance from an irregular polygon’s centroid to its border.

There are of course problems with the radius-based approach: It ignores the original shape of the municipality and forces the shape of a circle, which can lead into problems regarding the municipality’s shape: The further it truly is from a circle, the poorer the estimated distance. However as the Alko stores were built near populations, and the populations tend to cluster, I believe that this circle radius approach is not that big of a sin and I believe that it is truer expression of the distance than zero-distance could ever be. Further efforts are also considered in tackling the population vs. geographical centroid issues: As the true distance to the Alko store lies probably nearer zero than the generated radius does, an additional arbitrary factor, which gets smaller with each Alko store in the municipal area, is used in reducing the radius-distances of these municipalities that housed Alko stores.

[13] I actually live with it quite well: we can presume that the population centroids can be expressed as the geographical centroids of the municipalities. Individual municipality-wise this could not be further away from truth, but after considering the sheer amount of municipalities, I’m sure that the individual differences will balance out. That is unless there are some major endogeneities related to the choices of municipal population centers (e.g. the coastline might be one).

Appendix C: Descriptive statistics for the diff-in-diff groups

Alko pre-1969 (n=73)

Population dens.	Mean	St.dev.	Max	Min	Median	Kurtosis	Skewness
1960	202.24	379.01	2868.50	3.51	67.45	34.41	5.21
1975	277.69	476.27	3182.49	3.29	72.98	19.49	3.83
Dist. to Alko							
1960	9.17	15.08	98.43	0.14	4.49	17.81	3.80
1975	4.99	4.04	17.98	0.11	3.79	1.16	1.16
Licens. dumm.							
1960	0.90	0.30	1.00	0.00	1.00	6.02	-2.80
1975	1.00	0.00	1.00	1.00	1.00	0.00	-

Alko post-1969 (n=57)

Population dens.	Mean	St.dev.	Max	Min	Median	Kurtosis	Skewness
1960	15.22	22.90	171.34	0.33	9.73	39.91	5.89
1975	19.86	63.91	486.94	0.36	8.06	53.55	7.22
Dist. to Alko							
1960	50.80	41.32	269.20	6.39	42.38	13.73	3.17
1975	11.73	6.89	42.48	2.14	10.16	6.98	2.30
Licens. dumm.							
1960	0.16	0.37	1.00	0.00	0.00	1.78	1.93
1975	0.98	0.13	1.00	0.00	1.00	57.00	-7.55

No Alko (n=139)

Population dens.	Mean	St.dev.	Max	Min	Median	Kurtosis	Skewness
1960	13.74	9.36	81.42	0.52	11.97	19.79	3.30
1975	12.66	13.43	129.08	0.45	9.01	41.39	5.35
Dist. to Alko							
1960	32.98	27.93	249.42	1.97	26.14	26.58	4.10
1975	25.96	14.44	112.57	1.97	24.02	9.67	2.21
Licens. dumm.							
1960	0.04	0.20	1.00	0.00	0.00	18.93	4.54
1975	0.85	0.36	1.00	0.00	1.00	1.91	-1.97

Table C.1. Descriptive statistics for the Diff-in-diff groups

Appendix D: Graphs for the diff-in-diff approach

This appendix focuses on the judging whether there are notable trend changes among the quasi-experimental groups after the Keskiolutlaki's entry into force and if Differences-in-differences based approach is a valid choice.

A quick recap on the group composition before we start with the overall crimes:

Treatment group: Alko post-69

Control group I: Alko pre-69

Control group II: No Alko

All the crime levels are annual averages for the above group averages and are announced as crimes per 1000 persons, unless stated otherwise.

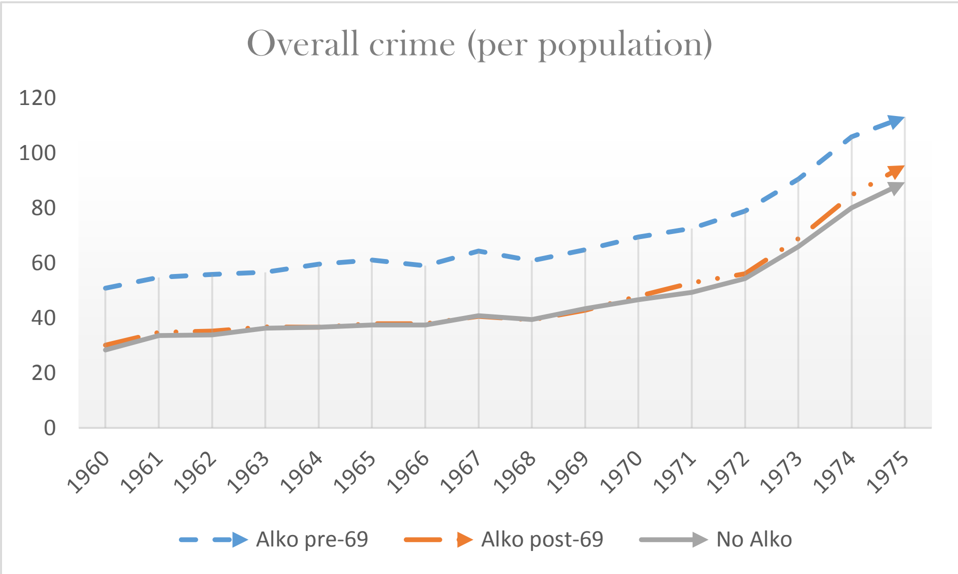

Figure D.1. Group trends: Overall crime per population

For overall crimes, there is a very minor convergence between the treatment group and the control group 1 after the legislative change and before that, the group trends seem to follow each other even better than I had anticipated. There is also evidence of fixed group level differences.

In the next two figures (G.2 and G.3), an interesting, yet galling phenomenon regarding the statistical bookkeeping can be sighted at the turn of 1964 and 1965. At 1965 something has changed regarding the statistical bookkeeping that concerns traffic crimes and honestly I do not have a clue what it is. However this is overridden by just limiting the observations for traffic crimes and other crimes to the years 1965-1975.

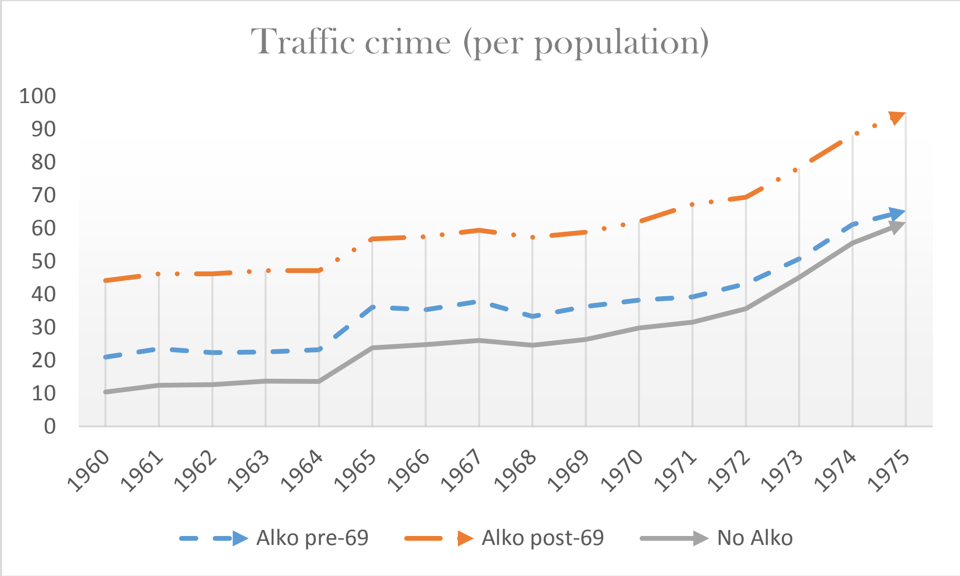

Figure D.2. Group trends: Traffic crime per population

First thing that is noticeable is that the base traffic crime rate in the treatment municipalities is a lot higher than in the other groups and the trends follow each other quite well until the 70s hits them with the varying attempts at speed limit legislation. There is very little change immediately after 1969. Most interesting convergence in the crime rates happens with the municipalities that had no Alko during the period as in the 70s they start to catch the municipalities that received Alko stores post-1969. To speculate these differences, I represent plausible reasons for this considering appendix C, which contains descriptive statistics for these groups.

There was little need for cars in the bigger municipalities (Alko-pre 69 group), as public transportation was pretty well established and thus the population adjusted car numbers most likely were smaller in the cities.

As for the treatment group and control group II, which had smaller population densities, there probably was more need for a car. What is peculiar, is that their initial levels of traffic crime are on the different ends of the scale. To speculate reasons behind this, I am relying on the word of contemporary people: Even in the 60s, the “really rural” municipalities were quite poor and very few families owned automobiles before the 70s. In these municipalities horses were a more likely source of transportation than two-stroke engines (Luukkonen, 2015).

So basically I am presuming that people in the treatment group had it better than the people in the control group II (the Alko stores could work as a proxy of this). Presumably if there were more cars in the treatment group, the surveillance of the traffic police would have been concentrated there. Thus the base traffic crime –level in the 'No Alko' -municipalities might have started to rise because of increasing car pool and/or the increased law enforcement by the traffic police.

In order to better estimate the reasons behind these crime levels or trend changes, some additional statistics especially regarding municipal car fleets would come handy. However fascinating, we are not interested in the trend changes of traffic crimes if they are most likely related with the changes in speed limit legislation, and not alcohol availability. Thus diff-in-diff –approach is not encouraged with traffic crimes.

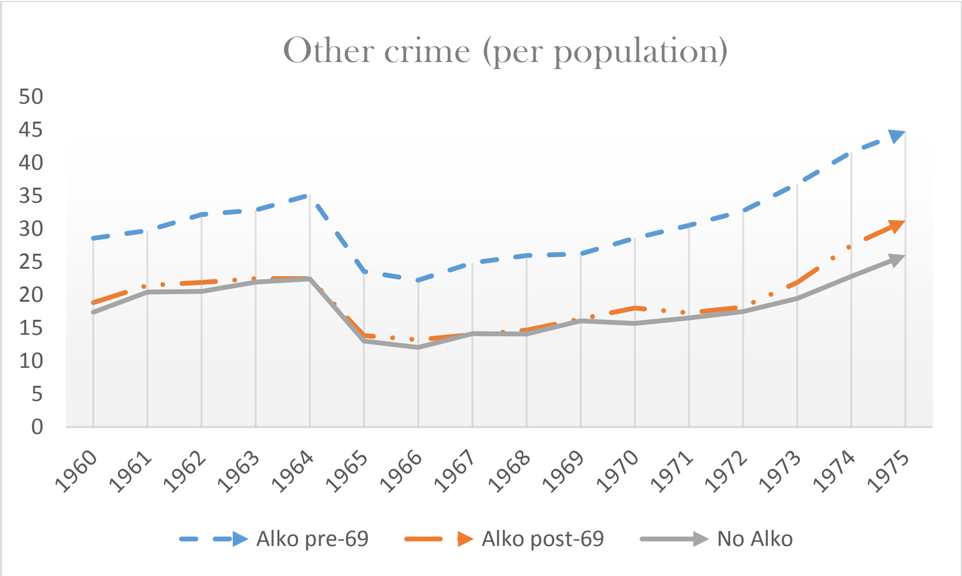

Figure D.3. Group trends: Other crime per population

As for the population adjusted other crimes, which were formed by deducing all the other crime types from overall crimes, the trends seem to follow each other quite slavishly. The treatment group's trend seems to diverge from the other rural municipalities after 1969, and as such diff-in-diff regression is chosen as the primary tool to analyze other crimes.

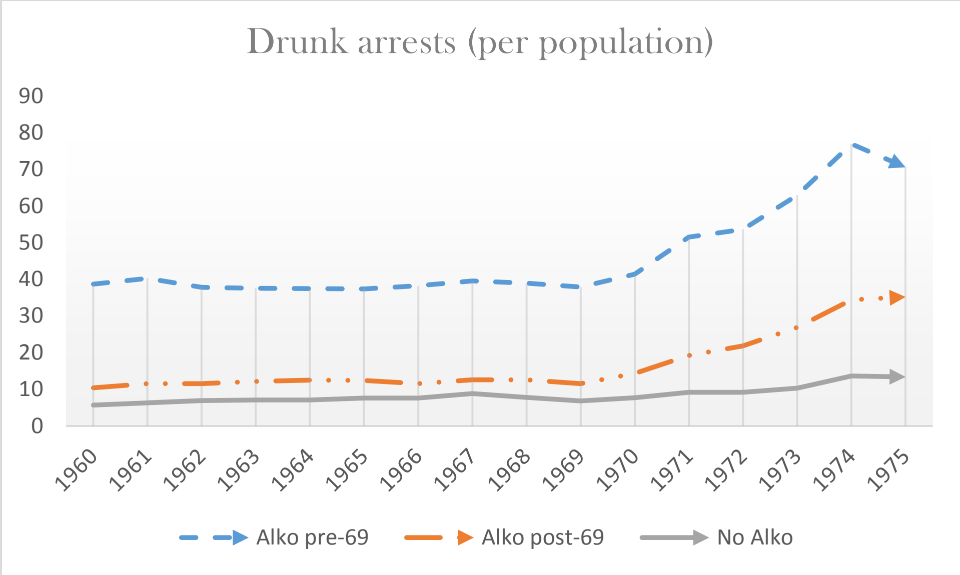

Figure D.4. Group trends: Drunk arrests per population

The figure containing population adjusted drunk arrests is the most beautiful in terms of the trend change. This captures the essence of the diff-in-diff setup: The trends were on a steady level before the legislative change and the treatment group takes a clear turn after 1969 (as does control I, but this was expected).

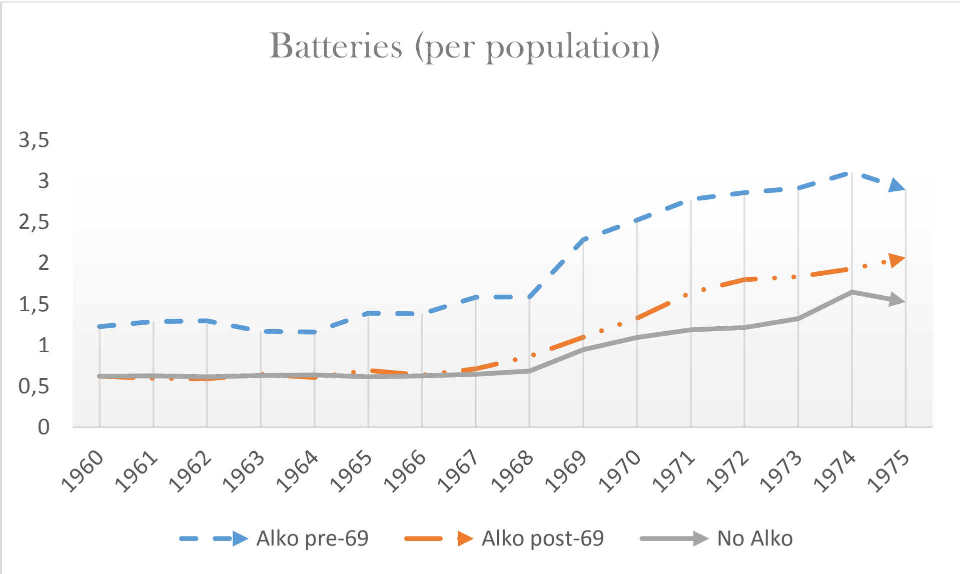

Figure D.5. Group trends: Batteries per population

For population adjusted batteries, there also seems to enough preliminary evidence for the diff-in-diff approach, even though the trend of the treatment municipalities seems to break off from the other rural municipalities before the legislative change.

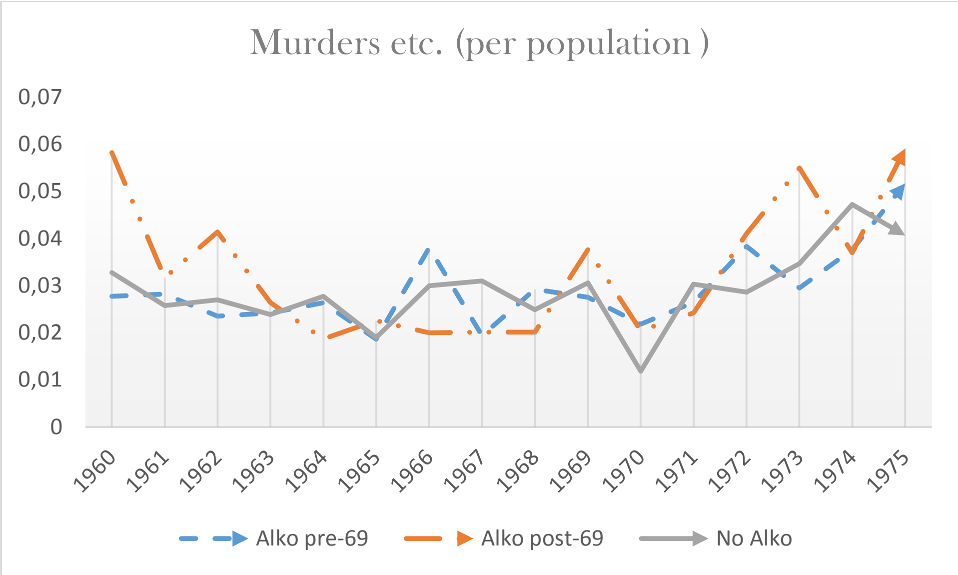

Figure D.6. Group trends: Fatal offences at attempts at these per population

While it might be a bit controversial to adjust small crime counts of zero with population, as small counts such as 1 attempt to kill another person might inflate the figures in population-wise small municipalities, I deem it necessary in order to see the trends (and it is not that big of a sin as the numbers are averages).

First impression: The population adjusted “murder counts” seem to follow a random walk process more than having clear trends and thus there is little evidence for diff-in-diff-approach here. Thus the group setup is neglected in the regression explaining the fatal offences. Speculation about the lack of trend changes (or trends in general) is presented below.

If one deems it necessary to kill someone, it does not look at the time and place. People with a tendency to commit killings seem to be “distributed” randomly across the population and this attempt to divide municipalities is not enough to identify plausible differences.

Appendix E: Robust standard errors in fighting heteroscedasticity

Instead of trying to prove that the variables are homoscedastic[14], I am going to apply robust standard errors, as it is highly likely that a time series data set this large would have heteroscedasticity present. In general it does not matter whether or not there is heteroscedasticity detected, now we just do apply the robust standard errors in an effort to be "conservative".

The problems related to heteroscedasticity can be defined with a simple linear regression (Wooldridge, 2013):

$$y_i = \beta_0 + \beta_1 x_i + u_i \qquad (1)$$

We assume that the first four Gauss-Markov assumptions hold. If the errors contain heteroscedasticity, the $Var(u_i|x_i) = \sigma_i^2$, where the subscript i on σ_i^2 indicates that the variance of the error depends upon the particular value of x_i. The fitted OLS estimator can be written as:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \qquad (2)$$

Without the homoscedasticity assumption and conditioning the values x_i in the sample, we can show that

$$Var(\hat{\beta}_i) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})\sigma_i^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \qquad (3)$$

When $\sigma_i^2 = \sigma^2$ for all i, the formula is reduced to the usual form $\frac{\sigma^2}{SST_x}$, where $SST_x = \sum_{i=1}^{n}(x_i - \overline{x})^2$. Thus equation (3) shows that the variance formula is no longer valid when heteroscedasticity is present. As the standard error of $\hat{\beta}_1$ is directly based on estimating its

[14] The variance of a variable is unequal across the range of values of a second variable that predicts it. This can have implications with standard errors and test-statistics being biased. (Taylor, 2013)

variance, a way to estimate equation (3) is needed. This can be taken care with the OLS residuals $\hat{u}_i$ from the initial regression of y on x, making (4) a valid estimator of $Var(\hat{\beta}_i)$:

$$\frac{\sum_{i=1}^{n}(x_i-\overline{x})\hat{u}_i^2}{SST_x} \qquad (4)$$

If equation (4) multiplied by the sample size n, it converges in probability to $\frac{E[(x_i-u_x)^2u_i^2]}{(\sigma_x^2)^2}$, which is the probability limit of n times (3). Thus the square root of (4) is called the heteroscedasticity-robust standard error for $\hat{\beta}_i$. This can be applied to multiple regression model as well.

Appendix F: Skewed distributions and logarithmic transformation

As the crime distributions are quite skewed, some adjustment can be seen as appropriate. In order to fight skewed distributions and more extreme values logarithmic transformation on certain variables is considered.

When the dependent variable is larger than zero, models using log(y) as the dependent variable often satisfy the assumptions related to linear regression better than the level of y would. Strictly positive values also often have distributions that are both skewed and heteroskedastic and taking the log can mitigate both problems. Taking logarithmic values also often narrows the range of values, which makes the analysis less sensitive to outliers. However logarithmic transformation is something not to be taken lightly as it can lead to more extreme values with variables of which values are between zero and one and it is nigh useless if the variable takes on zero or negative values. (Wooldridge, 2013)

Considering the tables containing the descriptive statistics presented in chapter 5, logarithmic transformation is applied with the following crimes: Overall crime, traffic crime, drunk arrests and other crimes. Logarithmic transformation is also applied with the distances to Alko and population densities. For the few zero valued observations in these, Stata simply treats as missing values and for the observations that take values between zero and one, they are also few in numbers and are just considered as a necessary evil. Ultimately I believe that these steps are necessary in order to transform the data more “regressible”. As for the other crime types, The Poisson regression takes care of batteries and ‘fatal offences and attempts at these’ as they contain too many zero valued observations that cannot be conveyed as logs.

An example on how logarithmic transformation affects the descriptive statistics with population adjusted overall crimes:

Crime/pop	Mean	Kurtosis	Skewness
1960	34.88	2.85	1.52
1975	97.15	7.18	1.92
Ln(crime/pop)			
1960	3.43	-0.28	0.23
1975	4.49	0.45	0.04

Table F.1. The effect of logarithmic transformation on overall crime/population

As we can see, the distribution after the logarithmic transformation is not normal, but it is way closer to one than distributions without the logarithmic transformation.

As for the regressions, the coefficients of the transformed variables can be interpreted as follows (Wooldridge, 2013) (Benoit, 2011):

$$\log(y) = \beta_0 + \beta_1 \log(x_1) + \beta_2 x_2$$

For the explanatory variables that underwent logarithmic transformation, the interpretation is simple: Expected percentage change in y when x increases by some percentage. One percent change in x_1 can be interpreted as β_1 % change in y holding x_2 fixed. Hence this is an elasticity.

For the variables that did not undergo transformation, the interpretation requires a bit algebra:

$$\%\Delta y = 100 * [\exp(\beta_2 \Delta x_2) - 1],$$

This implies that one unit change in x_2 can be interpreted as $100 * [\exp(\beta_2) - 1]\%$ change in y. To account for more than one unit increase in x_2 , we must include c in the component. Each one unit increase in x_2 multiplies the expected value of y by $e^{c\hat{\beta}}$

We can also directly take β_2 as an approximation of the percentage change if the values of the coefficients are small ($e^{\hat{\beta}} \approx 1 + \hat{\beta}$), which means that $100 * (\hat{\beta} - 1)$ is the approximated expected percentage change in y for one unit increase in x.

Appendix G: Inconsistent standard errors caused by correlation

The correlation between observations is a problem that messes up the standard errors and thus trustworthy interpretation of the regression results.

Three factors make correlation between observations highly important especially for differences-in-differences based estimation: I) DID usually relies on fairly long time series, II) most commonly used dependent variables in DID estimation are typically highly positively correlated and III) the treatment variable usually changes itself very little within a unit of observation over time. These three factors can reinforce themselves in a fashion, such that the estimated standard errors could seriously understate the actual standard deviations. (Bertrand;Duflo;& Mullainathan, 2004)

In essence, the problem related to correlation can be seen summed by 'h*ow much information the model presumes it has*' versus '*how much information the model truly has due to the non-independent observations*'. If correlation between observations is neglected, this leads to interpreting all the observations as independent sources of information, while truly there are less independent observations than the model assumes. (Miles, 2014)

To further distinguish between different types of correlation, the correlation between observations applies for both correlation in time (autocorrelation, the outcome of observation $y_{c,i,t}$ is correlated with $y_{c,i,t-1}$) or intra-cluster correlation (the outcome $y_{1,1,t}$ is correlated with the outcome of $y_{1,2,t}$ but not $y_{2,1,t}$. The subscripts denominate i for individual, c for cluster and t for point of time). The intra-cluster correlation can be seen as a presence of an observed shock that affects all the observations in one group while not the others, while serial correlation is regarded when the errors of observations of certain units are somehow temporally correlated with each other. Serial correlation in the errors may be due to measurement errors, misspecifications in the model, or due to omitted variables.

The fixed effects takes care of the common random effects that are time-invariant on a municipal level. That being said, should time-variant random shocks affect some municipalities, it is highly unlikely that I will manage to identify and cluster these municipalities in a fashion that takes care of such shocks. Based on the historical literature evidence I have absorbed and the word of contemporary people, it is also highly unlikely that I have missed such a shock that affected crime in certain municipalities. The quasi-experimental group setup introduced in chapter 5.4.1 could be a plausible source for intra-cluster correlation, although I seriously doubt that there is a mechanism for this: The clusters are geographically quite divided and I cannot imagine how some random unobserved shock would affect only one cluster and not the others (¡Alko stores are not unobserved!). If the observations would consist of individual human beings committing crimes or not committing them, instead of municipalities (which in a sense are aggregated individuals), I would be more concerned as the individuals living in the same municipality could somehow be correlated intra-class-wise. My conclusion with the intra-class correlation is as follows: Finland in the turn of 70s was a quite homogenous place and with this, my identifying assumption is that the observed crime rates in different municipalities are independent observations (or independent enough) from each other. Thus the plausibility of unobserved shocks is dismissed in the scope of this study.

Instead a more severe issue is serial correlation within the municipal level observations: Are the errors in municipal variables in different municipalities correlated in time? Temporal independence within municipalities is something that is hard to swallow with sixteen years of observations. Unfortunately there is a cure for this: Clustered standard errors as a way to fight serial correlation. Without going that much into the specifics, the clustered standard errors adjust the number of observations according to detected correlation within the unit that contains the observations (Miles, 2014). This works for temporally correlated observations as well: Stata does this clustering with '*vce(robust)*'-command if the municipal id has been specified as the identifying variable in the time series. This clusters the individual observations and thus this vce-estimator is robust individual-wise instead of observation-wise and is consistent in the presence of heteroscedasticity and autocorrelation (Drukker, 2009).

Appendix H: Interpreting the regression coefficients

In order to give a clear interpretation of how the findings are analyzed, I will present an overview on how the coefficients of the regression should be interpreted pre- and post- 1969. Basically the dummy variables indicating different groups of the quasi-experiment are interacted with different variables indicating alcohol supply. The following variables are needed for each group, while this example focuses on the treatment group:

Explained variable:

Crime level in municipality X

Explanatory variables that measure the effect of alcohol supply pre-legislative change:

Treatment (omitted due to fixed effects model)

β_3 $*$Distance to Alko

β_4 $*$License dummy

β_5 $*$(Treatment group*Distance to Alko)

β_6 $*$(Treatment group*License dummy)

The variables expressing alcohol supply are used in the overall regressions in measuring how the alcohol supply affects crime. If there are differences on how the quasi-experimental groups reacted on the alcohol supply, the interaction terms should capture these differences (β_5 and $\beta_6$respectively). But wait, there is more:

Explanatory variables that measure the effect of alcohol supply post- legislative change:

The variables above, and:

β_7 $*$Keskiolutlaki

β_8 $*$(Distance to Alko *Keskiolutlaki)

β_9 $*$(License dummy*Keskiolutlaki)

β_{10} $*$(Treatment group*Keskiolutlaki)

β_{11} $*$(Treatment group*Distance to Alko post Keskiolutlaki)

β_{12} $*$(Treatment group*License dummy post Keskiolutlaki)

After the legislative change, a dummy variable expressing the Keskiolutlaki's entry into force is thrown in the fray. This dummy variable by its own expresses the effect of the medium-strength beer joints (β_7) on crime. This dummy is also used to create interaction variables with the other forms of alcohol supply that had existed before the legislative change (β_8 and β_9). These interaction variables are used in order to measure how the presumed roles of different alcohol supply (Alko stores and licensed premises) changed after the legislative change. The Keskiolutlaki dummy is also used with the quasi-experimental groups in order to check whether their response to medium-strength beer was different.

To ma, quasi-experimental group interaction variables are also used in determining whether the groups reacted differently to the Alko stores and licensed premises after the legislative change (β_{11} and β_{12}). To put it other words; if the groups indeed reacted differently to alcohol supply pre-1969, then this effect most likely would have changed with the availability of medium-strength beer post-1969. Thus the coefficients are interacted with the Keskiolutlaki dummy after the legislative change. Yes, this might sound confusing.

In order to interpret these results, all the coefficients should be added together. For example in order to determine how the alcohol supply affected a treatment municipality after the legislative change that had no licensed premises (dummy value zero) and a seven kilometer distance to Alko, we take:

$$\boldsymbol{\beta_3 * 7 + \beta_4 * 0 + \beta_5(1 * 7) + \beta_6(1 * 0)} + \beta_7 * 1 + \beta_8(7 * 1) + \beta_9(0 * 1) + \beta_{10}(1 * 1) + \beta_{11}(1 * 1 * 7) + \beta_{12}(1 * 0 * 1)$$

The bolded part of the equation represents the effect of alcohol supply before the legislative change. In a nutshell, this kind of an equation should be able to distinguish the magnitude and the significance of the following things:

Pre-1969

I) The effect of Alko stores and licensed premises on municipal crime.

II) The differences of the quasi-experimental groups' reaction to alcohol supply.

Post-1969

I) The effect of medium-strength beer on alcohol induced crime (the dummy itself) and the effect of Keskiolutlaki on the crime induced by the Alko stores and licensed premises (interaction variables with the variables indicating these).

II) The differences of the quasi-experimental groups' reaction to medium-strength beer and the change in the quasi-experimental groups' reaction to alcohol supply that existed pre-legislative change.

Appendix i: Regression results

OVERALL CRIME

```
Fixed-effects (within) regression               Number of obs     =      4,304
Group variable: Munid                           Number of groups  =        269

R-sq:                                           Obs per group:
     within  = 0.6581                                         min =         16
     between = 0.3755                                         avg =       16.0
     overall = 0.0075                                         max =         16

                                                F(11,268)         =     171.84
corr(u_i, Xb)  = -0.8074                        Prob > F          =     0.0000

                                     (Std. Err. adjusted for 269 clusters in Munid)
```

LNcrimeP1000	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
LNpopulationdensity	-.3901895	.0816023	-4.78	0.000	-.5508525	-.2295264
Time	.0372338	.0027852	13.37	0.000	.0317502	.0427174
Speed68	-.099679	.0129059	-7.72	0.000	-.1250889	-.0742692
Speed72	-.0466967	.0196498	-2.38	0.018	-.0853843	-.0080091
Speed73	.0905621	.0240595	3.76	0.000	.0431925	.1379318
Speed74	.2697095	.0270967	9.95	0.000	.2163599	.323059
Licensedummy	.0545173	.0229664	2.37	0.018	.0092998	.0997349
LNdistancetoAlko	-.0393711	.0201645	-1.95	0.052	-.0790721	.0003299
Keskiolutlaki	.1294129	.0532386	2.43	0.016	.0245938	.2342321
LNDALKOPOST	-.0037841	.0126339	-0.30	0.765	-.0286585	.0210903
LicensePOST	-.0618577	.0398857	-1.55	0.122	-.1403868	.0166715
_cons	4.679875	.234418	19.96	0.000	4.21834	5.14141
sigma_u	.78442631					
sigma_e	.21889019					
rho	.92775904	(fraction of variance due to u_i)				

Table i.1. Overall crime without the group setup

```
Fixed-effects (within) regression               Number of obs     =      4,304
Group variable: Munid                           Number of groups  =        269

R-sq:                                           Obs per group:
     within  = 0.6611                                         min =         16
     between = 0.4102                                         avg =       16.0
     overall = 0.0404                                         max =         16

                                                F(20,268)         =     109.86
corr(u_i, Xb)  = -0.8901                        Prob > F          =     0.0000

                                (Std. Err. adjusted for 269 clusters in Munid)
```

LNcrimeP1000	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Time	.0378863	.0027972	13.54	0.000	.0323791	.0433935
Speed68	-.0976185	.0130043	-7.51	0.000	-.123222	-.072015
Speed72	-.0447795	.0196314	-2.28	0.023	-.0834309	-.0061281
Speed73	.0919637	.0241887	3.80	0.000	.0443397	.1395877
Speed74	.2699609	.0272784	9.90	0.000	.2162536	.3236681
LNpopulationdensity	-.4043454	.0856876	-4.72	0.000	-.573052	-.2356389
Licensedummy	.0366029	.0330746	1.11	0.269	-.0285162	.1017219
LNdistancetoAlko	.1177221	.0935006	1.26	0.209	-.066367	.3018113
Keskiolutlaki	.121165	.1055113	1.15	0.252	-.0865715	.3289015
LNDALKOPOST	-.0004318	.0324245	-0.01	0.989	-.0642709	.0634072
LicensePOST	-.0407914	.04688	-0.87	0.385	-.1330913	.0515085
Ctrl1	0	(omitted)				
Treat	0	(omitted)				
Ctrl2	0	(omitted)				
Ctrl1XLicence	-.02225	.1192147	-0.19	0.852	-.2569664	.2124664
TreatXLicence	.0049563	.0447848	0.11	0.912	-.0832185	.0931312
Ctrl2XLicence	0	(omitted)				
Ctrl1XLNAlko	-.2259329	.1019881	-2.22	0.028	-.4267327	-.0251331
TreatXLNAlko	-.0933831	.1034761	-0.90	0.368	-.2971124	.1103463
Ctrl2XLNAlko	0	(omitted)				
Ctrl1XKeskiolutlaki	-.0498414	.1077748	-0.46	0.644	-.2620344	.1623515
TreatXKeskiolutlaki	.3275848	.2029045	1.61	0.108	-.0719048	.7270745
Ctrl2XKeskiolutlaki	0	(omitted)				
Ctrl1XAlkoPOST	.0125152	.0378606	0.33	0.741	-.0620267	.0870572
TreatXAlkoPOST	-.0706254	.0524679	-1.35	0.179	-.1739272	.0326764
Ctrl2XAlkoPOST	0	(omitted)				
Ctrl1XLicensePOST	0	(omitted)				
TreatXLicensePOST	-.0845188	.09383	-0.90	0.369	-.2692564	.1002188
Ctrl2XLicensePOST	0	(omitted)				
_cons	4.44386	.2959437	15.02	0.000	3.861189	5.02653
sigma_u	.99574855					
sigma_e	.21816172					
rho	.95419681	(fraction of variance due to u_i)				

Table i.2. Overall crime with the group setup

DRUNK ARRESTS

```
Fixed-effects (within) regression               Number of obs     =      4,295
Group variable: Munid                           Number of groups  =        269

R-sq:                                           Obs per group:
     within  = 0.2841                                         min =         13
     between = 0.2754                                         avg =       16.0
     overall = 0.1309                                         max =         16

                                                F(7,268)          =      75.35
corr(u_i, Xb)  = -0.8046                        Prob > F          =     0.0000

                                 (Std. Err. adjusted for 269 clusters in Munid)
```

LNdrunkarrestP1000	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
LNpopulationdensity	-.7454859	.1761963	-4.23	0.000	-1.092391	-.3985809
Time	.0588701	.0037059	15.89	0.000	.0515737	.0661665
Licensedummy	-.0797262	.0430432	-1.85	0.065	-.164472	.0050195
LNdistancetoAlko	-.2526898	.0389766	-6.48	0.000	-.3294291	-.1759504
Keskiolutlaki	-.2441758	.1310895	-1.86	0.064	-.5022721	.0139205
LNDALKOPOST	-.1194233	.029981	-3.98	0.000	-.1784516	-.060395
LicensePOST	.3605819	.0954913	3.78	0.000	.1725734	.5485905
_cons	4.944266	.5094583	9.70	0.000	3.941216	5.947315
sigma_u	1.4913954					
sigma_e	.42627414					
rho	.92447552	(fraction of variance due to u_i)				

Table i.3. Drunk arrests without the group setup

```
Fixed-effects (within) regression               Number of obs     =      4,295
Group variable: Munid                           Number of groups  =        269

R-sq:                                           Obs per group:
     within  = 0.2975                                         min =         13
     between = 0.3453                                         avg =       16.0
     overall = 0.1761                                         max =         16

                                                F(16,268)         =      43.81
corr(u_i, Xb)  = -0.8395                        Prob > F          =     0.0000
                             (Std. Err. adjusted for 269 clusters in Munid)
```

LNdrunkarrestP1000	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Time	.0599916	.0037512	15.99	0.000	.0526061	.0673771
LNpopulationdensity	-.7194436	.1643337	-4.38	0.000	-1.042993	-.3958944
Licensedummy	-.1359544	.0674369	-2.02	0.045	-.2687278	-.003181
LNdistancetoAlko	-.0404486	.1707021	-0.24	0.813	-.3765364	.2956392
Keskiolutlaki	.1086698	.2885343	0.38	0.707	-.4594125	.676752
LNDALKOPOST	-.2386655	.0895435	-2.67	0.008	-.4149637	-.0623674
LicensePOST	.4539103	.1153289	3.94	0.000	.2268444	.6809763
Ctrl1	0	(omitted)				
Treat	0	(omitted)				
Ctrl2	0	(omitted)				
Ctrl1XLicence	.33026	.1893185	1.74	0.082	-.0424808	.7030008
TreatXLicence	.0328685	.0790523	0.42	0.678	-.1227741	.1885111
Ctrl2XLicence	0	(omitted)				
Ctrl1XLNAlko	-.106131	.1798683	-0.59	0.556	-.4602656	.2480036
TreatXLNAlko	-.162198	.1824857	-0.89	0.375	-.5214858	.1970899
Ctrl2XLNAlko	0	(omitted)				
Ctrl1XKeskiolutlaki	-.6431321	.3001561	-2.14	0.033	-1.234096	-.0521681
TreatXKeskiolutlaki	.4026596	.3683001	1.09	0.275	-.3224699	1.127789
Ctrl2XKeskiolutlaki	0	(omitted)				
Ctrl1XAlkoPOST	.25721	.0938497	2.74	0.007	.0724335	.4419865
TreatXAlkoPOST	-.0077973	.1088342	-0.07	0.943	-.222076	.2064814
Ctrl2XAlkoPOST	0	(omitted)				
Ctrl1XLicensePOST	0	(omitted)				
TreatXLicensePOST	-.4471297	.1378267	-3.24	0.001	-.7184904	-.175769
Ctrl2XLicensePOST	0	(omitted)				
_cons	4.36616	.5476807	7.97	0.000	3.287856	5.444464
sigma_u	1.5806167					
sigma_e	.42274981					
rho	.93324133	(fraction of variance due to u_i)				

Table i.4. Drunk arrests with the group setup

OTHER CRIME

```
Fixed-effects (within) regression               Number of obs     =      2,959
Group variable: Munid                           Number of groups  =        269

R-sq:                                           Obs per group:
     within  = 0.3435                                         min =         11
     between = 0.3444                                         avg =       11.0
     overall = 0.2106                                         max =         11

                                                F(7,268)          =      67.75
corr(u_i, Xb)  = 0.1038                         Prob > F          =     0.0000

                                  (Std. Err. adjusted for 269 clusters in Munid)
```

LNothercrimeP1000	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
LNpopulationdensity	-.004733	.1529516	-0.03	0.975	-.3058725	.2964066
Time	.0743268	.0046624	15.94	0.000	.0651473	.0835063
Licensedummy	-.0148052	.0354284	-0.42	0.676	-.0845586	.0549482
LNdistancetoAlko	-.0434554	.0288059	-1.51	0.133	-.1001699	.0132592
Keskiolutlaki	-.1066528	.0742606	-1.44	0.152	-.252861	.0395555
LNDALKOPOST	-.0031502	.0188358	-0.17	0.867	-.040235	.0339347
LicensePOST	.0362672	.053371	0.68	0.497	-.0688125	.1413469
_cons	2.263684	.4283442	5.28	0.000	1.420336	3.107031
sigma_u	.40146263					
sigma_e	.30178332					
rho	.63895014	(fraction of variance due to u_i)				

Table i.5. Other crime without the group setup

```
Fixed-effects (within) regression               Number of obs     =      2,959
Group variable: Munid                           Number of groups  =        269

R-sq:                                           Obs per group:
     within  = 0.3461                                         min =         11
     between = 0.1478                                         avg =       11.0
     overall = 0.0076                                         max =         11

                                                F(16,268)         =      36.37
corr(u_i, Xb)  = -0.4394                        Prob > F          =     0.0000

                                (Std. Err. adjusted for 269 clusters in Munid)
```

LNothercrimeP1000	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Time	.0749436	.0046747	16.03	0.000	.0657397	.0841475
LNpopulationdensity	-.0128119	.1530459	-0.08	0.933	-.3141372	.2885134
Licensedummy	-.024761	.05184	-0.48	0.633	-.1268265	.0773045
LNdistancetoAlko	.0985563	.1916949	0.51	0.608	-.2788632	.4759758
Keskiolutlaki	-.2624772	.1585717	-1.66	0.099	-.574682	.0497275
LNDALKOPOST	.0501601	.0518046	0.97	0.334	-.0518357	.1521559
LicensePOST	.0425889	.0644833	0.66	0.510	-.0843695	.1695472
Ctrl1	0	(omitted)				
Treat	0	(omitted)				
Ctrl2	0	(omitted)				
Ctrl1XLicence	.0277824	.1215066	0.23	0.819	-.2114464	.2670113
TreatXLicence	.0080494	.0724478	0.11	0.912	-.1345899	.1506887
Ctrl2XLicence	0	(omitted)				
Ctrl1XLNAlko	-.0973158	.2007629	-0.48	0.628	-.4925889	.2979573
TreatXLNAlko	-.1233312	.1994786	-0.62	0.537	-.5160756	.2694132
Ctrl2XLNAlko	0	(omitted)				
Ctrl1XKeskiolutlaki	.1875414	.1664457	1.13	0.261	-.140166	.5152488
TreatXKeskiolutlaki	.2880423	.2643012	1.09	0.277	-.2323285	.8084131
Ctrl2XKeskiolutlaki	0	(omitted)				
Ctrl1XAlkoPOST	-.0948626	.056881	-1.67	0.097	-.2068531	.0171278
TreatXAlkoPOST	-.0836945	.073282	-1.14	0.254	-.2279761	.0605871
Ctrl2XAlkoPOST	0	(omitted)				
Ctrl1XLicensePOST	0	(omitted)				
TreatXLicensePOST	-.026213	.1171254	-0.22	0.823	-.256816	.20439
Ctrl2XLicensePOST	0	(omitted)				
_cons	2.012808	.5427185	3.71	0.000	.9442737	3.081342
sigma_u	.52461478					
sigma_e	.30167776					
rho	.7514965	(fraction of variance due to u_i)				

Table i.6. Other crime with the group setup

TRAFFIC CRIME

```
Fixed-effects (within) regression               Number of obs     =      2,959
Group variable: Munid                           Number of groups  =        269

R-sq:                                           Obs per group:
     within  = 0.6178                                         min =         11
     between = 0.2353                                         avg =       11.0
     overall = 0.0042                                         max =         11

                                                F(11,268)         =     142.26
corr(u_i, Xb)  = -0.8085                        Prob > F          =     0.0000

                                (Std. Err. adjusted for 269 clusters in Munid)
```

LNtrafficcrimeP1000	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
LNpopulationdensity	-.4059022	.114409	-3.55	0.000	-.6311569	-.1806475
Time	.0522891	.0064645	8.09	0.000	.0395614	.0650168
Speed68	-.15805	.0188471	-8.39	0.000	-.1951573	-.1209428
Speed72	-.10641	.0290124	-3.67	0.000	-.1635313	-.0492887
Speed73	.0441525	.0367751	1.20	0.231	-.0282523	.1165573
Speed74	.2149866	.0459927	4.67	0.000	.1244337	.3055394
Licensedummy	.0247319	.0259734	0.95	0.342	-.026406	.0758698
LNdistancetoAlko	-.0520907	.0207065	-2.52	0.012	-.0928588	-.0113227
Keskiolutlaki	-.0439266	.0550685	-0.80	0.426	-.1523485	.0644954
LNDALKOPOST	.0482739	.0139108	3.47	0.001	.0208855	.0756622
LicensePOST	-.0300241	.0364896	-0.82	0.411	-.1018668	.0418187
_cons	4.155712	.3312683	12.54	0.000	3.503492	4.807931
sigma_u	.83550138					
sigma_e	.24276386					
rho	.92214725	(fraction of variance due to u_i)				

Table i.7. Traffic crime without the group setup

```
Fixed-effects (within) regression               Number of obs     =      2,959
Group variable: Munid                           Number of groups  =        269

R-sq:                                           Obs per group:
     within  = 0.6221                                         min =         11
     between = 0.2379                                         avg =       11.0
     overall = 0.0226                                         max =         11

                                                F(20,268)         =      89.25
corr(u_i, Xb)  = -0.8912                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 269 clusters in Munid)
```

LNtrafficcrimeP1000	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Time	.0522649	.0065113	8.03	0.000	.0394452	.0650847
Speed68	-.1594091	.0188107	-8.47	0.000	-.1964447	-.1223736
Speed72	-.1068259	.0290431	-3.68	0.000	-.1640075	-.0496443
Speed73	.0439884	.0367952	1.20	0.233	-.028456	.1164328
Speed74	.2148792	.0460924	4.66	0.000	.1241299	.3056284
LNpopulationdensity	-.4331193	.1125429	-3.85	0.000	-.6546999	-.2115387
Licensedummy	.0450644	.036598	1.23	0.219	-.0269918	.1171205
LNdistancetoAlko	.1312148	.0872402	1.50	0.134	-.0405486	.3029781
Keskiolutlaki	.0370981	.100256	0.37	0.712	-.1602913	.2344876
LNDALKOPOST	.0234151	.0302846	0.77	0.440	-.036211	.0830412
LicensePOST	-.0437608	.0449918	-0.97	0.332	-.1323431	.0448216
Ctrl1	0	(omitted)				
Treat	0	(omitted)				
Ctrl2	0	(omitted)				
Ctrl1XLicence	-.0486717	.1458988	-0.33	0.739	-.3359253	.2385819
TreatXLicence	-.0297483	.0503311	-0.59	0.555	-.128843	.0693463
Ctrl2XLicence	0	(omitted)				
Ctrl1XLNAlko	-.2411155	.1004333	-2.40	0.017	-.4388542	-.0433768
TreatXLNAlko	-.1293797	.1009895	-1.28	0.201	-.3282134	.069454
Ctrl2XLNAlko	0	(omitted)				
Ctrl1XKeskiolutlaki	-.1382411	.1015565	-1.36	0.175	-.3381912	.061709
TreatXKeskiolutlaki	.1546746	.2358421	0.66	0.512	-.3096643	.6190135
Ctrl2XKeskiolutlaki	0	(omitted)				
Ctrl1XAlkoPOST	.0851391	.0378486	2.25	0.025	.0106207	.1596575
TreatXAlkoPOST	-.0230573	.059514	-0.39	0.699	-.1402318	.0941171
Ctrl2XAlkoPOST	0	(omitted)				
Ctrl1XLicensePOST	0	(omitted)				
TreatXLicensePOST	-.0151235	.0942271	-0.16	0.873	-.2006431	.170396
Ctrl2XLicensePOST	0	(omitted)				
_cons	3.915144	.3501004	11.18	0.000	3.225848	4.604441
sigma_u	1.0731473					
sigma_e	.24179743					
rho	.95168543	(fraction of variance due to u_i)				

Table i.8. Traffic crime with the group setup

BATTERIES

```
Iteration 0:   log pseudolikelihood = -21726.074
Iteration 1:   log pseudolikelihood = -13787.212
Iteration 2:   log pseudolikelihood = -13785.675
Iteration 3:   log pseudolikelihood = -13785.675

Conditional fixed-effects Poisson regression    Number of obs      =      4,304
Group variable: Munid                           Number of groups   =        269

                                                Obs per group:
                                                             min =         16
                                                             avg =       16.0
                                                             max =         16

                                                Wald chi2(7)       =     887.87
Log pseudolikelihood  = -13785.675              Prob > chi2        =     0.0000

                                  (Std. Err. adjusted for clustering on Munid)
```

Battery	IRR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Time	1.032699	.0058449	5.68	0.000	1.021307	1.044219
Populationdensity	.9993615	.0001241	-5.14	0.000	.9991184	.9996047
DistancetoAlko	.9948936	.0010629	-4.79	0.000	.9928125	.996979
Licensedummy	1.066303	.0607145	1.13	0.260	.9537042	1.192195
Keskiolutlaki	1.376803	.0916012	4.81	0.000	1.208481	1.56857
DistancePOST	.9984512	.00216	-0.72	0.474	.9942266	1.002694
LicensePOST	1.124761	.0773018	1.71	0.087	.9830139	1.286949
ln(Population)	1	(exposure)				

Table i.9. Batteries without the group setup

```
Iteration 0:   log pseudolikelihood = -21726.074
Iteration 1:   log pseudolikelihood = -13787.212
Iteration 2:   log pseudolikelihood = -13785.675
Iteration 3:   log pseudolikelihood = -13785.675

Conditional fixed-effects Poisson regression    Number of obs     =      4,304
Group variable: Munid                           Number of groups  =        269

                                                Obs per group:
                                                              min =         16
                                                              avg =       16.0
                                                              max =         16

                                                Wald chi2(7)      =     887.87
Log pseudolikelihood  = -13785.675              Prob > chi2       =     0.0000

                                   (Std. Err. adjusted for clustering on Munid)
------------------------------------------------------------------------------
                  |               Robust
          Battery |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------+-----------------------------------------------------------
             Time |   .0321761   .0056598     5.68   0.000      .021083    .0432691
Populationdensity |  -.0006387   .0001241    -5.14   0.000     -.000882   -.0003954
   DistancetoAlko |  -.0051195   .0010684    -4.79   0.000    -.0072135   -.0030255
     Licensedummy |   .0641973   .0569393     1.13   0.260    -.0474017    .1757962
     Keskiolutlaki |   .3197643   .0665318     4.81   0.000     .1893643    .4501642
     DistancePOST |    -.00155   .0021634    -0.72   0.474    -.0057902    .0026901
      LicensePOST |   .1175709   .0687273     1.71   0.087    -.0171321    .2522739
   ln(Population) |          1  (exposure)
------------------------------------------------------------------------------
```

Table i.10. Batteries without the group setup (IRR)

```
note: you are responsible for interpretation of non-count dep. variable
note: Ctrl2 omitted because of collinearity
note: TreatXLicence omitted because of collinearity
note: TreatXLNAlko omitted because of collinearity
note: Ctrl2XKeskiolutlaki omitted because of collinearity
note: Ctrl2XAlkoPOST omitted because of collinearity
note: Ctrl1XLicensePOST omitted because of collinearity
note: Ctrl2XLicensePOST omitted because of collinearity
note: Ctrl1 dropped because it is constant within group
note: Treat dropped because it is constant within group

Iteration 0:   log pseudolikelihood = -21726.074
Iteration 1:   log pseudolikelihood = -13618.697
Iteration 2:   log pseudolikelihood = -13602.826
Iteration 3:   log pseudolikelihood = -13602.824
Iteration 4:   log pseudolikelihood = -13602.824

Conditional fixed-effects Poisson regression     Number of obs     =      4,304
Group variable: Munid                            Number of groups  =        269

                                                 Obs per group:
                                                              min =         16
                                                              avg =       16.0
                                                              max =         16

                                                 Wald chi2(16)     =    1117.32
Log pseudolikelihood  = -13723.048               Prob > chi2       =     0.0000

                                   (Std. Err. adjusted for clustering on Munid)
```

Battery	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Time	.0310755	.0059275	5.24	0.000	.0194579	.0426932
Populationdensity	-.0006073	.0001664	-3.65	0.000	-.0009333	-.0002812
DistancetoAlko	-.0049325	.0017378	-2.84	0.005	-.0083386	-.0015265
Licensedummy	.0249639	.0677999	0.37	0.713	-.1079215	.1578493
Keskiolutlaki	.201197	.0800776	2.51	0.012	.0442478	.3581461
DistancePOST	.0026828	.0018293	1.47	0.142	-.0009026	.0062681
LicensePOST	.1246642	.0800748	1.56	0.120	-.0322795	.2816079
Ctrl1XLicense	.1755235	.2006377	0.87	0.382	-.2177191	.5687661
TreatXLicense	.2756584	.1258251	2.19	0.028	.0290457	.5222711
Ctrl1XAlko	.0023875	.0051893	0.46	0.645	-.0077832	.0125583
TreatXAlko	-.0011045	.0024393	-0.45	0.651	-.0058855	.0036765
Ctrl1XKeskiolutlaki	.1113764	.1199899	0.93	0.353	-.1237994	.3465523
TreatXKeskiolutlaki	.4124956	.1881285	2.19	0.028	.0437704	.7812207
TreatXlicensePOST	-.4515076	.168816	-2.67	0.007	-.7823808	-.1206343
Ctrl1xAlkoPOST	.0049729	.0150239	0.33	0.741	-.0244734	.0344192
TreatxAlkoPOST	-.0055765	.0038023	-1.47	0.142	-.0130289	.0018758
ln(Population)	1	(exposure)				

Table i.11. Batteries with the group setup

```
Conditional fixed-effects Poisson regression    Number of obs      =      4,304
Group variable: Munid                           Number of groups   =        269

                                                Obs per group:
                                                             min =         16
                                                             avg =       16.0
                                                             max =         16

                                                Wald chi2(16)      =    1117.32
Log pseudolikelihood  = -13723.048              Prob > chi2        =     0.0000

                                      (Std. Err. adjusted for clustering on Munid)
```

Battery	IRR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Time	1.031563	.0061146	5.24	0.000	1.019648	1.043618
Populationdensity	.9993929	.0001663	-3.65	0.000	.9990671	.9997188
DistancetoAlko	.9950796	.0017293	-2.84	0.005	.9916961	.9984747
Licensedummy	1.025278	.0695138	0.37	0.713	.8976981	1.17099
Keskiolutlaki	1.222866	.0979241	2.51	0.012	1.045241	1.430675
DistancePOST	1.002686	.0018342	1.47	0.142	.9990978	1.006288
LicensePOST	1.132768	.0907062	1.56	0.120	.9682359	1.325259
Ctrl1XLicense	1.19187	.239134	0.87	0.382	.8043514	1.766086
TreatXLicense	1.317398	.1657617	2.19	0.028	1.029472	1.685852
Ctrl1XAlko	1.00239	.0052017	0.46	0.645	.992247	1.012638
TreatXAlko	.9988961	.0024366	-0.45	0.651	.9941318	1.003683
Ctrl1XKeskiolutlaki	1.117816	.1341266	0.93	0.353	.883557	1.414183
TreatXKeskiolutlaki	1.510583	.2841838	2.19	0.028	1.044742	2.184137
TreatXlicensePOST	.6366676	.1074797	-2.67	0.007	.4573159	.886358
Ctrl1xAlkoPOST	1.004985	.0150988	0.33	0.741	.9758237	1.035018
TreatxAlkoPOST	.994439	.0037812	-1.47	0.142	.9870556	1.001878
ln(Population)	1	(exposure)				

Table i.12. Batteries with the group setup (IRR)

FATAL OFFENCES AND ATTEMPTS AT THESE

```
note: 10 groups (160 obs) dropped because of all zero outcomes

Iteration 0:   log pseudolikelihood = -2890.8572
Iteration 1:   log pseudolikelihood = -2859.6378
Iteration 2:   log pseudolikelihood = -2859.6313
Iteration 3:   log pseudolikelihood = -2859.6313

Conditional fixed-effects Poisson regression    Number of obs      =      4,144
Group variable: Munid                           Number of groups   =        259

                                                Obs per group:
                                                              min =         16
                                                              avg =       16.0
                                                              max =         16

                                                Wald chi2(7)       =      45.96
Log pseudolikelihood  = -2859.6313              Prob > chi2        =     0.0000

                                         (Std. Err. adjusted for clustering on Munid)
```

Murderetc	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Time	.0378045	.0169601	2.23	0.026	.0045634	.0710456
Populationdensity	-.0006635	.0005219	-1.27	0.204	-.0016865	.0003595
DistancetoAlko	-.001589	.0027453	-0.58	0.563	-.0069698	.0037918
Licensedummy	-.3416741	.123216	-2.77	0.006	-.583173	-.1001752
Keskiolutlaki	-.5713627	.2157687	-2.65	0.008	-.9942615	-.1484638
DistancePOST	.0045664	.0035442	1.29	0.198	-.0023801	.0115129
LicensePOST	.5985313	.1903691	3.14	0.002	.2254147	.9716479
ln(Population)	1	(exposure)				

Table i.13. Fatal offences (& attempts) without the group setup

```
note: 10 groups (160 obs) dropped because of all zero outcomes

Iteration 0:   log pseudolikelihood = -2890.8572
Iteration 1:   log pseudolikelihood = -2859.6378
Iteration 2:   log pseudolikelihood = -2859.6313
Iteration 3:   log pseudolikelihood = -2859.6313

Conditional fixed-effects Poisson regression    Number of obs     =      4,144
Group variable: Munid                           Number of groups  =        259

                                                Obs per group:
                                                              min =         16
                                                              avg =       16.0
                                                              max =         16

                                                Wald chi2(7)      =      45.96
Log pseudolikelihood  = -2859.6313              Prob > chi2       =     0.0000

                                 (Std. Err. adjusted for clustering on Munid)
```

Murderetc	IRR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Time	1.038528	.0176135	2.23	0.026	1.004574	1.07363
Populationdensity	.9993367	.0005216	-1.27	0.204	.9983149	1.00036
DistancetoAlko	.9984122	.002741	-0.58	0.563	.9930544	1.003799
Licensedummy	.7105798	.0875548	-2.77	0.006	.5581246	.9046789
Keskiolutlaki	.5647553	.1218565	-2.65	0.008	.3699966	.8620312
DistancePOST	1.004577	.0035604	1.29	0.198	.9976227	1.011579
LicensePOST	1.819445	.346366	3.14	0.002	1.252842	2.642295
ln(Population)	1	(exposure)				

Table i.14. Fatal offences (& attempts) without the group setup (IRR)

```
Iteration 0:   log pseudolikelihood = -2890.8572
Iteration 1:   log pseudolikelihood = -2851.8811
Iteration 2:   log pseudolikelihood = -2851.3485
Iteration 3:   log pseudolikelihood = -2851.3397
Iteration 4:   log pseudolikelihood = -2851.3397

Conditional fixed-effects Poisson regression    Number of obs     =      4,144
Group variable: Munid                           Number of groups  =        259

                                                Obs per group:
                                                              min =         16
                                                              avg =       16.0
                                                              max =         16

                                                Wald chi2(16)     =      66.12
Log pseudolikelihood  = -2851.3397              Prob > chi2       =     0.0000

                                    (Std. Err. adjusted for clustering on Munid)
```

Murderetc	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Time	.0375662	.0175216	2.14	0.032	.0032245	.071908
Populationdensity	-.0008338	.0005062	-1.65	0.100	-.0018259	.0001584
DistancetoAlko	-.0095837	.0097587	-0.98	0.326	-.0287103	.0095429
Licensedummy	-.176254	.1971295	-0.89	0.371	-.5626207	.2101128
Keskiolutlaki	-.6596543	.2377814	-2.77	0.006	-1.125697	-.1936112
DistancePOST	.0065782	.0053319	1.23	0.217	-.003872	.0170285
LicensePOST	.4965713	.2337962	2.12	0.034	.0383392	.9548034
Ctrl1XLicense	-1.395193	.6774475	-2.06	0.039	-2.722966	-.06742
TreatXLicense	-.3184232	.252287	-1.26	0.207	-.8128966	.1760502
Ctrl1XAlko	-.0278273	.0265711	-1.05	0.295	-.0799056	.0242511
TreatXAlko	.0100665	.0101352	0.99	0.321	-.0097982	.0299311
Ctrl1XKeskiolutlaki	.2663883	.2035201	1.31	0.191	-.1325036	.6652803
TreatXKeskiolutlaki	.8647168	.5742479	1.51	0.132	-.2607884	1.990222
TreatXlicensePOST	-.4057588	.5532978	-0.73	0.463	-1.490203	.678685
Ctrl1xAlkoPOST	-.0320385	.0199871	-1.60	0.109	-.0712125	.0071355
TreatxAlkoPOST	-.0087656	.0092499	-0.95	0.343	-.026895	.0093638
ln(Population)	1	(exposure)				

Table i.15. Fatal offences (& attempts) with the group setup

```
Iteration 0:   log pseudolikelihood = -2890.8572
Iteration 1:   log pseudolikelihood = -2851.8811
Iteration 2:   log pseudolikelihood = -2851.3485
Iteration 3:   log pseudolikelihood = -2851.3397
Iteration 4:   log pseudolikelihood = -2851.3397

Conditional fixed-effects Poisson regression    Number of obs      =      4,144
Group variable: Munid                           Number of groups   =        259

                                                Obs per group:
                                                              min =         16
                                                              avg =       16.0
                                                              max =         16

                                                Wald chi2(16)      =      66.12
Log pseudolikelihood  = -2851.3397              Prob > chi2        =     0.0000

                                  (Std. Err. adjusted for clustering on Munid)
```

Murderetc	IRR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Time	1.038281	.0181924	2.14	0.032	1.00323	1.074556
Populationdensity	.9991666	.0005058	-1.65	0.100	.9981757	1.000158
DistancetoAlko	.9904621	.0096656	-0.98	0.326	.9716979	1.009589
Licensedummy	.838405	.1652744	-0.89	0.371	.5697141	1.233817
Keskiolutlaki	.51703	.1229402	-2.77	0.006	.3244261	.8239782
DistancePOST	1.0066	.0053671	1.23	0.217	.9961354	1.017174
LicensePOST	1.643078	.3841454	2.12	0.034	1.039084	2.59816
Ctrl1XLicense	.2477853	.1678615	-2.06	0.039	.0656797	.9348025
TreatXLicense	.7272949	.183487	-1.26	0.207	.4435713	1.192498
Ctrl1XAlko	.9725563	.0258419	-1.05	0.295	.9232035	1.024547
TreatXAlko	1.010117	.0102378	0.99	0.321	.9902496	1.030384
Ctrl1XKeskiolutlaki	1.305242	.2656429	1.31	0.191	.8758998	1.945036
TreatXKeskiolutlaki	2.374333	1.363456	1.51	0.132	.7704439	7.317158
TreatXlicensePOST	.6664709	.3687569	-0.73	0.463	.225327	1.971284
Ctrl1xAlkoPOST	.9684693	.0193569	-1.60	0.109	.931264	1.007161
TreatxAlkoPOST	.9912727	.0091691	-0.95	0.343	.9734635	1.009408
ln(Population)	1	(exposure)				

Table i.16. Fatal offences (& attempts) with the group setup (IRR)