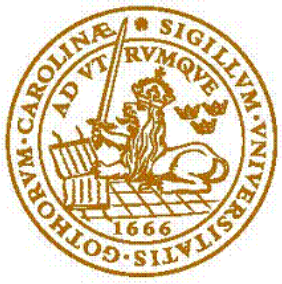

LUND UNIVERSITY

School of Economics and Management

The Influence of Income Inequality on Suicides - A study across age groups in Swedish counties

By Danna Garcés Delgado

Department of Economics

Master Thesis I - June 2014

Supervisor: Martin Nordin

Abstract

The *National Board of Health and Welfare* together with the *Public Health Agency of Sweden* suggests that a reduction in income inequality could decrease suicide rates. The thesis aims at testing if this statement is accurate and if the influence of income inequality on suicides is different across age groups. The estimations are done using a panel dataset consisting of all Swedish counties for the period 2007-2012 together with a fixed effects model. The main finding is that the only age group with significant results are young females, where income inequality is negatively correlated with suicide rates. Although insignificant for young males the coefficient is negative in their case too.

Keywords: Suicide, regional Gini-coefficients, panel dataset, Swedish counties, fixed effects

LIST OF CONTENTS

1. INTRODUCTION 1
1.1 AIM AND RESEARCH QUESTIONS 2
2. THEORETICAL BACKGROUND 3
3. PREVIOUS RESEARCH 5
4. DATA 6
4.1 DESCRIPTIVE STATISTICS 8
5. METHOD 11
5.1 MODEL SPECIFICATION 11
5.2 EMPIRICAL FRAMEWORK 12
5.3 POTENTIAL PROBLEMS 13
6. RESULTS 15
6.1 FEMALE RESULTS 15
6.2 MALE RESULTS 17
6.3 DISCUSSION 19
7. CONCLUSION 21
REFERENCE LIST 23

1. INTRODUCTION

In 2006, the *National Board of Health and Welfare* together with the *Public Health Agency of Sweden* released "Förslag till nationellt program för suicidprevention", which is a national program on how to prevent suicides in Sweden. The government decided the making of such a program was urgent due to the worrying signs that suicide rate for young people aged 15-24 was increasing (Socialstyrelsen, 2006). One of the suggestions made in the program was that society should aim at reducing social inequalities, e.g. measured with the Gini-coefficient. The reasoning is that individuals with low educational levels and low incomes have the highest risk of committing suicide. The circumstances they live in as well as the opportunities they have to improve their well being are the best when there are low social differences. Thus, a reduction in social inequality could help to reduce suicides. The suggestion is relevant to address because: 1. Income inequality in Sweden is increasing (OEDC, 2011), and 2. The effect of income inequality on suicide rates might differ across age groups.

There are two main theories about how income inequality can influence health; 1. The weak income hypothesis, says that health depends on where the individual in economic terms stands in relation to others, 2. The strong income hypothesis states that income inequality itself affects irrespectively of income level (Grönkvist, Johansson, & Niknami, 2012). The literature covering this relationship is vast and the results are contradicting, they depend on the type of data (individual or aggregated), the setting (within or between countries) and what variables are used for health and inequality (Bergh, Nilsson, & Waldenström, 2012).

If a causal effect of inequality on suicide rates is found then reducing inequality is in fact an achievable economic solution that politicians could put into practice. There are several reasons for why reducing suicide rates is beneficial for the society as a whole. Mental health problems impose large costs to society in the shape of sick leave and increased welfare costs. It would also of course reduce the immeasurable emotional costs incurred by individuals close to the suicide victim. However, reducing inequality in Sweden (a country that in international standards already has very low income inequality) could potentially have a significant long run effect on the Swedish economy. Thus it is important to further study the relationship between income inequality and suicide rates so that the potential causal effect can better be understood.

1.1 AIM AND RESEARCH QUESTIONS

The aim of the thesis is to study the asymmetry in the association between income inequality and suicide rates across age groups in Sweden. The fact that the association might be different across age groups suggests that political measures aiming at preventing suicides, such as reducing income inequality, might affect age groups differently. If this is the case politicians ought to take this into account when deciding on efficient prevention plans.

The research questions are:

- Does income inequality at the county level influence suicide rates?
- How does the influence of income inequality on suicide rate for youth aged 15-24 differ from the rest of the population?

The research questions are answered with a panel dataset consisting of all 21 Swedish counties for the period 2007-2012. The empirical model is a fixed effect model. Throughout the thesis, the suicide rates for both genders are studied separately. The reason is that it is more informative to do it this way since suicide rates in general differ a lot between females and males (Socialstyrelsen, 2006).

This thesis contributes to literature by testing the impact of income inequality across age groups for both sexes. To my knowledge there have been no similar studies in Sweden. The main findings are that income inequality at the county level only influences the suicide rates for young females. The relationship is unexpectedly negative. Although the relationship is insignificant for young males the coefficient is negative in their case too. Thus the results imply that the influence of income inequality differs across age groups, and that an increasing income inequality does not seem to increase suicide rates for any age group.

The thesis contains seven sections; section 2, gives a theoretical background on suicides; 3, offers an overview of previous research; 4, describes the data; 5, describes the method and the model specification; 6, presents the results and a discussion; at last section 7 concludes the thesis.

2. THEORETICAL BACKGROUND

In 1897 the sociologist Emile Durkheim released his famous book *Le Suicide* with a new theory about the causal factors of suicides (Durkheim, 1952). He did a quantitative study with suicide rates and other variables, such as unemployment and divorce rates, for several European countries. From his results he concluded that there are no specific individual factors that can solely cause suicide. For example is depression an underlying factor that all suicide victims have in common, but it is not a causal factor since it is impossible to give an exact measure of depression that will cause an individual to take his or her own life. He instead claimed that the causal effect comes from social structures. He further claimed that societies have constant suicide rates as long as the social structure is unchanged. One could argue that income inequality is a social structure of society therefore it is interesting to see if suicides have been influences by the increasing income inequality in Sweden.

The mechanisms affecting an individual to take the decision of committing suicide can be hard to study because the set of causal factors might change from one individual to another. In the national report there are a few main types of factors that are believed to have an effect. The factors range from; individual, such as genes and stress levels; inter personal, e.g. life crisis; social structural, such as social cohesion; and ultimately access to methods to use in order to commit suicide like weapons (Socialstyrelsen, 2006). The focus of this thesis is on a social structural factor measured as income inequality.

There are economic approaches to the theory about suicides as well, that partly build upon sociology and psychology. Hamermesh & Soss (1974) emphasize that suicides cannot solely be explained with an economic approach, but there is a value in analyzing suicides partly with hypothesis from economic theory. They create a utility function for the average individual in a group, and assume that an individual commits suicide when the discounted lifetime utility reaches zero. In a preliminary working paper by Becker & Posner (2004) the author's base their work on Hamermesh & Soss paper, and further develops the utility-maximizing approach with an emphasis on the link between mental disorder and suicide.

To better understand how income inequality can affect health there is, as already mentioned, two main theories; the weak and the strong hypothesis:

The weak income hypothesis

The assumption made in the weak income hypothesis is that individuals compare their relative position in society with individuals that are better off and exclude everyone that is less advantaged than them. Being relatively deprived can affect health through increased psychosocial stress (Wilkinson R. G., 1997). Thus the health of the relatively poorest individual is at the largest risk of being adversely impacted by income inequality. The weak income hypothesis is partly built upon the idea of the absolute income hypothesis which says that the relationship between an individual's health and income is concave. A higher absolute value of income will increase the individual's health since the individual can afford to purchase better health in the form of better nutrition and access to health care. The effect of an increase in income is however diminishing as income gets higher. The other part of the weak income hypothesis states that an individual's health depends on the income inequality in the community or population, that is; the relative income of the individual matters as well. (Wagstaff & van Doorslaer, 2000)

The strong income hypothesis

The assumption made in the strong income hypothesis is that the everyone is affected by income inequality, regardless of income level. A potential mechanism behind the strong hypothesis is through the decrease in trust between individuals living in a less cohesive society (Kawachi, Kennedy, & Lochner, 1997). A reduced feeling of social cohesion can increase mortality rates via disinvestment in social capital (ibid.)

The suggestions in the national prevention plan is built upon the weak hypothesis, they assume that the poorest individuals are worse off when income inequality increases. Due to the data in hand it is not possible to control for the income level of the individuals committing suicide hence it is not possible to test the weak income hypothesis, only the strong income hypothesis.

3. PREVIOUS RESEARCH

The weak income hypothesis is methodologically tested by estimating income inequality for different income earning groups. The support for the weak income hypothesis depends on the level of data used. Studies use either individual data or aggregated data, such as cross country or cross-county panel data. Individual data typically proves a positive relationship between income inequality and suicide risk i.e. people living in poor areas are more prone to commit suicide compared to people living in richer areas (Inagaki, 2010). A study in Sweden showed that poorer areas in Stockholm have higher suicide rates in comparison to richer areas (Ferrada-Noli & Åseberg, 1997). A possible explanation is the psychosocial impact of low social status, where low status is associated with violence, shame and depression (Wilkinson, 2006).

There are also studies that show that there is no support for the weak income hypothesis. A Swedish study shows that since Sweden has very low variation in income inequality there is no significant effect of income inequality on health (Grönkvist, Johansson, & Niknami, 2012). The study is based on the fact that newly arrived refugees were randomly settled to different counties by swedish authorities. The results showed that municipalites with higher income inequalties had no statistically significant effect on the risk of being hospitalized (ibid.)

Other studies even show that it is the richest people that are the ones that are worse off when income inequality increases. Weich et al (2004) finds that causal factors of mental illness are stronger with higher inequality, and that the causal factors become stronger for richer individuals. Mellor and Milyo (2002) also find that the richer groups feel less happy when income inequality goes up.

Wilkinson (1997) argues that in developed countries the income inequality itself affects health of *all* individuals. The reason for this is that mortality is affected more by the distribution of income than the absolute value of the living standards in the society. The results differ depending on the type of research. In a cross country analysis it is found that more people suffer from mental illnesses in more unequal countries (Wilkinson & Picket, 2009). One study shows that the relationship differs between genders, the results implicated that mental illness among males is not correlated with income inequality while females mental illness is (ibid.). A study in New York found that income inequality affected age groups differently, where young poeple between 15-35 reacted positively to income inequality (Miller et. al., 2005).

4. DATA

A dataset is manually put together from various public Swedish sources. The dataset consists of aggregated data from 21 counties (n=21) in Sweden, during a period of six years (t=6).

A list of the variables used in the following empirical analysis can be found in Table 1 together with a short description and the source. All the variables are collected yearly for each county and when possible also by gender.

Table 1. List of variables used in empirical analysis

Variable name	Description	Source
Suicide	Suicide per 100 000 across age group and by gender.	The National Board of Health and Welfare, and Statistics Sweden
Gini	Measure of income inequality calculated from earnings before tax (excluding capital earnings). Collected both by gender and the entire population.	Statistics Sweden
Average income	Calculated from earnings before tax (excluding capital earnings). Collected both by gender and the entire population.	Statistics Sweden
Unemployment	Share of unemployed individuals aged 15-74 years that are part of the workforce. Collected by gender.	Statistics Sweden
Alcohol consumption	Litres of alcoholic beverages (recalculated as 100% alcohol) per 15+ year old individual sold by Systembolaget.	Public Health Agency of Sweden
High school eligibility	The share of 9^{th} graders that finished elementary school with grades enough to be eligible to attend high school. Collected by gender.	Public Health Agency of Sweden

The dependent variable *Suicide* is constructed by first adding the two variables *intentional self-harm* (ICD: X60-X84) and *event of undetermined intent* (ICD: Y10-Y34)[1] from the cause of death registry collected by the National Board of Health and Welfare. Intentional self-harm is when an individual purposely self-inflicts poisoning or injury. Events of undetermined intent are deaths where there is not sufficient information on the cause of death to distinguish between accident, self-harm and assault. Although the numbers of event of undetermined intent are very low, they are included in the *Suicide* variable in order to follow international praxis (Ferrada-Noli, 1997). The number of suicides is calculated for all Swedish citizens above 15 years registered in Sweden regardless whether the death occurred in Sweden or abroad. Individuals' temporarily visiting Sweden and asylum seekers are excluded. The numbers of suicides for individuals under the age of 15 are excluded from the data. The reason is that suicides in this age category are almost non-existent. In order to construct the suicide rates the population size was collected from Statistics Sweden. The suicide rates are constructed by gender and age group, and are throughout the thesis analysed separately.

The explanatory variable of interest, income inequality, is measured with the *Gini-coefficient* constructed with income before tax, excluding capital earnings. It is an index that goes from 0 to 1, where 0 is complete equality, and 1 is complete inequality. The Gini-coefficient is calculated by Statistics Sweden, and they only started doing this at the county level in 2007 which is the reason for why the time period in this thesis is somewhat short. Only using the Gini-coefficient in the analysis of the impact of income inequality on health does not say anything about the income level in the county, therefore *Average income* is added as a control. Statistics Sweden uses the same income measure when they calculate the Gini-coefficient and the average income. The two variables are collected for each gender and for the entire population, separately. There are other measures of income inequality that can be used for the purpose of studying its effect on health, however the Gini-coefficient is the most common. If the suggestion made in the national prevention plan is accurate, the expectation of this variable is to be positively correlated with suicide rates. That is, an increase in the Gini-coefficient should lead to an increase in suicide rates. A higher absolute value of average income is expected to be negatively correlated with suicide rates, since a higher absolute level of incomes are expected to improve health.

[1] ICD is the International Classification of Diseases that all members in the World Health Organisation use to code deaths.

Unemployment is measured as the share of unemployed individuals aged 15-74 years that are part of the workforce. It is collected by gender. The variable is included in the model since studies show that unemployment is an important factor of suicide (Gerdtham & Johannesson, 2002). The expectation of the variable is a negative relationship with suicide rates.

The amount of alcohol litres sold by Systembolaget is a proxy for *Alcohol consumption*. It is measured as the amount of litres sold per individual above 15 years. There are studies that show that alcohol has an effect on the amount of suicides, and particularly for young individuals (Ramstedt, 2001). The expectation is a positive relationship, where higher alcohol consumption leads to more cases of suicide.

High school eligibility refers to the shares of pupils eligible to start high school after finishing 9th grade[2]. The variable is collected by gender. Not being able to fulfil the educational requirements in 9th grade is found to increase suicidal attempts for both genders (Jablonska et.al., 2009). The expectation of the variable is that an increase in share of pupils that fulfil the educational requirements the less suicides will take place.

4.1 DESCRIPTIVE STATISTICS

In Table 2, the descriptive statistics for the variables included in the model are presented. The variables that are collected by gender are presented next to one another in two columns while the population variables are presented at the bottom of the table. For each variable there are four values, the mean, standard error, minimum and maximum. The calculations are done for three different components; *overall*, *between* and *within*.

The overall component is calculated from the total amount of observations, i.e. for the 21 counties during a period of 6 years, $(N = 126)$. The between component calculates the averages of each county, $(n = 21)$. Lastly, the within component is also calculated over all observations, it takes each observation x_i, subtracts the county average $\bar{x}_i$ and adds the global mean $\bar{\bar{x}}$ (the global mean is the mean from the overall component).

[2] The requirements changed in 2011; before that the requirement was at least a passing grade in Swedish, English and Math. Now there are five additional subjects the pupils have to pass in order to be eligible to start high school.

Table 2. Descriptive statistics for both genders

		Males				Females			
Variable		Mean	Std. D	Min	Max	Mean	Std.D	Min	Max
Youth Suicide	overall	19.083	11.38	0	64.613	7.088	6.374	0	26.487
	between		5.311	11.786	33.125		2.886	0	14.201
	within		10.12	-2.484	51.494		5.712	-2.569	22.145
Work Suicide	overall	30.116	7.760	12.537	62.364	11.896	5.101	0	42.448
	between		4.297	22.236	38.828		2.228	8.844	16.574
	within		6.518	12.822	53.652		4.610	-4.456	37.992
Pensioners Suicide	overall	31.687	14.27	4.845	97.276	9.731	7.071	0	42.808
	between		7.974	16.494	52.362		3.224	3.551	15.023
	within		11.941	-2.660	76.602		6.326	-1.274	42.713
Total Suicide	overall	28.620	6.670	16.397	52.977	10.571	3.951	0	36.370
	between		4.504	19.716	37.643		1.734	7.043	13.021
	within		5.001	17.204	43.953		3.568	-1.544	34.826
Gini - coefficient	overall	0.324	0.028	0.285	0.408	0.312	0.021	0.276	0.37
	between		0.028	0.298	0.405		0.020	0.288	0.363
	within		0.008	0.307	0.335		0.008	0.295	0.324
Average income	overall	280257	22201	228533	365907	206249	14543	182416	269630
	between		20557	242492	348234		11601	196577	252169
	within		9338	261565	300446		9073	188695	223710
Unemployment	overall	7.795	1.874	3.700	11.800	7.637	1.408	4.3	11.5
	between		1.098	5.917	9.433		0.92	6.267	9.317
	within		1.534	3.912	10.879		1.082	5.021	9.971
High school	overall	86.802	2.273	80	91	89.508	1.926	85	93
	between		1.9	82.833	90		1.671	85.333	92.167
	within		1.304	82.968	89.302		1.013	86.175	92.341

Descriptive statistics for population variables

Variable		Mean	Std. D	Min	Max
Gini-coefficient	overall	0.330	0.024	0.290	0.399
	between		0.024	0.303	0.395
	within		0.007	0.313	0.340
Average income	overall	242792	17606	205913	316607
	between		a	220241	298855
	within		9230	225179	260821
Alcohol	overall	5.793	1.307	3.5	9.8
	between		1.302	3.983	9.07
	within		0.284	5.093	6.526

By looking at Table 2, we can see that the mean of suicides is higher for males than for females in all age groups. The highest mean for males during this time period, is for pensioners with a mean values of almost 32 suicides per 100 000. The highest for females is among the work group, with a value of almost 12 suicides. The male youth suicide rate is

surprisingly high, with a mean of 19 suicides. By comparing the different min and max values for the between component we can get an estimate on the variation between the county averages. Both for females and males the between values show that counties suicide rates vary a lot. For example, by looking at the female youth suicide the county averages varied between 0 to 14 suicides. By looking at the standard deviation values for the within component we can get an estimate if the suicide rates over time. If the value is 0 it means the value has been constant. The standard deviations for males within suicides are higher than for females meaning that male suicide rates are more volatile over time than female suicide rates.

The Gini-coefficient for males is slightly larger than for females, which means that income inequality among males is somewhat larger than among females. The highest Gini-coefficient for males is 0.408 while the lowest measured Gini for males is 0.285, and 0.276 for females. Males also have a higher variation between counties. The very low standard deviation in the within components is the same for both genders (0.008), the low values imply that the Gini-coefficient has had a very low variation during the six years studied in this thesis. In comparison to the population Gini, presented at the bottom of the table, we can see that the global mean of the population Gini, 0.330, is quite representative for both the gender specific Gini's global means.

The Gini-coefficient in itself does not provide any information about the income level. Although both genders have similar Gini-coefficients, the average income between the genders differs quite much. The global mean average income of males is 74 000 SEK higher than for females. Males also have a higher variation between counties, where the county with highest average income has a value of 100 000 SEK more than the than the lowest average income county. In comparison to the population average income we can see that population average income is almost 242 800 SEK, which is about 36 500 SEK higher than the average female income, and about 37 400 SEK lower than the male average income.

Comparing the unemployment rates between genders, we see that they are quite similar. There is a higher variation between counties for male unemployment and within males than for females. Looking at the high school variable between genders we see that females have a higher share of pupils eligible to high school. The variation between the counties varies a lot for both genders, the lowest county average is measured for males with only 82.8% of boys being eligible to start high school, while females have the highest value with 92.2% being

eligible. Lastly the alcohol consumption for the entire population varies a lot between counties, where the min and max county averages are 3.98 and 9.07 respectively.

5. METHOD

With the panel data in hand it is suitable to conduct a fixed effect model. Since there is heterogeneity problems when comparing different counties a fixed effect model helps to hold all unobserved time invariant factors on the county level constant. In other words, the idea of the fixed effect estimator is to see if a change in income inequality *within* a county has any influence on the suicide rates in that county, while keeping all the potential factors that are time invariant and unobserved fixed. When heterogeneity problems are present an ordinary least squares (OLS) is biased if the omitted explanatory variables are correlated with the regressors included in the model (Kennedy, 2008). The regressions are done using clustered standards errors on the county level. To cluster by county means that the observations are independent across counties while they are allowed to correlate within the county. The estimated coefficients are not affected by clustering, only the standard errors (Stata, 2013).

5.1 MODEL SPECIFICATION

We want to test if the dependent variables suicide rates in different age groups are determined differently by the county's income inequality. A set of control variables based on previous research and theoretical relevance are included in the model in order to try to reduce omitted variable bias.

The main group of interest is youth aged 15-24. Suicide rates among pensioners has always been relatively high in comparison to the rest of the population, therefore it also seemed justified to make individuals above 65 as one group called *pensioner*. Different age groups for the ages 25-64 in five year classes were tested, however there were no logical cut off points, thus they are kept as one group called *work*. The total suicides by gender are also constructed into a group called *total*. Thus, in total there are three different age groups, and the total used as the dependent variable for each gender.

Specification of model 1:

$$Sui_{ita} = \beta_0 + \beta_1 Ginipop_{it} + \beta_2 Incomepop_{it} \sum \beta_n X_{it} + \delta_i + \tau_t + \varepsilon_{it} \quad (1)$$

Where *i* represents county (i = 1, 2, 3, ..., 21), *t* represents year (t = 2007, 2008, ..., 2012) and *a* represents age group (a = *Youth* 15-24; *Work* 25-64; *Pensioners* 64+; *Total* 15+). The dependent variable Sui_{ita} is amount of suicides per 100 000. The explanatory variable is income inequality for the whole population measured by $Ginipop_{it}$. $Incomepop_{it}$ is the average income for the whole population. X_{it}, is a vector of *n* control variables. δ_i, and τ_t are county and time fixed effects respectively. ε_{it}, is the error term.

A second model is also estimated which is similar to model (1). The difference is that the variables for income inequality and average income are now gender specific.

Specification of model 2:

$$Sui_{ita} = \beta_0 + \beta_1 Ginigender_{it} + \beta_2 Incomegender_{it} \sum \beta_n X_{it} + \delta_i + \tau_t + \varepsilon_{it} \quad (2)$$

5.2 EMPIRICAL FRAMEWORK

Before doing the regressions the average income variable is logged in order to make it comparable to the rest of the variables, i.e. measured in relative terms. The variable *High school* is only relevant as a control for the group *Youth*.

In the regression analysis that follows there are six dependent variables used. These are suicide rates for the three age-groups discussed above and separately for females and males. Additionally there are two main model specifications are tested for both females and males. The first model contains the Gini-coefficient for the whole population and population average income as the explanatory variables along with other controls. This implies that individuals compare themselves with everyone within the county and not just individuals of the same gender. The second model contains a gender specific Gini-coefficient and gender specific average income as the explanatory variables along with other controls. This would imply that

individuals compare themselves to same sex individuals instead of comparing themselves with everyone in the county.

Table 3 shows an overview of the model specification with the two variables that differ between the models; it excludes the control variables and the error term. All the regressions are done using Stata.

Table 3. An overview of the model specification (excluding controls and the error term)

Model	Female	Male
1	$Sui\ female_{\ ita} = \beta_0 + \beta_1 Gini\ pop_{it} + \beta_2 Income\ pop_{it}$	$Sui\ male_{ita} = \beta_0 + \beta_1 Gini\ pop_{it} + \beta_2 Income\ pop_{it}$
2	$Sui\ female_{\ ita} = \beta_0 + \beta_1 Gini\ female_{it} + \beta_2 Income\ female_{it}$	$Sui\ male_{ita} = \beta_0 + \beta_1 Gini\ male_{it} + \beta_2 Income\ male_{it}$

5.3 POTENTIAL PROBLEMS

Swedish public data sources are generally considered reliable. However, that does not preclude the possibility of problems with the data. Measurement errors may exist along with general problems involving the concepts and definitions of the variables used.

A measurement error in suicide rates could be present if there are some suicides that are classified as accidents. This is partly solved by the fact that *event of undetermined intent* are included in the number of suicides. But if for example the social stigma is high for males to commit suicide and they do it by crashing their car there could still be some suicides for males that are impossible to distinguish from accidents.

A potential problem of using a Gini-coefficient constructed with income before tax is that it could give an overestimation of the "true" income inequality since there is a redistribution of welfare once taxes are paid. On the other hand, the Gini-coefficient used excludes income from capital earnings which could underestimate the income inequality. However if there are no differences in the measurement error over time it is not a problem.

The variable with largest measurement error in this data set is probably alcohol consumption. It is very hard to measure the real consumption of individuals. Ideally there would be a

measurement of consumption both by gender and age group. It is safe to assume that the real alcohol consumption is higher since not all consumption is obtained from purchases from Systembolaget. Consumption through import and visits to bars and restaurants are not included in this measure. However, if the drinking habits across age groups are time invariant this would not pose a problem since it is controlled for in the fixed effect model.

With the data at hand it is not possible to control for the income or educational level of the individuals who have committed suicide. Thus it is not possible to fully test the relevance of the suggestion made in the suicide prevention plan. In other words, we cannot test if counties with larger increase in income inequality have more suicides among individuals with low socioeconomic status. This thesis thus tests as already mentioned the influence of income inequality on suicide rates for *all* individuals irrespective of income and education level of the suicidal victims.

One important note to discuss when using aggregated data is ecological fallacy (Morgenstem, 1982). That is to make a causal interpretation at the individual level when the estimations use aggregated data about the regions the individuals live in. In this case it means it is impossible to make any causal linkage between income inequality at the county level and the inhabitant's suicide risk, because it is impossible to exclude confounding effects. However as is pointed out by Kaplan et.al (1996), income inequality is a variable that can only be measured with aggregated data because it is a property of a population and not an individual.

A potential limitation is that the Gini variable is not lagged in the regression models. This means that the influence of income inequality is measured at the same time (year) as the suicide events. If there is a lagged effect of income inequality on health it could be problematic not to lag the Gini in the models. The time period of the Gini is too short in the data used in order to try estimate lags properly.

Another potential problem could perhaps be if people who commit suicide are more prone to migrate between counties, then the exposure to inequality is not a factor affecting suicide.

6. RESULTS

The results are presented by gender. First, the female results are presented for the two models, and thereafter are the male results presented. The last part of this section provides a discussion of the main findings.

6.1 FEMALE RESULTS

Table 4. Regression output for females, with population income and Gini

	Female suicides			
VARIABLES	(1) Youth 15-24	(2) Work 25-64	(3) Pensioner 65+	(4) Total Suicide 15+
Gini population	-752.160*	-172.833	374.013	-157.481
	(421.466)	(493.140)	(422.892)	(366.358)
Log of average income pop	-157.471*	-59.863	-166.979	-110.584
	(86.741)	(105.150)	(204.313)	(93.252)
Unemployment rate female	1.709**	-0.128	-1.047	-0.130
	(0.813)	(0.536)	(0.953)	(0.565)
High School eligibility female	1.253			
	(0.764)			
Alcohol consumption pop	10.153**	-1.283	-3.520	-0.143
	(4.491)	(4.930)	(3.609)	(2.825)
Constant	2,010.451*	812.868	1,976.483	1,426.482
	(1,157.879)	(1,375.143)	(2,621.434)	(1,238.264)
Observations	126	126	126	126
R-squared	0.145	0.118	0.045	0.079
Number of regions	21	21	21	21

Note: Each column represents a separate regression where the dependent variables are suicide per 100 000 in each age group. County and time fixed effects are included. Clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In Table 4, the regression output for females using population income and population Gini is displayed. The only group that shows significant results is youth. The Gini-coefficient is significant at the 10%-level. It says that a unit increase in the Gini-coefficient would lead to a decline of 752 suicides per 100 000 habitants. This is not in accordance with expectations and is further discussed in the next section. Average income is significant at the 10%-level. It says that a 1% increase in average income leads to 1.57 less suicides per 100 000. The result is in

accordance with expectations where a higher average income leads to a reduction of suicides. Unemployment rate is significant at the 5%-level, it says that a one percentage unit increase in unemployment would lead to an increase of 1.71 suicides per 100 000. Alcohol is also significant at 5%-level, it says that a one litre increase of alcohol per person leads to an increase of 10.15 suicides per 100 000. Both unemployment and alcohol are in accordance with expectations.

Next we estimate the same model but with Gini and income switched from population variables to being gender specific ones.

Table 5. Female regression output with gender specific income and Gini

	Female suicides			
VARIABLES	(1) Youth 15-24	(2) Work 25-64	(3) Pensioner 65+	(4) Total Suicide 15+
Gini female	-945.970*	-836.555	391.749	-554.587
	(479.375)	(623.984)	(621.881)	(467.789)
Log of average income female	-217.995	-279.968	-146.682	-240.120
	(150.169)	(192.470)	(310.754)	(166.506)
Unemployment rate female	2.112**	0.195	-1.050	0.116
	(0.789)	(0.606)	(0.967)	(0.598)
High school eligibility female	1.405*			
	(0.724)			
Alcohol consumption pop	6.133*	-2.701	-2.128	-1.442
	(3.388)	(3.390)	(3.549)	(1.686)
Constant	2,768.341	3,681.361	1,694.906	3,104.544
	(1,897.660)	(2,482.290)	(3,947.049)	(2,152.215)
Observations	126	126	126	126
R-squared	0.156	0.168	0.041	0.111
Number of regions	21	21	21	21

Note: Each column represents a separate regression where the dependent variables are suicide per 100 000 in each age group. County and time fixed effects are included. Clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In Table 5, the regression output for females using gender specific income and Gini is displayed. The results are similar to when using population Gini and income in the sense that the only group with significant results is youth. Gini is significant at the 10%-level, it says that unit increase in Gini leads to a decrease of almost 946 suicides per 100 000. This is again an unexpected result discussed in the next section. Average income is no longer significant but still has a negative coefficient. Unemployment is significant at the 5%-level. A one percentage unit leads to a 2.11 increase of suicide per 100 000. Unlike previous model, high school is significant. It is significant at the 10%-level. It says that a one percentage unit increase of pupils eligible to start high school leads to an increase of 1.41 suicides per 100 000. Lastly is alcohol significant at the 10%-level, where a one litre increase leads to 6.13 more suicides per 100 000.

In sum, the results for females are the following. The only group with significant results in both models is youth. Both have a significant negative relationship between income inequality and suicide rates which is unexpected. The coefficient is also higher when using gender specific income and Gini. Unemployment is significant in both models but slightly stronger in the second one. Alcohol is also significant in both models, however much stronger in the first model. The difference between the models is that income is only significant in the first one, the coefficient sign is however negative in the second model as well.

6.2 MALE RESULTS

Now the results for males are presented. First the regression output of the first model with population Gini and income is presented and later the output with gender specific Gini and income.

From Table 6, we conclude that there are no significant results for any of the age groups. However, the Gini-coefficient for youth is negative as it is in the results for female youth.

Table 6. Regression output for males, with population income and Gini				
	Male suicides			
VARIABLES	(1) Youth 15-24	(2) Work 25-64	(3) Pensioner 65+	(4) Total Suicide 15+
Gini pop	-474.149 (623.535)	-163.324 (513.832)	-361.040 (663.635)	-281.488 (375.981)
Log of average income pop	-59.650 (189.283)	34.546 (224.401)	71.433 (618.249)	23.621 (237.251)
Unemployment male	0.887 (1.811)	-1.133 (0.839)	-0.341 (2.128)	-0.642 (0.794)
High school eligibility male	-0.160 (1.176)			
Alcohol consumption pop	1.196 (6.564)	-2.668 (4.656)	-11.999 (8.549)	-3.952 (3.399)
Constant	904.841 (2,439.588)	-322.377 (2,857.396)	-668.364 (7,682.380)	-148.279 (3,003.793)
Observations	126	126	126	126
R-squared	0.043	0.043	0.156	0.085
Number of regions	21	21	21	21

Note: Each column represents a separate regression where the dependent variables are suicide per 100 000 in each age group. County and time fixed effects are included. Clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Next we estimate the same model but with Gini and income switched from population variables to being gender specific ones.

In Table 7, the regression output for males, with gender specific income and Gini is displayed. As in previous model no results are significant. However, also in this case is the Gini coefficient negative.

Table 7. Regression output for males, with gender specific income and Gini

	Male suicides			
	(1)	(2)	(3)	(4)
VARIABLES	Youth 15-24	Work 25-64	Pensioner 65+	Total Suicide 15+
Gini male	-219.515	-456.520	-268.068	-388.038
	(552.682)	(423.060)	(705.280)	(310.958)
Log of average income male	-11.649	49.620	46.689	35.453
	(168.716)	(144.324)	(434.418)	(161.542)
Unemployment male	0.815	-0.749	-0.389	-0.434
	(1.914)	(0.882)	(2.078)	(0.817)
High School eligibility male	-0.174			
	(1.204)			
Alcohol consumption pop	0.503	0.037	-11.863	-2.392
	(8.325)	(4.876)	(8.933)	(3.224)
Constant	236.701	-442.612	-402.415	-278.248
	(2,199.566)	(1,844.012)	(5,420.478)	(2,050.673)
Observations	126	126	126	126
R-squared	0.040	0.057	0.155	0.096
Number of regions	21	21	21	21

Note: Each column represents a separate regression where the dependent variables are suicide per 100 000 in each age group. County and time fixed effects are included. Clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In sum for males are the results the following. There are no significant result in any of the models. However is the Gini effect for youth negative in both models as in the models with female suicide rates.

6.3 DISCUSSION

The most interesting result is the negative relationship between the Gini-coefficient and suicide rates for young females. It is significant at the 10%-level in both models, i.e. model 1, using population income and Gini, and model 2, using gender specific variables instead. The results imply that when income inequality increases suicide rates go down for young females. And although the relationship is insignificant for males, the effect is negative for them as well. This result is unexpected since it contradicts the assumption made by the *Board of*

Health and Welfare, and the *Public Health Agency of Sweden.* In their suicide prevention plan one of the suggestions is that society should aim at reducing income inequality since this would help to decrease suicide rates. Assuming the results are true it means that their suggestion is not accurate for any of the age groups, especially for young females. The conclusion of this would thus be that reducing income inequality is not an effective measure to decrease suicide rates, and in particularly not for young females.

If the results are true could one explanation for why the result is significantly negative for young females be that in times of increasing income inequality there has been an increased worry about youth mental state. Thus, as a precaution plan more and better professional help is provided. This would then lead to a reduction of suicide rates although income inequality has increased. And the explanation for young males could be that they have accepted the help to a less extent than females, thus still showing a negative relation but an insignificant one.

Also supporting the potential explanation above, is that females attempt suicide more often than males. Thus, it could be that help is given more to females than males since females after a suicidal attempt are in contact with medical care. (Socialstyrelsen)

Another significant result that is unexpected is the positive relationship between female youth suicide rates and the share of female pupils eligible to start high school. If the result is true it means that when the share of females that fulfil the educational requirements increase it leads to more suicides. An explanation for this could be that as a consequence of better grades, stress levels increase which deteriorate mental health ultimately leading to more suicides. This relationship is however only significant in model 2, using gender specific income and Gini. In model 1, although insignificant it has the same positive coefficient. A possible interpretation for the variable only being significant in model 2 is that the effect only becomes significant when the model controls for reference group. In other words, the positive relation between high school eligibility and suicides is only captured when females compare their income level to other females, and not the entire population.

A possible reason for why we get so few significant results, could be that the within variation of the Gini coefficients are very low. Thus, there is not enough variation in income inequality during the time period studied with the data at hand. Although the time period is short, the advantage of having a panel dataset is that we are able to study the dynamics in several counties. However, here the dynamics of the Gini between counties is quite similar.

It is interesting that Durkheim (1951) already in his time saw that the male suicide rates were much higher than for females. The fact that this is still present could support the idea that factors causing suicide are due to social causes if the stigmatization for males to seek help for mental issues is larger than for females. The expectation for how genders are "supposed" to act are based on social structures. Also, if this is true it might imply that the measurement error for male suicides is potentially higher if more males wanting to commit suicide make it look like accidents.

7. CONCLUSION

The influence of income inequality on suicide rates differs across age groups, this is supported by the fact that no models using total suicide rates show significant results. And from the results it is very interesting that the only group that differed from the rest of the population are youth. From the results it is also possible to conclude that females and males are influenced differently by income inequality. The results also suggest that the reference group matters for females, and that they compare themselves more to other females than the entire population.

Doing studies on suicide prevention has implications on future potential costs. Society could benefit of investing in the most profitable suicide prevention plan today rather than paying the costs in the future. Assume that there exists an intergenerational transmission of mental illnesses; this would imply that children to parents with poor mental health have a larger risk of also developing the same health issues in the future. Thus, a suicide prevention plan aiming at reducing mental illness among youth today helps to break a vicious cycle of "creating" new generations of children more prone to committing suicide. Health care expenses (such as therapy, hospitalization and medication), and costs due to loss of work (such as social benefits and decreased production) could diminish considerably if less people suffered from suicidal thoughts.

A suggestion for future studies is to study the relationship between income inequality and suicides at the county level in Sweden for a longer time period. In particular since income inequality has been increasing in Sweden. It would also be interesting to lag the Gini, to see if there is a delayed influence of income inequality on health.

Another suggestion for future studies is to test the strong income hypothesis by constructing a panel data set over regions in Sweden, and regress share of suicides on share of immigrants in the county. One would, though, have to show that the increase in the share of immigrants has lead to a less cohesive society, e.g. by decreasing trust levels and language barriers.

REFERENCE LIST

Becker, G. S., & Posner, R. (2004). Suicide: An Economic Approach (Working paper).

Bergh, A., Nilsson, T., & Waldenström, D. (2012). *Blir vi sjuka av inkomstskillnader?* Lund: Studentlitteratur AB.

Durkheim, E. (1951). *Suicide - A study in sociology.* Norfolk: The Free Press.

Ferrada-Noli. (1997). A cross-cultural breakdown of Swedish. *Acta Psychiatr Scand* , 108-116.

Ferrada-Noli, M., & Åseberg, M. (1997). Psychiatric Health, Ethnicity and Socioeconomic Factors Among Suicides in Stockholm. *Psychological Reports 81* , pp. 323-332.

Gerdtham, U.-G., & Johannesson, M. (2002). A note on the effect of unemployment on mortality. *Journal of Health Economics 22 (2003)* , 505–518.

Grönkvist, H., Johansson, P., & Niknami, S. (2012). Income inequality and health: Lessons from a refugee residential assignment program. *Journal of Health Economics 31* , pp. 617-629.

Hamermesh, D. S., & Soss, N. M. (1974). An Economic Theory of Suicide. *Journal of Political Economy, 82(1)* , 83-98.

Inagaki, K. (2010). Income Inequality and the Suicide Rate in Japan: Evidence from Cointegration and La-Var. *Journal of Applied Economics 14(1)* , pp. 113-133.

Jablonska, B., Lindberg, L., Lindblad, F., Rasmussen, F., & Östberg, V. (2009). School performance and hospital admissions due to self-inflicted injury: a Swedish national cohort study. *International Journal of Epidemiology, 38(5)* , 1334-1341.

Kaplan, G., Pamuk, E., Lynch, J., Cohen, R., & Balfour, J. (1996). Inequality in income and mortality in the United States: analysis of mortality and the potential patways. *British Medical Journal 312 (7037)* , pp. 999-1003.

Kawachi, I., Kennedy, B. P., & Lochner, K. (1997). Social Capital, Income Inequality and Mortality. *American Journal of Public Health* , 1491-1498.

Kennedy, P. (2008). *A Guide to Econometrics 6E.* Wiley-Blackwell.

Miller, J. R., T, P., Ahern, J., Tracy, M., Tardiff, K., Vlahov, D., et al. (2005). Income inequality and risk of suicide in New York City neighbourhoods: A multilvel case-control study. *Suicide Life-Threatening Behaviour 35 (4)* , pp. 448–459.

Morgenstem, H. (1982). Uses of ecologic analysis in epidemiologic research. *American Journal of Public Health* , 1336-44.

OECD. (2006). *Divided We Stand: Why Inequality Keeps Rising. Country Note: Sweden.* Retrieved March 17, 2014, from http://www.oecd.org/social/soc/49564868.pdf

Ramstedt, M. (2001). Alcohol and suicide in 14 European countries. *Addiction 96(1)* , 59-75.

Socialstyrelsen. (2006, December). *Förslag till nationellt program för suicideprevention, Artikelnr 2006-107-23.* Retrieved April 2014, from Socialstyrelsen: www.socialstyrelsen.se

Stata. (2013). *Base Reference Manual.* Stata Press.

Wagstaff, A., & van Doorslaer, E. (2000). Income Inequality and Health: What does the Literature Tell Us? *Annual Rev. Public Health 21* , pp. 543-567.

Wilkinson, R. G. (1997). Health inequalities: relative or absolute material standards? *British Medical Journal 314 (7080)* , 591-595.

Wilkinson, R., & Picket, K. (2009). *The Spirit Level - Why Equality is Better for Everyone.* London: Penguin Books.