The effect of maternal alcohol consumption on the birth weight of babies

Author: Fanni Mercédesz Schipek
SNR: 2092268

Supervisor: Dr. Mery Ferrando

Tilburg School of Economics and Management (TiSEM)

Master's Thesis

MSc Economics

June 20, 2023

10452 words

Abstract

This paper aims to find empirical evidence on the effect of maternal alcohol consumption on the birth weight of babies. Three different dependent variables are employed to capture the effect: birth weight in grams, percentage of babies born with low birth weight and percentage of babies born with very low birth weight. Two staggered difference-in-differences models were implemented to estimate the effect. In the intention-to-treat estimation, changes in alcohol tax rates showed a 0.044 increase in the number of drinks consumed per week and this increase was significant. In the main regressions, alcohol tax is used as a proxy for maternal alcohol consumption. The main results found no evidence on the effect of maternal alcohol consumption on birth weights in grams and in the percentage of babies born with low birth weight, however, the model estimated a statistically significant decrease in the proportion of babies born with very low birth weight. The study had several limitations, thus it is important to consider the results with caution.

Contents

List of Tables

1 Introduction

In the United States, the prevalence of alcohol consumption among pregnant women is estimated to be approximately 12% (range 10.2%-16.2%), according to the Centers for Disease Control and Prevention (US). Furthermore, they measured that more than 50% of non-pregnant women drink alcohol occasionally (Denny et al., 2009). As alcohol use became more popular among women of childbearing age, it can easily lead to unintentional fetal alcohol exposure during the early gestational period (Pruett et al., 2013).

Alcohol exposure during the prenatal period can lead to devastating consequences, commonly referred to as fetal alcohol effects. The most serious consequence of fetal alcohol exposure, fetal alcohol syndrome, was first reported in the United States in 1973 by Jones and Smith. Fetal alcohol syndrome (FAS) is a condition that can develop in a fetus when a pregnant woman drinks alcohol during the gestational period. This syndrome can cause a group of symptoms that happen together as the result of a particular disease or abnormal condition. First in the literature Jones and Smith (1973) characterized a range of distinct abnormalities in children born to women who drank heavily during pregnancy. Furthermore, Pruett et al. (2013) mention several other consequences that can be attributed to maternal alcohol consumption such as abnormalities of organ systems, behavioural and intellectual deficits, and fetal death.

A baby's birth weight is a commonly used measure for their health and development. Low birth weight is considered to be less than 2500 grams, while babies born with less than 1500 grams are diagnosed with very low birth weight. Low birth weight can be caused by several factors, such as premature birth or intrauterine growth restriction (IUGR), which happens when a baby does not grow well during pregnancy due to problems with the placenta, the mother's health, or the baby's condition (Children's Hospital of Philadelphia, 2023).

Addictive activities and habits can affect the mother's health condition, such as smoking or drinking. Several studies have analysed the effect of maternal smoking and drinking during pregnancy on the abnormal conditions of the newborn. However, the evidence is mixed. In most cases, authors do not find a clear effect. The most common limitations of previous studies were recall bias/interviewer bias (Little et al., 1980; Verkerk et al., 1993; Windham et al., 1995; Mariscal et al., 2006), incomplete control for confounders (Olsen et al., 1983; Windham et al., 1995), low sample size or not representative sample (Miyake et al., 2014; Nykjaer et al., 2014) and low statistical power (Miyake et al., 2014).

However, in previous studies, multivariate linear and logistic regressions are the most commonly employed methods next to the instrumental variable research design. In my thesis, I implement a new methodology to examine the effect of maternal alcohol consumption on the birth weight of newborns: a difference-in-differences approach. I estimate two models: first, I calculate the effect of change in state alcohol taxes on maternal alcohol consumption and then, I estimate the effect of change in state alcohol taxes on birth weight. In the second estimation, the alcohol tax is used as a proxy for maternal alcohol consumption. Compared to the previous instrumental variable studies, these models do not estimate the classic two-stage instrumental variable regressions, they are calculated by Callaway & Sant'Anna's difference-in-differences estimation method. To measure birth weight, I use three different variables: the birth weight of babies in grams, the percentage of babies born with low birth weight and the percentage of babies born with very low birth weight.

To conclude, my research focuses on two questions. The first research question aims to provide evidence on the effect of maternal alcohol consumption on the birth weight of babies. In addition, the second research question examines the effect of maternal alcohol consumption on the proportion of babies born with low or very low birth weight. My hypothesis for the first research question is that prenatal alcohol consumption has a negative effect on birth weight, thus I expect decreasing birth weight with an increase in alcohol drinking. For the second research question, I assume that an increase in maternal alcohol consumption would lead to an increase in the proportion of babies born with (very) low birth weight.

The organization of this paper is as follows: Section 2 presents a literature review on the historical background of studying alcohol exposure as well as empirical evidence, Section 3 introduces data and data sources, Section 4 defines the empirical methodology employed for data analysis, Section 5 presents and interprets the results of the analysis, while Section 6 provides a conclusion and discussion of the study.

2 Literature review

2.1 Historical background

The study of alcohol exposure in utero started in the early 1970s. In 1973, two studies were published that first described the association between prenatal alcohol exposure and birth defects in the medical community (Jones and Smith, 1973; Jones et al., 1973). Several years later other studies showed that the detrimental effects of alcohol on fetuses had been known for centuries (Erb and Andresen, 1978; Clarren and Smith, 1978; Tenbrinck and Buchin, 1975). The earliest references to alcohol intoxication resulting in a damaged newborn date back to ancient Greek and Roman times (Green, 1974).

The beliefs about prenatal alcohol exposure have changed over the centuries. In ancient times, they believed that the deformity of newborns was caused by intoxication at the moment of conception rather than alcohol consumption during the gestational period (Calhoun and Warren, 2007). In the 1700s, the Royal College of Physicians of London (1726) stated that the offspring of alcoholic women are "*weak, feeble, and distempered*". At the end of the 19th century, a British medical officer, William Sullivan, examined children of imprisoned alcoholic mothers. Sullivan (1899) found that alcohol had a direct toxic effect on embryos (Golden, 2009). Until the late 1900s, no other significant medical studies were published in this association (Calhoun and Warren, 2007).

From the 1970s, a large number of medical case studies and articles focused on the relationship between maternal behaviour and the health conditions of newborns. According to these reports, there is a correlation between the extent of defects in children and the severity of maternal alcohol consumption, and factors such as malnutrition, smoking, and inadequate prenatal care can also contribute to the development of morphological differences (Majewski, 1981; Bingol et al., 1987). Later on, doctors also recognized that the physical and neurobehavioral outcomes of maternal drinking were variable, ranging from the classic form to a few minor anomalies (Calhoun and Warren, 2007).

Jones and Smith (1973) diagnosed a new disease and introduced a new term in the medical literature which is related to prenatal alcohol exposure: fetal alcohol syndrome. Fetal alcohol syndrome (FAS) is a birth defect that occurs when a person was exposed to alcohol before birth. Authors recognised several anomalies that are connected to prenatal alcohol exposure, such as developmental delay, prenatal and postnatal growth deficiency, short palpebral fissures, epicanthal folds, small jaws and flattened midface, joint anomalies, and altered palmer crease patterns (Jones et al., 1973). The diagnoses of fetal alcohol syndrome have been expanded and clarified over the years, however, all diagnostic schemas lean on defects in three distinct areas: prenatal and postnatal growth deficiency, central nervous system dysfunction, and facial anomalies (Riley et al., 2011).

Riley et al. (2011) summarised the facts that have become clear over the years regarding alcohol exposure in utero. First of all, alcohol is a teratogen that may cause damage to the developing embryo and fetus. Secondly, the range of outcomes associated with gestational alcohol exposure is broader than with fetal alcohol syndrome. Thirdly, the most significant consequences of gestational alcohol exposure are those on cognitive and behavioural outcomes and brain development. Lastly, prenatal alcohol consumption is a major public health issue with profound effects at a staggering economic cost.

2.2 Empirical evidence of maternal alcohol consumption on infant birth weight

The first empirical analyses in this association date back to the 1970s. Little (1977) examined the effect of moderate maternal alcohol consumption on infant birth weight in Seattle, Washington. She found a significant correlation relationship between maternal drinking before pregnancy and in late pregnancy and infant birth weight. Her study showed that one ounce of alcohol before pregnancy was associated with 91 grams decrease in birth weight, while one ounce consumed in late pregnancy was associated with 160 grams decrease. However, the results of this study are not completely representative as the sample was taken from a distinctive population of women who were members of a health maintenance organization. Furthermore, the sample was small and robust results were not calculated.

Later, Little et al. (1980) analysed whether the history of maternal alcoholism is related to birth weight, even if the mother refrains from consuming alcohol during the gestational period. They interviewed three cohort groups: the abstinent alcoholics, the drinking alcoholics, and the controls. They compared both alcoholics groups to the non-alcoholic control group and found that in the case of abstinent alcoholics there is a decrease of 258 grams in mean birth weight and an additional 235 grams of decrease for drinking alcoholic women. They concluded that both the history of maternal alcohol consumption and heavy drinking may have a risk to the optimal development of the fetus. This result shows a correlation between maternal alcohol consumption and birth weight.

Based on the previous studies by Little et al. (1977, 1980), Olsen et al. (1983) conducted a study with a large sample of Danish women to re-examine the effect of alcohol consumption and birth weight. Their study also investigated the relationship between conception time and the weight and length of newborn babies. Their findings are similar to the results of previous papers, these indicate that alcohol drinking, even in low consumption, can have an effect on birth weight. However, this study has several limitations. Firstly, health implications cannot be assessed from these results as only babies with good general health conditions were included in the study, secondly, the relation between alcohol and low birth weight might be confounded.

Verkerk et al. (1993) performed a study in the same association in the Netherlands. They believed that their study had higher statistical power than many previous ones to detect a possible effect of alcohol consumption on pregnancy outcomes because of numerous reasons. They collected data from a low-risk population, which limited random variation in outcome measures. Furthermore, their sample size was large with a high percentage of alcoholic mothers, and alcohol use was measured for each semester. Their results showed no significant relationship between alcohol use in general and a decrease in birth weight and only a limited effect of moderate alcohol consumption on birth weight. However, they found a significant relationship in this association in the subgroup of women who smoked 20 or more cigarettes per day. A possible reason for this finding is the residual confounding of smoking. Other studies (Wright et al., 1983; Olsen et al., 1991) showed a synergistic effect of alcohol and smoking and the risk of delivering a baby with low birth weight.

At the beginning of the 2000s, the effects of moderate alcohol consumption on fetuses remained unclear and evidence was mixed. Mariscal et al. (2006) carried out a case-control study in Spain to examine the effect of different patterns of alcohol drinking and interaction with smoking during the gestational period on low birth weight. Their study showed a slightly synergistic effect of tobacco smoking and alcohol drinking on risk for low birth weight, especially for small-for-gestational-age babies. They found a positive relationship between weekday drinkers consuming at least 12 g/day and low birth weight, however, for those babies whose mothers were only weekend drinkers, alcohol consumption showed a borderline protective relationship with low birth weight.

Chen (2012) conducted a study that aimed to address the methodological limitations in the literature. He used a sibling fixed-effects model to adjust for unobserved heterogeneity among the mothers in the sample. He analysed the relation between alcohol and birth weight and the effect of alcohol on infant behavioural outcomes (such as positive mood, fearfulness, and difficultness). His results suggest that prenatal alcohol consumption is a risk factor for infant behavioural outcomes, but behavioural outcomes seem to be more affected by alcohol use than birth weight.

Two British cohort studies (McCarthy et al., 2013; Nykjaer et al., 2014) focused on the timing of alcohol exposure. McCarthy et al. (2013) highlighted the importance of timing and degree of alcohol exposure and they collected exceptionally detailed data to examine the effect. As a result, they found no significant association between occasional and low amounts of alcohol and birth weight. However, these results contradict the findings of Nykjaer et al. (2014). Their data showed that low levels of alcohol exposure, especially in the first trimester, have a negative effect on fetal growth and alcohol also increases the odds of babies being born preterm.

Patra et al. (2011) conducted a systematic review and meta-analyses and revealed that heavy prenatal alcohol consumption increases the risk of low birth weight, whereas light and moderate drinking show no effect. This analysis, however, did not contain data from Asian countries, therefore Miyake et al. (2014) presented the first epidemiological study on this association in Japan. No evidence was found in the relationship between maternal alcohol consumption and low birth weight, but their study had several limitations. Their subjects were not representative of the whole Japanese population and their study did not have significant statistical power.

Studies have shown conflicting results in the association between prenatal alcohol use and birth weight so far. While several American studies showed a significant relationship between alcohol consumption and low birth weight, other studies conducted in Europe and Asia did not show clear and consistent results. Nevertheless, most studies had either methodological or data limitations. Considering these facts, I will use a different methodology and a much larger database to examine the association in question. However, there is a chance that I may not find an effect similar to previous studies.

3 Data

This empirical research is based on secondary data and Table 1 contains the sources for the main variables of interest. Both research questions require the same data and resources. An 8-year time period (1989-1996) is used to conduct the analysis.

Table 1: Data sources

Variable	**Source**
Birth weight	Natality Detail Files (1989-1996)
Maternal alcohol consumption	Natality Detail Files (1989-1996)
State alcohol beverage taxes	TPC analysis of state revenue department and legislative websites (accessed January 27, 2022); Federation of Tax Administrators, the Tax Foundation, the Council of State Governments, the Advisory Commission on Intergovernmental Relations, the Distilled Spirits Council of the United States.
Demographic and socioeconomic status variables	Natality Detail Files (1989-1996)

3.1 Natality Detail Files

Natality Detail Files provide information on live births in the United States. Live birth is defined by World Health Organization as "*every product of conception that gives a sign of life after birth, regardless of the length of the pregnancy*" (ICPSR, 1996).

These databases are cross-sectional, recorded yearly, and contain one hundred per cent of birth certificates (at the individual level) in the given year. As the records are collected from birth certificates, most of the information can be considered to be reported correctly and accurately. The data is limited to births that occurred within the United States to all American residents and non-residents. The analysis uses Natality Detail Files between the calendar years 1989 and 1996.

Natality Detail Files include information regarding birth outcomes, such as birth weight, period of gestation, sex of infant, and possible abnormal conditions of newborns. Furthermore, they also contain information on demographic details, such as the mother's age and race, father's age and race as well as maternal alcohol consumption and smoking during pregnancy. Natality Detail Files do not contain detailed information on socioeconomic status, however, the mother's education and marital status and the father's education can serve as a proxy for socioeconomic status.

3.2 Outcome variable: Birth weight

This study focuses on whether maternal alcohol consumption has an effect on birth weight and causes abnormal conditions in the newborns, such as low birth weight or very low birth weight.

The birth weight of the newborns is directly collected from U.S. Birth certificates. To maintain consistency, Natality Detail Files contain birth weight in grams. To define low birth weight, I use the definition used in the Ninth Revision of the International Classification of Diseases (IDC-9). IDC-9 specifies low birth weight as less than 2500 grams. Furthermore, very low birth weight is considered to be less than 1500 grams. In the analysis, I employ three outcome variables. "Birth weight" is a continous variable and measured in grams. "Low birth weight" is a binary variable, it takes the value of 1 if the baby has a birth weight lower than 2500 grams, 0 otherwise. "Very low birth weight" is a binary variable, it takes the value of 1 if the baby has a birth weight lower than 1500 grams, 0 otherwise.

3.3 Maternal alcohol consumption

Birth certificates in the United States contain a checkbox to classify a mother as a drinker during pregnancy and show the average number of drinks consumed per week. When drinking status is not reported or is not consistent with the number of drinks reported, the status is modified to be consistent. The mother is categorized as a drinker, for example, if the drinking status is given as "no" but at least one drink a week is documented. A mother is classified as a drinker when one drink per week is reported. If the number of drinks is recorded as a span (e.g. 4-5 drinks per week), then the lower number is used.

The analysis used a continuous variable to measure maternal alcohol consumption. Variable "drink" shows the average number of drinks per week. Values between 0 and 97 show the average number of drinks per week, while 98 shows 98 or more drinks per week. This variable has unknown or not stated values, those are considered to be missing values (approximately 4.6 per cent of cases).

The certificates of 48 states and the District of Columbia included information on alcohol use. South Dakota and California did not include information on alcohol consumption on their birth certificates, thus these states are not included in the analysis.

The accuracy of this data should be considered with limitations. The birth certificate question regarding alcohol use has no time reference and does not encourage trustworthiness in reporting very low alcohol consumption. However, based on the literature, very low alcohol consumption does not have an effect on the abnormal conditions of newborns. But, in the case of moderate and high alcohol consumption, the records can also be underestimated because of recalling bias.

3.4 State alcohol beverage taxes

State alcohol excise tax rates are collected from the webpage of the Tax Policy Center (TPC) Urban Institute & Brookings Institution. The sources of the data collection are the TPC analysis of state revenue department and legislative websites (accessed January 27, 2022); Federation of Tax Administrators, the Tax Foundation, the Council of State Governments, the Advisory Commission on Intergovernmental Relations, the Distilled Spirits Council of the United States.

In general, the state beer tax is a Specific Excise Tax Per Gallon of 5% Alcohol, the state distilled spirits tax is a Specific Excise Tax Per Gallon of 40% Alcohol, and the state wine tax is a Specific Excise Tax Per Gallon for 12% Alcohol. All types of alcohol taxes are measured in dollars and determined at the state level.

Tax Policy Center used the following criteria for collecting tax rates in special circumstances. For states with an additional state-wide local tax rate, local tax rates are included. When there are multiple rates in a given year, they used the rate as of January 1st. When there is more than one tax rate per type of alcohol, the following criteria are used to select a single rate: for beer, the rate for beer over 3.2% alcohol by volume (ABV) is selected, and they select bottled and canned over barrel and keg. For wine, the tax rate assigned to still as opposed to sparkling wine is chosen, and the rate for the lowest ABV is selected. For spirits, they choose the rate associated with 50% ABV or less. Their database contained "n.a." for states where state governments directly regulate the sales of wine and distilled spirits distilled. These states are known as monopolies or control states.

3.5 Control variables

Natality Detail Files report the possible demographic and socioeconomic status variables that can be used as control variables. The demographic variables are represented by age and race of the mother, age and race of the father, and sex of the infant, while socioeconomic status is measured by the education level of the mother and marital status of the mother. The education level of the father is omitted from the analysis as Natality Detail Files do not include this variable after 1994.

The age of the mother is represented by a variable named "Mother age" and measured in single years. This variable takes values between 10 and 49 because it is edited for lower and upper limits. When the mother's age is recorded to be under 10 or 50 years or over, it is not taken into consideration and stated as a missing value. Similarly, the age of the father is named "Father age" and is measured in single years. This variable takes values between 10 and 98 and if the age is reported to be under 10, it is considered not stated.

The race of the mother and the race of the father are represented in variables named "Mother race" and "Father race" respectively. These variables take 3 different values: white, black, and other. In the "other" category the following races are included: American Indian (includes Aleuts and Eskimos), Chinese, Japanese, Hawaiian (includes part-Hawaiian), Filipino, Other Asian or Pacific Islander, and Guamanian. The category white contains births reported as white and Hispanic.

The sex of an infant is a dummy variable, named "Sex". It has only two categories: male and female. It takes the value of 1 if the infant is female, 0 otherwise.

The education of the mother is a continuous variable, named "Mother education". It takes a value between 0 and 17 and the numbers show the total number of years completed in education. Only years spent in "regular" schools—public schools or its equivalent in recognized private or parochial institutions—are taken into consideration. Trade or business schools are not classified as "regular" schools. People who have finished only a partial year in high school or college are counted as having completed the highest preceding grade. If the mother has no formal education, this variable takes the value of 0. Values between 1 and 8 show the years spent in elementary school, while values between 9 and 12 determine the additional years spent in high school (after elementary school). If the variable takes a value of 13 or above, it shows the number of years spent in college. When the education level has a status of not stated or is not compatible with this coding specification, it is treated as missing.

The marital status of the mother is a dummy variable, referred to as "Marital status". It has two categories: married or unmarried. The variable takes the value of 1 if the mother is married, 0 otherwise. This is not reported by five states, however, in these cases, marital status is derived from the comparison of the child's and parent's surnames. The accuracy of this variable in those states is not completely clear as spouses are entitled to keep their married names after divorce in the United States.

3.6 Treatment and control group identification

Natality Detail Files give information about live births in all states of the United States of America. This research primarily focuses on states where there was a change only once in alcohol beverage taxes and on states where there was no change in alcohol taxes in the period under review. Changes in alcohol beverage taxes are taken into consideration in states where all kinds of alcohol taxes are changed at the same time. Based on the data collected by Tax Policy Center (TPC), four states changed their alcohol taxes once between 1989 and 1996, and 40 states did not have a change in alcohol taxes. In the remaining 7 states, there were changes more times during this period or there was a change in one of the three alcohol types only. To be consistent, I left out these states from the analysis. Furthermore, South Dakota and California are also omitted from the analysis due to missing information on maternal alcohol consumption.

As can be seen, 42 states are involved in total. Out of 42 states, the following 3 states are included in the Treatment group: Connecticut, Delaware, and Rhode Island. California would have been the fourth state in the Treatment group. The remaining 39 states form the Control group: Alabama, Alaska, Arizona, Arkansas, Colorado, District of Columbia, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, North Carolina, North Dakota, Oklahoma, Oregon, Pennsylvania, South Carolina, South Dakota, Tennessee, Texas, Utah, Vermont, Virginia, West Virginia, and Wyoming.

After the identification of treatment and control groups, I revised the data collected by Tax Policy Center (TPC). In the case of multiple rates in a given year, TPC used the tax rates as of January 1st. As this would result in incorrect estimates, I used the exact months when the policy was introduced. In Connecticut, the new policy was implemented on 23 March 1989 (Lohman and Pinho, 2011), in Delaware, the taxes changed on 01 September 1990 (Delaware Division of Revenue, 2019), while in Rhode Island, the tax rates are effective from 30 June 1989 (Rhode Island General Laws, 1989). From Natality Detail Files I used the birth date (year and month) of newborns to capture time-fixed effects and to ensure consistency, I identified the start months of treatment by setting them 9 months after the policy changes. The rationale behind this decision is that in Natality Detail Files mothers reported drinking during the prenatal period and babies conceived after the tax change were typically born 9 months later. Based on the previously mentioned reasons, the start dates of treatment are determined as follows: January 1990 for Connecticut, June 1991 for Delaware, and April 1990 for Rhode Island.

3.7 Descriptive statistics

Table 2 and Table 3 show the descriptive statistics for both research questions. The tables provide information on observable characteristics between the treatment and the control groups in the period before the first state became treated. The birth weight of the babies is on average higher in the treatment group and the difference in means is significant based on a result of a t-test (t=7.62). The proportion of babies with low birth weight (less than 2500 grams) is higher in the control group, but the proportion of babies with very low birth weight (less than 1500 grams) is lower in the control group. Furthermore, as can be seen, the proportion of drinkers is moderately higher in the control group as well as the number of drinks consumed per week. All the differences in means are significantly different from zero at all significance levels, except for the proportion of babies with very low birth weights. This means that there could be pre-existing differences between the treatment and the control groups that may impact the treatment effect.

In the case of demographic variables, greater differences can be seen between the two groups. In treatment states the average age of mothers and fathers are higher and these differences are significant. However, 78% of observations for the "Mother age" variable and 18% of observations for the "Father age" variable are missing. Thus, the inferences should be treated with caution about the significance of mean differences. The proportion of white mothers is also higher in the treatment group and as a result, the proportion of black mothers is lower than in the control group. Furthermore, the proportion of female infants is approximately the same in both groups. Taking the socioeconomic status variables into consideration, the result is similar to the previously mentioned variables. The proportion of married mothers is slightly lower in the treatment group and mothers are higher educated on average. These differences also reflect on the possible presence of pre-existing differences between the treatment and the control groups.

Table 3 contains the rates of drinking participation. In most cases, the rate of drinking participation is higher in the control group than in the treatment group. In the education variable, the proportions slightly increase from no education to college for the control group and the participation rate for drinking alcohol is the highest among mothers who went to high school in the treatment group. The rate of drinking participation is the highest among black women for the treatment group and the highest among white women for the control group. Furthermore, it can be seen that the participation rate among unmarried women is also higher on average than among married women for both groups.

Table 2: Descriptive statistics I.

	Treatment	Control	t-test	p-value
$Birth\ weight$	3351.32	3334.05	7.62	0.000
	(615.58)	(607.83)		
$<= 2500\ grams$	0.0697	0.0728	-3.30	0.001
	(0.25)	(0.26)		
$<= 1500\ grams$	0.0142	0.0134	1.94	0.051
	(0.12)	(0.11)		
$Drinker$	0.0360	0.0409	-6.47	0.000
	(0.19)	(0.20)		
$Drinks\ per\ week$	0.0585	0.0706	-3.88	0.000
	(0.76)	(0.82)		
$Mother's\ age$	26.89	25.80	9.35	0.000
	(5.53)	(5.61)		
$Father's\ age$	30.44	29.43	41.77	0.000
	(6.05)	(6.65)		

Table 3: Descriptive statistics II.

				Rate of drinking participation	
		Treatment	Control	Treatment	Control
Education of mother	*No education*	0.0011	0.0014	0.0000	0.0163
	Elementary school	0.0270	0.0416	0.0248	0.0245
	High school	0.4481	0.5538	0.0307	0.0351
	College	0.4298	0.3641	0.0285	0.0406
Race of mother	*White*	0.8433	0.7820	0.0278	0.0363
	Black	0.1326	0.1864	0.0377	0.0336
	Other	0.0241	0.0316	0.0174	0.0317
Marital status	*Married*	0.7367	0.7388	0.0248	0.0319
	Unmarried	0.2633	0.2612	0.0403	0.0462
Sex of child	*Female*	0.4872	0.4878	0.0281	0.0354
	Male	0.5128	0.5122	0.0296	0.0359

4 Methodology

In this section, I present the empirical strategy which will be used to examine the effect of maternal alcohol consumption on the abnormal conditions of newborns. In the literature, multivariate linear regression or multivariate logistic regression analyses are the most commonly used techniques to diagnose the effect of maternal alcohol consumption during pregnancy on abnormalities of newborns. However, these methods have several limitations, such as inadequate control for confounding factors.

Another prevalent research method for analysing the effect of alcohol consumption on an outcome variable is the instrumental variable research design. Previous studies have used different instruments to infer the effect of parental addictive behaviour on birth outcomes. Wehby et al. (2011) estimated the effect of prenatal smoking on birth weight with genetic instruments, while other studies have relied on other instruments, such as changes in cigarette taxes, prices, or smoking laws (Wehby and Scholder, 2013).

In this research, I would like to estimate a difference-in-differences model that has not been used before – to my best knowledge – to analyse these questions. The main regression equation calculates the effect of change in state alcohol taxes on birth outcomes. In this context, alcohol tax is used as a proxy for maternal alcohol consumption. In addition, I calculate the effect of change in state alcohol taxes on maternal alcohol consumption as an intention-to-treat estimation. Compared to the previous instrumental variable studies, this research does not estimate the classic two-stage IV regressions, it calculates difference-in-differences models while using alcohol taxes as an intention-to-treat effect. Difference-in-differences approach is an adequate estimation methodology for the given research question as the main goal is to estimate the causal effect of a policy, change in alcohol taxes. Furthermore, there are two distinct groups, treatment and control as well as pre-treatment and post-treatment periods. Moreover, there is time variation in the treatment effect, inducing the estimation of a staggered difference-in-differences model. The validity of using alcohol tax as a proxy for maternal alcohol consumption is supported by the following. First, the own-price elasticity of all alcoholic beverages is negative (-0.51), meaning that a 1% increase in alcohol prices leads to a 0.5% decrease in alcohol consumption based on a meta-analysis (Wagenaar et al., 2009). This meta-analysis also calculated the own-price elasticity separately for specific alcoholic beverages which was -0.46 for beer, -0.69 for wine and -0.80 for distilled spirits. These results suggest that a tax increase leads to a decrease in alcohol consumption. Second, Daley et al. (2012) examined the effect of a 25-cent-per-drink alcohol tax increase on drinking behaviour. They found that tax increase results in a 9.2% reduction in alcohol consumption, including an 11.4% reduction among heavy drinkers. Furthermore, Zhang (2010) conducted a correlation study and ran an OLS estimation on the relationship between tax changes and birth outcomes and found that a one-cent increase in beer taxes induces a 1-2 percentage points decrease in low-birth-weight.

4.1 Difference-in-differences

4.1.1 Two-way fixed effects

The simplest form of the difference-in-differences approach is the estimation of two-way fixed effects. It aims to control for differences between groups and time (Huntington-Klein, 2021). The regression equation is the following:

$$Y_{ist} = \alpha + \beta(Treated_s X Post_t) + \gamma Treated_s + \delta Post_t + \epsilon_{ist} \tag{1}$$

where γ shows the group-fixed effect, while δ shows the time- fixed effect. In the simplest form, there are only two groups: "treated" and "untreated" and two time periods: "before treatment" and "after treatment". The variable "Treated" is binary and it takes the value of 1 if you are in the treatment group and 0 otherwise. Similarly, the variable "Post" is a binary variable, it takes the value of 1 for the treatment period and 0 otherwise. The estimated difference-in-differences effect is shown by β (Huntington-Klein, 2021).

The two-way fixed effects method assumes two criteria: parallel trends assumption and no other policies pursued around the same time. Parallel trends assumption in a difference-in-differences design means that "if no treatment had occurred, the difference between the treated group and the untreated group would have stayed the same in the post-treatment period as it was in the pre-treatment period" (Huntington-Klein, 2021). Unfortunately, there are no formal tests for parallel trends assumption as it relies on a counterfactual observation. Huntington-Klein (2021) suggests two tests – a test of prior trends and a placebo test – that can provide some evidence to make parallel trends assumption plausible. If covariate-specific trends are present, ignoring them would lead to biases when interpreting causal effects (Callaway and Sant'Anna, 2021).

Two-way fixed effects is excessively intuitive when we need to control for group and time differences. It also allows the involvement of multiple groups instead of just one treated and one untreated group and many time periods to estimate long-term effects. However, it has a really important drawback: it does not work properly with staggered treatment timing. If there is time variation in applying treatment between groups, the two-way fixed effects is not a good estimator for capturing treatment effects (Huntington-Klein, 2021).

4.1.2 Rollout Design and Multiple Treatment Periods

A rollout design is suitable to estimate a difference-in-differences model with multiple groups in the "treated" group while assuming that the groups get treated at different times (Huntington-Klein, 2021). In this case, the "before treatment" and "after treatment" periods are not the same for all groups, and the two-way fixed effects regression is not an appropriate design anymore. If the treatment effects evolve, the "negative weight problem" can arise by using two-way fixed effects estimation (Goodman-Bacon, 2021). This means that, in general, β of the two-way fixed effects estimation represents the weighted average of different treatment effect parameters, however, some of these weights can be negative (Callaway and Sant'Anna, 2021). This may result in particularly problematic cases, e.g. the treatment effect is positive for all units, but the estimated βs of the two-way fixed effects estimation are negative (Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021). In addition, even when the weights are not negative, they are determined by the two-way fixed effects estimation approach and they are sensitive to factors such as the group sizes, the timing of treatment and the overall number of time periods (Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021). Application of two-way fixed effects estimation method with multiple treatment periods leads to using already-treated groups as an untreated group and having a continuously-treated group as a comparison causes parallel trends assumption to break (Huntington-Klein, 2021). If we do an analysis of a balanced panel without using control variables, Goodman-Bacon decomposition can show how much of a problem is caused by using two-way fixed effects in a difference-in-differences model (Huntington-Klein, 2021).

4.1.3 Callaway and Sant'Anna's estimation method

Callaway and Sant'Anna (2021) identify procedures for estimating treatment effect parameters using difference-in-differences when the following 3 specifications hold: first, when there are multiple time periods, second, when there is a variation in treatment timing, and third, when the "parallel trends assumption" may only be valid after conditioning on observed covariates. Their method is suitable for estimating causal parameters that allow for arbitrary treatment heterogeneity and dynamic effects. Furthermore, they avoid the possibility of interpreting the results of the classic two-way fixed effects regressions as causal effects in difference-in-differences frameworks.

Contrary to the canonical difference-in-differences method where causal parameters are reduced to the ATT, Callaway and Sant'Anna (2021) focus on the disaggregated causal parameter called as group-time average treatment effect. The group-time average treatment effect is the average treatment effect for group g at time t and the group is defined by the period when the units are first treated (Callaway and Sant'Anna, 2021). Group-time average treatment effect parameters have a really important advantage: they do not limit heterogeneity in terms of observed covariates, the time period in which units are first treated, or the development of treatment effects over time (Callaway and Sant'Anna, 2021). As a result, these causal parameters are appropriate to directly examine treatment effect heterogeneity and/or to build other aggregated causal parameters (Callaway and Sant'Anna, 2021).

Callaway and Sant'Anna (2021) propose several conditions that are related to treatment anticipation behaviour and conditional parallel trends assumption under which the group-time average treatment effects are non-parametrically point-identified. This framework makes researchers possible to involve covariates in a staggered difference-in-differences model with multiple groups and multiple time periods. Their concept consists of six assumptions. Their first assumption claims that at time t=1, nobody is treated, and once a unit is treated, that unit will remain treated in the following period (Callaway and Sant'Anna, 2021). The second assumption implies random sampling and states that the variables of interest are independent and identically distributed across individuals. The third assumption restricts treatment anticipation for all "eventually treated" groups (Callaway and Sant'Anna, 2021). They state that "no-anticipation" is assumed when the treatment path is not known in advance and/or when units do not "choose" their treatment status (Callaway and Sant'Anna, 2021). Assumptions 4 and 5 generalize the parallel trend assumption for multiple time periods and multiple groups. These assumptions are conditional on covariates. The main difference between the two assumptions is that one is based on a "never-treated" control group while the other one is based on a "not-yet-treated" control group. My study focuses on the staggered difference-in-differences method with a "never-treated" control group. In the case of a "never-treated" control group, the parallel trends assumption is interpreted as the following: in the absence of treatment, the average outcome for the group that became treated in period g and for the "never-treated" control group would have followed parallel trends, conditional on covariates (Callaway and Sant'Anna, 2021). It is important to note that covariate-specific trends are allowed and the relationship between treatment timing and the outcomes is not restricted (Callaway and Sant'Anna, 2021). The last one is the overlap assumption which says that a positive fraction of the population becomes treated in period g and the generalized propensity score is uniformly bounded away from one for all g and t (Callaway and Sant'Anna, 2021).

Callaway and Sant'Anna (2021) propose three different estimation methods that can be adapted to staggered difference-in-differences models: the outcome regression (OR), the inverse probability weighting (IPW), and the doubly-robust (DR) method. In the outcome regression method, it is required to appropriately model the comparison group's outcome evolution to estimate the group-time average treatment effects (Callaway and Sant'Anna, 2021). The inverse probability weighting approach is completely different from the outcome regression method. In IPW, it is required to properly model the conditional probability of unit i being in group g given their covariates X and unit i should be either in group g or in a relevant comparison group (Callaway and Sant'Anna, 2021). The doubly-robust method is the combination of outcome regression and inverse probability weighting. In DR, it is only necessary to specify either the comparison group's outcome evolution or the propensity score model (Callaway and Sant'Anna, 2021; Sant'Anna and Zhao, 2020). The doubly-robust method has one advantage compared to the other two approaches: it has additional robustness against model misspecifications (Callaway and Sant'Anna, 2021).

4.2 Empirical models

As the treatment timing is different between groups, Callaway & Sant'Anna's estimation method is appropriate to calculate the average treatment effect. The main regression equation is the following:

$$Y_{it} = \alpha + \beta Treated_{it} + \gamma X_i + \lambda_t + \mu_i + \epsilon_{it} \tag{2}$$

where the outcome variable (Y_{it}) is represented by three different variables: (1) the birth weight of the newborns in state i and month t, (2) the proportion of babies born with low birth weight in state i and month t and (3) the proportion of babies born with very low birth weight in state i and month t. $Treated_{it}$ is a binary variable for states i that changed the tax rates in month t and

β shows the difference in birth weight between Treated and Untreated states after the change in alcohol tax rates relative to the pre-treatment period. The year-fixed effect is captured by λ_t and the state-fixed effect is set by μ_i. Moreover, γX_i shows the vector of control variables and ϵ_{it} represents the error term. Standard errors are computed by using the multiplier-type bootstrap method proposed by Callaway and Sant'Anna (2021).

Before estimating the main regression of interest, I calculate an intention-to-treat model, which estimates the effect of change in state alcohol taxes on maternal alcohol consumption. The following regression equation captures the effect:

$$Y_{it} = \alpha + \beta Treated_{it} + \gamma X_i + \lambda_t + \mu_i + \epsilon_{it} \tag{3}$$

where the outcome variable is maternal alcohol consumption in state i and month t, measured in drinks per week. β shows the difference in maternal alcohol consumption between Treated and Untreated states after the change in alcohol tax rates. All other terms are the same as in the main regression equation.

After estimating group-time average treatment effects, it is useful to check pre-treatment trends to assess the validity of parallel trends assumption. To capture this, an event-study analysis is performed with the following specifications:

$$Y_{it} = \alpha + \sum_{\tau=-28}^{-1} \beta_\tau Treated_{it} + \sum_{\tau=1}^{83} \beta_\tau Treated_{it} + \gamma X_i + \lambda_t + \mu_i + \epsilon_{it} \tag{4}$$

where event-time τ ranges from -28 to 83 and $\tau = 0$ is excluded as it is the last period before treatment. β_τ shows the difference in birth weight between Treated and Untreated states in period τ relative to the last period before the tax was changed. All other factors are the same as in the main regression equation.

5 Results

5.1 The effect of maternal alcohol consumption on birth weight

Before analysing the effect of tax increases on birth weight, it is useful to analyse the relationship between maternal alcohol consumption and birth weight. To see whether there is a correlation relationship between the variables, I estimated four models by ordinary least square, however, each model differs in terms of the included covariates. The first model is the simplest, it only contains one independent variable: maternal alcohol consumption. The second model includes the group of demographic variables: age and race of the mother, age and race of the father, and sex of the newborns. The third model adds the group of socioeconomic variables: marital status of the mother and education of the mother. The last model includes both demographic and socioeconomic status variables next to the main independent variable, except the variables that have a high proportion of missing values. The age of the mother variable is omitted because it has approximately 78% missing values, while the age and race of the father variables have approximately 17% missing values. Including variables with a high proportion of missing values significantly reduces the number of observations while estimating OLS regressions.

The results can be seen in Table 4. The estimations are significant at all significance levels. The results suggest that the estimated reduction in birth weight with one additional drink per week ranges between 12.31 and 21.38 grams on average. However, if Model 2 is excluded as it has significantly fewer observations, it can be concluded that the estimated reduction in birth weight with one additional drink per week ranges between 17.45 and 21.38 grams on average. It is important to note that adding covariates to the regression induces an increase in the estimate. This means that the coefficient is not completely stable and omitted variable bias can be present (Oster, 2019).

Comparing the results to the previous correlation studies, no clear conclusion can be drawn on the correlation relationship between maternal alcohol consumption and birth weight. The study by Little (1977) concluded that 28 grams of alcohol (one ounce) can reduce birth weight by 91-160 grams, while the study by Little et. al (1980) found 258-493 grams decrease in birth weight, depending on whether the mother is abstinent alcoholist or drinking alcoholist. However, in the meta-analysis, Patra et al. (2011) determined a cut-off value at which alcohol consumption may cause adverse effects on birth outcomes. This cut-off value is on average 7 to 10 alcoholic drinks per week, which is equal to 70 to 126 grams per week. In this study, the proportion of mothers drinking more than 10 drinks per week is 0.055%, which supports that the decrease is not that significant.

However, this study contains all live births for 8 years in the United States and represents the American population appropriately. Thus, it can be concluded from this population that there is a negative relationship between the average drinks consumed per week during pregnancy and the birth weight of the babies. There is one limitation that should be mentioned. The accuracy of reported alcohol consumption should be considered with caution and can cause measurement error in the estimations. The measurement error can bias the estimates towards zero, resulting in an underestimation of the relationship between variables.

Table 4: Simple OLS estimates

	Dependent variable: Birth weight			
	Model 1	Model 2	Model 3	Model 4
Alcohol consumption	-21.38^{***}	-12.31^{***}	-18.52^{***}	-17.45^{***}
	(0.17)	(0.93)	(0.18)	(0.17)
Mother age		5.32^{***}		
		(0.14)		
Mother race (other)		103.54^{***}		116.84^{***}
		(4.94)		(0.80)
Mother race (white)		176.83^{***}		215.24^{***}
		(4.23)		(0.37)
Father age		-0.05		
		(0.12)		
Father race (other)		-46.44^{***}		
		(4.86)		
Father race (white)		48.05^{***}		
		(4.03)		
Sex of infant		118.63^{***}		118.52^{***}
		(1.06)		(0.26)
Marital status (Unmarried)			-181.64^{***}	-106.98^{***}
			(0.31)	(0.33)
Mother education			14.20^{***}	15.01^{***}
			(0.06)	(0.05)
Constant	$3,328.64^{***}$	$2,978.85^{***}$	$3,203.08^{***}$	$2,936.94^{***}$
	(0.13)	(3.29)	(0.74)	(0.83)
Observations	20,509,430	1,192,933	20,297,619	20,297,619
R^2	0.001	0.028	0.028	0.054
Adjusted R^2	0.001	0.028	0.028	0.054
Residual Std. Error	605.93	577.03	596.38	588.50
F Statistic	$15,381.08^{***}$	$4,355.23^{***}$	$196,758.80^{***}$	$192,214.90^{***}$

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

5.2 Validity of assumptions of Callaway and Sant'Anna (2021)

To fully understand and interpret the results of equation 2 and equation 3, it is important to check whether the assumptions are valid and hold. Callaway and Sant'Anna (2021) proposed two main assumptions that should hold while estimating the average treatment effect: limited treatment anticipation (assumption 3) and conditional parallel trends based on a "never-treated" group (assumption 4). I assume that there is no anticipation effect as increasing tax rates is not necessarily priori known and the individuals (mothers) do not actively select or determine their treatment status. To assess the validity of the parallel trends assumption, I conducted an event-study analysis. This analysis is performed on two models. Model 1 does not contain any control variables, while Model 2 contains 4 control variables: "Sex", "Mother education", "Mother race" and "Marital status". This means that the presence of unconditional parallel trends is tested by Model 1, while Model 2 tests for the presence of conditional parallel trends. The conditional parallel trends assumption proposed by Callaway and Sant'Anna (2021) is less strict compared to the two-way fixed effect model, it allows parallel trends to hold only after conditioning on covariates.

Figure 1 and Figure 2 show the results of the event-study analyses. The plot of Model 1 represents the results of the model without covariates, while the plot of Model 2 shows the results of the main model with control variables. The red dots and confidence intervals show the periods before the tax change was introduced. On the plot of the unconditional parallel trends assumption (Model 1) it can be seen that in some pre-treatment periods, the aggregate event-time average treatment effect is significantly different from zero, meaning that the unconditional parallel trends assumption does not hold. However, on the plot of conditional parallel trends assumption (Model 2) the aggregate event-time average treatment effect is significantly different from zero in four periods ($\tau = -21$, $\tau = -22$, $\tau = -17$, $\tau = -16$) only, suggesting that the conditional parallel trends assumption is likely to hold. Based on this result, I posit that the validity of the assumptions of Callaway and Sant'Anna (2021) is supported, however, the results should be interpreted with caution as the presence of pre-trends is not completely clear.

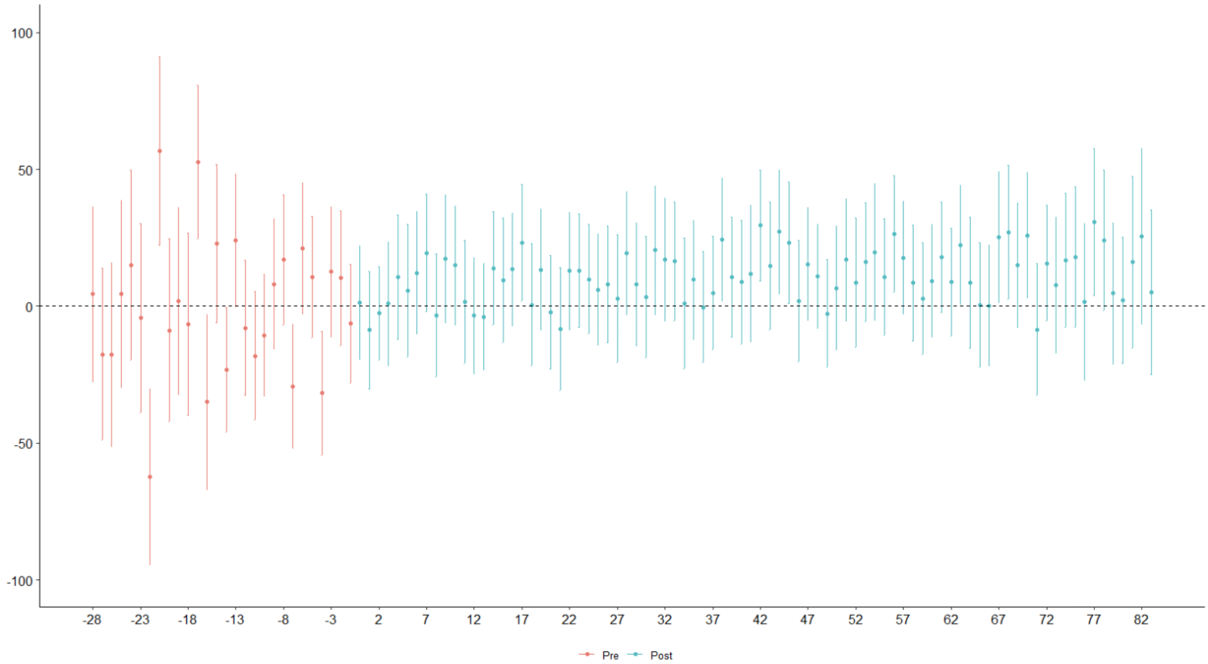

Figure 1: Event study plot for Model 1

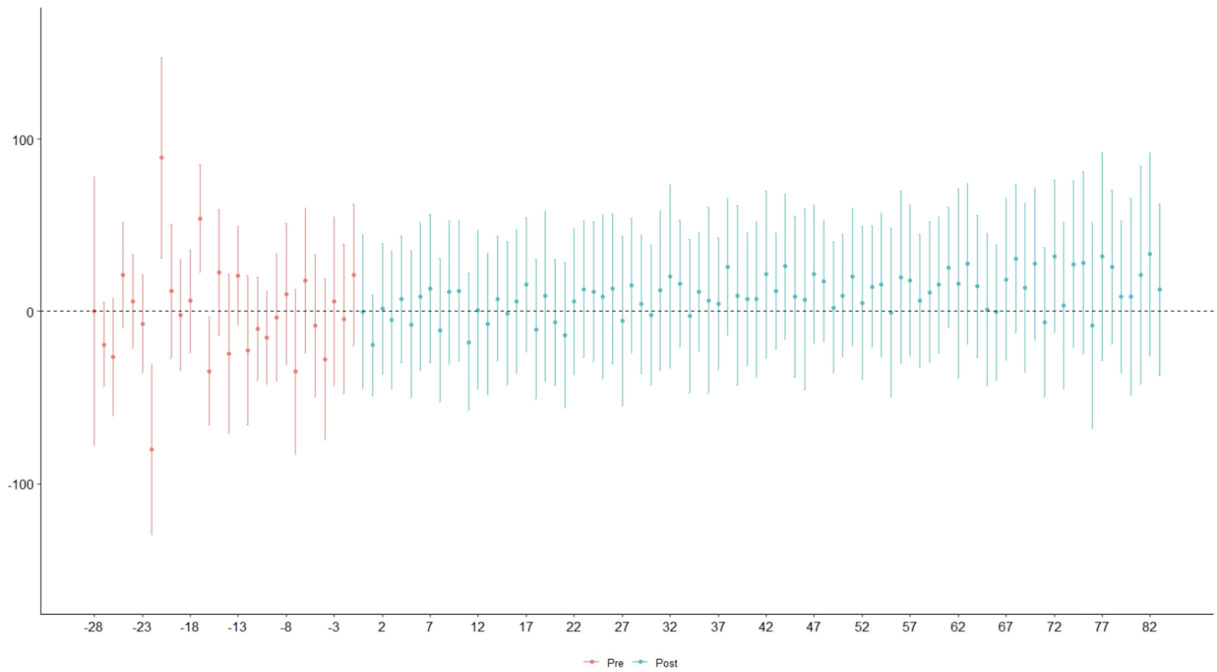

Figure 2: Event study plot for Model 2

5.3 The effect of changes in tax on maternal alcohol consumption

Before analysing the causal relationship between changes in alcohol beverage taxes and birth outcomes, it is essential to assess the effect of tax changes on maternal alcohol consumption as an intention-to-treat effect. To analyse this causal relation, I estimated two staggered difference-in-differences models using the method of Callaway and Sant'Anna. The estimation method is outcome regression and "never-treated" units form the control group. State-fixed effects and month-fixed effects are involved in both models. Model 1 is estimated without time-varying control variables, while Model 2 contains 4 control variables: "Sex", "Mother education", "Mother race" and "Marital status". "Mother age", "Father age" and "Father education" variables are excluded due to the high proportion of missingness. The dataset was aggregated on the state level and on a monthly basis before estimating the models. The aggregated data is weighted by state population. Standard errors are computed by using the multiplier bootstrap method.

Table 5 represents the results and the regression outputs of estimation equation 3. The first line (ATT) shows the average treatment effects for each treatment group and the aggregated effect across all groups. As the treatment states changed treatment status in different months, the treatment effects can be seen for each state individually. The aggregated group-specific average treatment effects range between 0.024 and 0.044 drinks per week. This means that maternal alcohol consumption increased by a range from 0.024 to 0.044 drinks per week on average after the increase in alcohol beverage taxes was implemented. The 95% confidence intervals of the aggregated ATTs do not include zero, meaning that these estimates are statistically significantly different from zero. However, looking at the group-specific ATTs, in Model 1, the group-specific ATTs range from 0.009 to 0.073, while in Model 2, the group-specific ATTs take values between 0.016 and 0.081. It can be seen that Delaware is the only state that has statistically significant results in both models. Furthermore, the ATTs of Delaware are significantly larger compared to the ATTs observed in the other two treatment states, thereby augmenting the overall aggregated effect.

The overall results are surprising, particularly due to the positive results and statistical significance of the aggregated estimates. The rationale behind employing alcohol tax changes as a proxy for alcohol consumption comes from previous studies (Wagenaar et al., 2009; Daley et al., 2012) that have shown evidence of the negative own-price elasticity of alcoholic beverages and reduction in consumption after the tax increase. Based on these studies, negative significant results were expected. However, as I previously mentioned, one of the main limitations of this data is the presumed measurement error in the number of drinks per week. The birth certificate question does not encourage mothers to report accurately, causing an underestimation of the recorded drinks. In total, only 1.6% of mothers reported drinking during pregnancy on the birth certificates, however, the estimated percentage of drinking mothers was 12.4% in 1991 and 16.3% in 1995 in the United States (Centers for Disease Control and Prevention, 1995). Furthermore, maternal alcohol consumption variable shows a high level of variability and is indicated to be too noisy to capture the effect.

Table 5: Results I.
Outcome variable: Maternal alcohol consumption

	Model 1			
	Connecticut	Delaware	Rhode Island	Aggregated
ATT	0.017	0.073	0.009	0.024
Confidence interval	$[0.004; 0.030]^*$	$[0.049; 0.098]^*$	$[-0.005; 0.22]$	$[0.014; 0.033]^*$
Standard error	0.007	0.013	0.007	0.005
FE: State	yes			
FE: Month	yes			
Controls	no			
Population weights	yes			

	Model 2			
	Connecticut	Delaware	Rhode Island	Aggregated
ATT	0.043	0.081	0.016	0.044
Confidence interval	$[-0.003; 0.090]$	$[0.042; 0.120]^*$	$[-0.021; 0.052]$	$[0.013; 0.074]^*$
Standard error	0.024	0.020	0.019	0.016
FE: State	yes			
FE: Month	yes			
Controls	yes			
Population weights	yes			

Note: Level of statistical significance: $^*\, p < 0.5$.

5.4 The effect of changes in tax on the birth weight of babies

The main question of interest is whether alcohol consumption affects birth weight. To analyse this effect with a difference-in-differences model, I use changes in alcohol tax as an intention-to-treat effect for alcohol consumption. I estimate two staggered difference-in-differences models using the method of Callaway and Sant'Anna. Similarly to the previous causal question, the estimation method is outcome regression and I only used "never-treated" units as controls. State-fixed effects and month-fixed effects are set in both models. In addition, Model 2 includes control variables. The dataset was aggregated on the state level and on a monthly basis before estimating the models. The aggregated data is weighted by state population. Standard errors are computed by using the multiplier bootstrap method.

Table 6 represents the results and the regression output of estimation equation 2. The first line (ATT) shows the aggregated group-time average treatment effects which is the main coefficient of interest. The aggregation method of ATT is "simple", assuming that the treatment effect is homogenous and there is no within-group variation. ATTs range between 9.36 and 10.64, indicating an increase in birth weights ranging from 9.36 grams to 10.64 grams in treatment states after the increase in tax rates. The confidence intervals range from -15.49 to 34.21 depending on the specification used, meaning that the overall effect of increasing the alcohol beverage taxes ranges from a decrease of approximately 15.49 grams in birth weight to an increase of 34.21 grams in the birth weight of the babies. As the mean of the control group is 3352.05, the treatment effect is negligible. In Model 1, the 95% confidence intervals of the aggregated group-time average treatment effects do not include zero, suggesting that these estimates are statistically significantly different from zero. On the other hand, when involving control variables in the model, the estimates are not significant as the 95% confidence intervals include zero.

In the presence of treatment heterogeneity across all individuals within each group, these estimates may not be correct, potentially leading to a Type II error (false negative). This means that even if the true effect exists, the analysis may fail to detect it. To handle this potential problem, I computed average treatment effects for each treatment group and then averaged the effect across all groups. As the treatment states changed treatment status in different months, the treatment effects can be seen for each state individually. Table 7 shows the results of the estimation of the group-specific average treatment effects. The specifications used and the models computed are the same as previously. As can be seen from the table, the aggregated group-specific average treatment effects range between 8.70 and 9.91 grams. The confidence interval ranges between -15.07 and 32.47 grams, meaning that the overall effect of increasing alcohol tax rates ranges from a decrease of 15.07 grams to an increase of 32.47 grams in the birth weights. The overall increase in birth weights is only statistically significant in Model 1 and this increase is negligible compared to the mean of the control group (3352.05). However, looking at the group-specific ATTs, it can be seen that the obtained results are startling, with a significant deviation from the expected findings. Connecticut is the only treatment state that has positive ATTs in both models. In Model 1, the estimates show a 20.31 grams increase in birth weights in Connecticut while in Model 2, this increase is slightly smaller, 19.12 grams, but the result is only significant in Model 1. The other two states do not have significant results. In Delaware, the ATTs range between -6.56 and -8.16, while in Rhode Island, the ATTs are slightly larger in absolute terms, ranging between -9.34 and -12.75. The rationale behind the negative ATTs can be the increase in maternal alcohol consumption, based on the estimation of the intention-to-treat regression.

Table 6: Results II.
Outcome variable: Birth weight

	Model 1	Model 2
ATT	10.64	9.36
Confidence interval	$[1.75; 19.54]^{*}$	$[-15.49; 34.21]$
Standard error	4.54	12.68
FE: State	yes	yes
FE: Month	yes	yes
Controls	no	yes
Population weights	yes	yes

Note: Level of statistical significance: $^{*}\ p < 0.5$.

Table 7: Results III.
Outcome variable: Birth weight

	Model 1			
	Connecticut	Delaware	Rhode Island	Aggregated
ATT	20.31	-8.16	-9.34	9.91
Confidence interval	$[4.56; 36.06]^*$	$[-26.74; 10.41]$	$[-26.34; 7.66]$	$[0.82; 19.00]^*$
Standard error	6.44	7.59	6.94	4.64
FE: State	yes			
FE: Month	yes			
Controls	no			
Population weights	yes			

	Model 2			
	Connecticut	Delaware	Rhode Island	Aggregated
ATT	19.12	-6.56	-12.75	8.70
Confidence interval	$[-20.03; 58.27]$	$[-18.89; 5.76]$	$[-44.96; 19.46]$	$[-15.07; 32.47]$
Standard error	18.93	5.96	15.57	12.13
FE: State	yes			
FE: Month	yes			
Controls	yes			
Population weights	yes			

Note: Level of statistical significance: $^* p < 0.5$.

In addition to the main outcome variable, I estimated the models with two other dependent variables: low birth weight and very low birth weight. Both variables are binary, they are stated as 1 if the baby has low or very low birth weight and 0 otherwise. Table 8 shows the results for low birth weight and Table 9 represents the results for very low birth weight. The specifications used and the models computed are the same as previously. As can be seen from Table 8, the aggregared group-specific average treatment effect is -0.001 for Model 1, while 0.0003 for Model 2. The treatment effect is negligible, 1.3% for Model 1 and 0.4% for Model 2. However, these estimates are not statistically significantly different from zero. In case of very low birth weight, Table 9 shows that the aggregared group-specific average treatment effect ranges between -0.003 and -0.002. The confidence intervals range from -0.004 to -0.001, meaning that the overall effect of increasing tax rates ranges from a decrease of 0.4 percentage points to a decrease of 0.1 percentage points in the proportion of babies born with very low birth weight. The confidence intervals do not include zero, suggesting that the decrease is statistically significantly different from zero. Furthermore, the group-specific ATTs are statistically significant for all treatment states in Model 1 and for Delaware in Model 2. Comparing the three treatment states, the treatment effect is prominent in Delaware. The mean of the control group is 0.014, demonstrating a 50% treatment effect in Delaware. The treatment effect of the aggregated results is 21% for Model 1, while 14% for Model 2.

Table 8: Results IV.
Outcome variable: Low birth weight

	Model 1			
	Connecticut	Delaware	Rhode Island	Aggregated
ATT	-0.003	-0.001	0.005	-0.001
Confidence interval	$[-0.007; 0.0004]$	$[-0.005; 0.003]$	$[0.001; 0.009]^{*}$	$[-0.004; 0.001]$
Standard error	0.002	0.002	0.002	0.001
FE: State	yes			
FE: Month	yes			
Controls	no			
Population weights	yes			

	Model 2			
	Connecticut	Delaware	Rhode Island	Aggregated
ATT	-0.001	-0.001	0.007	0.0003
Confidence interval	$[-0.008; 0.005]$	$[-0.004; 0.002]$	$[0.003; 0.010]^{*}$	$[-0.004; 0.004]$
Standard error	0.003	0.001	0.002	0.002
FE: State	yes			
FE: Month	yes			
Controls	yes			
Population weights	yes			

Note: Level of statistical significance: $^{*}\, p < 0.5$.

Table 9: Results V.
Outcome variable: Very low birth weight

	Model 1			
	Connecticut	Delaware	Rhode Island	Aggregated
ATT	-0.003	-0.007	-0.002	-0.003
Confidence interval	$[-0.004; -0.001]^*$	$[-0.009; -0.006]^*$	$[-0.003; -0.001]^*$	$[-0.004; -0.002]^*$
Standard error	0.001	0.001	0.001	0.000
FE: State	yes			
FE: Month	yes			
Controls	no			
Population weights	yes			

	Model 2			
	Connecticut	Delaware	Rhode Island	Aggregated
ATT	-0.001	-0.007	-0.001	-0.002
Confidence interval	$[-0.003; 0.001]$	$[-0.008; -0.006]^*$	$[-0.002; 0.0004]$	$[-0.003; -0.001]^*$
Standard error	0.001	0.001	0.001	0.001
FE: State	yes			
FE: Month	yes			
Controls	yes			
Population weights	yes			

Note: Level of statistical significance: $^* p < 0.5$.

To conclude, in this section two models were estimated with three outcome variables to capture the effect of an increase in alcohol tax rates on the birth weight of the babies. As the conditional trends assumption is likely to hold, I choose Model 2 to conclude from the findings. The overall results of the first model suggest that the increase in birth weights after the tax change was introduced is 8.70 grams. The confidence interval of the aggregated group-specific average treatment effect ranges between -15.07 and 32.47, suggesting that the overall effect of increasing alcohol tax rates ranges from a decrease of 15.07 grams to an increase of 32.47 grams in the birth weights. However, the confidence interval includes zero, meaning that the estimate is not statistically significantly different from zero. Furthermore, the overall results of the second model suggest that the proportion of babies born with low birth weight increases with 0.03 percentage points after the tax change was implemented. The confidence interval of the aggregated group-specific average treatment effect includes zero, meaning that the estimate is not statistically significantly different from zero. Moreover, the overall results of the third model demonstrate a 0.2 percentage points decrease in the proportion of babies born with very low birth weight. The confidence interval of this model ranges between -0.003 and -0.001, meaning that the overall effect of increasing alcohol tax rates ranges from a decrease of 0.3 percentage points to a decrease of 0.1 percentage points in the proportion of babies born with very low birth weight. The confidence interval do not include zero, thus the estimate is statistically significantly different from zero.

5.5 Robustness check

This subsection focuses on the robustness checks employed on the staggered difference-in-differences models. The methodology remains consistent with the previous models. I estimated four models with different specifications using the method of Callaway and Sant'Anna. The estimation method is outcome regression and "never-treated" units form the control group. State-fixed effects and month-fixed effects are involved in all models. Model 1 is estimated without using population weights and control variables. Model 2 is also estimated without using population weights, but it includes 4 control variables: “Sex”, “Mother education”, “Mother race” and “Marital status”. Compared to Model 2, Model 3 contains 2 additional variables: “Father age” and “Father education”. Model 4 contains all variables and uses population weights. The “Mother age” variable is not involved in any of the models, as the proportion of missingness is 78%.

Table 10 shows the results of the above-mentioned models. Aggregated group-specific average treatment effects are calculated for each specification. In the first block of the table, the results show the estimates for maternal alcohol consumption. ATTs range between 0.033 and 0.063. Compared to the original models, the estimates exhibit higher values without using population weights. This is probably due to the large differences in population between treatment and control states. Furthermore, upon the inclusion of all control variables, the ATTs became even larger. However, the estimates are statistically significantly different from zero. The second block shows the results of the estimates for birth weight. The results are not significant at any relevant level of significance. The aggregated group-specific average treatment effects range between -2.855 and 1.846. In comparison to the original models, the estimates show negligible effects of the tax changes on birth weights. The third block of the table shows the results for low birth weight. The results are not statistically significant, similarly to the original models. The aggregated group-specific average treatment effects range from -0.0004 to 0.0033. In Model 3 and Model 4, the treatment effect is larger than in the original models, however, the effect remains negligible (2.8% for Model 3 and 4.4% for Model 4). The last block of the table shows the results for very low birth weight. Compared to the original models, the estimates demonstrate slightly higher values without using population weights, while Model 4 shows negligible effect. The results are in line with the original model and suggest a decrease in the proportion of babies born with very low birth weight after the tax increase. The estimates are statistically significantly different from zero, except for Model 4.

Table 10: Robustness check

Outcome variable: Maternal alcohol consumption				
	Model 1	Model 2	Model 3	Model 4
ATT	0.033	0.056	0.063	0.061
Confidence interval	$[0.025; 0.041]^*$	$[0.037; 0.076]^*$	$[0.043; 0.084]^*$	$[0.026; 0.097]^*$
Standard error	0.004	0.010	0.010	0.018
FE: State	yes	yes	yes	yes
FE: Month	yes	yes	yes	yes
Controls	no	yes	yes	yes
Population weights	no	no	no	yes

Outcome variable: Birth weight				
	Model 1	Model 2	Model 3	Model 4
ATT	1.183	1.263	-2.855	1.846
Confidence interval	$[-10.446; 12.811]$	$[-15.526; 18.052]$	$[-21.791; 16.081]$	$[-29.073; 32.766]$
Standard error	5.933	8.566	9.661	15.776
FE: State	yes	yes	yes	yes
FE: Month	yes	yes	yes	yes
Controls	no	yes	yes	yes
Population weights	no	no	no	yes

Outcome variable: Low birth weight				
	Model 1	Model 2	Model 3	Model 4
ATT	-0.0004	0.0006	0.0021	0.0033
Confidence interval	$[-0.003; 0.003]$	$[-0.003; 0.004]$	$[-0.012; 0.005]$	$[-0.002; 0.008]$
Standard error	0.002	0.002	0.002	0.003
FE: State	yes	yes	yes	yes
FE: Month	yes	yes	yes	yes
Controls	no	yes	yes	yes
Population weights	no	no	no	yes

Outcome variable: Very low birth weight				
	Model 1	Model 2	Model 3	Model 4
ATT	-0.0040	-0.0029	-0.0022	-0.0004
Confidence interval	$[-0.005; -0.003]^*$	$[-0.005; -0.001]^*$	$[-0.003; -0.001]^*$	$[-0.002; 0.001]$
Standard error	0.001	0.001	0.001	0.001
FE: State	yes	yes	yes	yes
FE: Month	yes	yes	yes	yes
Controls	no	yes	yes	yes
Population weights	no	no	no	yes

Note: Level of statistical significance: $^*\, p < 0.5$.

6 Conclusion and Discussion

Prenatal alcohol consumption and thus alcohol exposure to babies during pregnancy is a relatively common problem in the United States. Alcohol exposure during the prenatal period can lead to different fetal alcohol effects, such as Fetal Alcohol Syndrome. Addictive habits can affect the mother's health conditions which may lead to premature birth or intrauterine growth restriction. As a consequence, the birth outcomes of the babies may be adversely influenced, resulting in e.g. low birth weight or behavioural deficits. The main goal of this research was to find empirical evidence on the effect of maternal alcohol consumption on the birth outcomes of babies. The birth outcomes were measured by three variables: the birth weight of babies in grams, low birth weight and very low birth weight.

To analyse the main research question, Callaway and Sant'Anna's difference-in-differences with multiple treatment periods approach was implemented. However, the results were estimated in two stages: first, I examined the effect of alcohol beverage tax changes on maternal alcohol consumption, second, I calculated the effect of maternal alcohol consumption on the birth outcomes. The first-stage regression was aimed to determine the intention-to-treat effect. In the second-stage regression changes in tax served as a proxy for maternal alcohol consumption.

The results of the first-stage regression showed an increase in maternal alcohol consumption after the increase in tax rates. Despite the results being significant, they were not anticipated due to the negative own-price elasticity of alcohol consumption. One possible explanation for the positive result is that studies showed that a low amount of alcohol consumption may have a "healthy drinker effect" or it may be beneficial (**?**). In this study, prenatal alcohol consumption can be considered low as the mean of the control group is 0.045 drinks per week. By employing tax changes as a proxy for alcohol consumption in the second-stage regression, I estimated two models with three outcome variables. In the model where birth weight was considered as the outcome variable, the results showed an insignificant and negligible increase in birth weights. In the model utilizing low birth weight as the dependent variable, the results suggested an insignificant and negligible increase in the proportion of babies born with low birth weight after the increase in tax rates. However, the last model demonstrated a 0.2 percentage point decrease in the proportion of babies born with very low birth weight after the changes in tax rates. This increase was significant, showing a 14% treatment effect.

There are several limitations of this study. Firstly, the presumed measurement error in the number of drinks per week. The data on alcohol consumption was not compiled in a careful manner as it was collected from birth certificates after pregnancy. Thus, the reports may have been affected by recall bias. Women have a tendency to underreport alcohol consumption during pregnancy, causing an underestimation of the results. Secondly, I only had data on average alcohol consumption during the whole period of pregnancy, but the time of alcohol exposure may be relevant to birth outcomes (Nykjaer et al., 2014). By incorporating trimester-specific data, more accurate analysis can be conducted. Thirdly, there were a small number of states that changed their alcohol beverage tax rates in the examined period. A longer time period and more states that change treatment status would give more precise estimates. Finally, Natality Detail Files do not contain detailed information on the socio-economic status of the mothers, furthermore, treatment and control states are different in observed characteristics. This implies that the impact of policy changes may vary across different states and a general conclusion cannot be drawn.

The general conclusion is that this study does not find clear evidence of the effect of maternal alcohol consumption on the birth weights of babies. While a significant decrease in the proportion of babies born with very low birth weight is supported, no significant effect is found in the evolution of birth weights and in the proportion of babies born with low birth weight. However, the results should be considered and interpreted with caution as there are several limitations in this study.

References

Bingol, N., Schuster, C., Fuchs, M., Iosub, S., Turner, G., Stone, R. K., and Gromisch, D. S. (1987). The influence of socioeconomic factors on the occurrence of fetal alcohol syndrome. *Advances in alcohol & substance abuse*, 6(4):105–118.

Calhoun, F. and Warren, K. (2007). Fetal alcohol syndrome: historical perspectives. *Neuroscience & Biobehavioral Reviews*, 31(2):168–171.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Centers for Disease Control and Prevention (1995). Alcohol consumption among pregnant and childbearing-age women: 1988 and 1995. CDC WONDER Online Database. Accessed: June 7, 2023. URL: https://wonder.cdc.gov/wonder/prevguid/m0047306/m0047306.asp.

Chen, J.-H. (2012). Maternal alcohol use during pregnancy, birth weight and early behavioral outcomes. *Alcohol and Alcoholism*, 47(6):649–656.

Children's Hospital of Philadelphia (2023). Low birthweight. Web. Accessed: May 06, 2023. URL: https://www.chop.edu/conditions-diseases/low-birthweight.

Clarren, S. and Smith, D. (1978). The fetal alcohol syndrome. *New England Journal of Medicine*, 298:1063–1067.

Daley, J. I., Stahre, M. A., Chaloupka, F. J., and Naimi, T. S. (2012). The impact of a 25-cent-per-drink alcohol tax increase. *American journal of preventive medicine*, 42(4):382–389.

Delaware Division of Revenue (2019). Delaware alcoholic beverage control. Web. Accessed: May 06, 2023. URL: https://financefiles.delaware.gov/docs/abc.pdf.

Denny, C. H., Tsai, J., Floyd, R., Green, P., et al. (2009). Alcohol use among pregnant and non-pregnant women of childbearing age-united states, 1991-2005. *Morbidity and Mortality Weekly Report*, 58(19):529–532.

Erb, L. and Andresen, B. D. (1978). The fetal alcohol syndrome (fas) a review of the impact of chronic maternal alcoholism on the developing fetus. *Clinical Pediatrics*, 17(8):644–649.

Golden, J. L. (2009). *Message in a bottle: The making of fetal alcohol syndrome*. Harvard University Press.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Green, H. G. (1974). Infants of alcoholic mothers. *American journal of obstetrics and gynecology*, 118(5):713–716.

Huntington-Klein, N. (2021). *The effect: An introduction to research design and causality*. CRC Press.

ICPSR (1996). *Natality Detail File*. Inter-university Consortium for Political and Social Research. Accessed on 1 May 2023. URL: https://www.icpsr.umich.edu/web/ICPSR/studies/3388.

Jones, K. and Smith, D. (1973). Recognition of the fetal alcohol syndrome in early infancy. *The Lancet*, 302(7836):999–1001.

Jones, K., Smith, D., Ulleland, C., and Streissguth, A. (1973). Pattern of malformation in offspring of chronic alcoholic mothers. *The Lancet*, 301(7815):1267–1271.

Little, R. E. (1977). Moderate alcohol use during pregnancy and decreased infant birth weight. *American Journal of Public Health*, 67(12):1154–1156.

Little, R. E., Streissguth, A. P., Barr, H. M., and Herman, C. S. (1980). Decreased birth weight in infants of alcoholic women who abstained during pregnancy. *The Journal of Pediatrics*, 96(6):974–977.

Lohman, J. and Pinho, R. (2011). Connecticut general assembly. Web. Accessed: May 06, 2023. URL: https://cga.ct.gov/2011/rpt/2011-R-0252.htm.

Majewski, F. (1981). Alcohol embryopathy: some facts and speculations about pathogenesis. *Neurobehavioral toxicology and teratology*, 3(2):129–144.

Mariscal, M., Palma, S., Llorca, J., Pérez-Iglesias, R., Pardo-Crespo, R., and Delgado-Rodríguez, M. (2006). Pattern of alcohol consumption during pregnancy and risk for low birth weight. *Annals of epidemiology*, 16(6):432–438.

McCarthy, F. P., O'keeffe, L. M., Khashan, A. S., North, R. A., Poston, L., McCowan, L. M., Baker, P. N., Dekker, G. A., Roberts, C. T., Walker, J. J., et al. (2013). Association between maternal alcohol consumption in early pregnancy and pregnancy outcomes. *Obstetrics & Gynecology*, 122(4):830–837.

Miyake, Y., Tanaka, K., Okubo, H., Sasaki, S., and Arakawa, M. (2014). Alcohol consumption during pregnancy and birth outcomes: the kyushu okinawa maternal and child health study. *BMC Pregnancy and Childbirth*, 14:1–7.

Nykjaer, C., Alwan, N. A., Greenwood, D. C., Simpson, N. A., Hay, A. W., White, K. L., and Cade, J. E. (2014). Maternal alcohol intake prior to and during pregnancy and risk of adverse birth outcomes: evidence from a british cohort. *J Epidemiol Community Health*, 68(6):542–549.

Olsen, J., Pereira, A. d. C., and Olsen, S. F. (1991). Does maternal tobacco smoking modify the effect of alcohol on fetal growth? *American Journal of Public Health*, 81(1):69–73.

Olsen, J., Rachootin, P., and Schiødt, A. V. (1983). Alcohol use, conception time, and birth weight. *Journal of Epidemiology & Community Health*, 37(1):63–65.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Patra, J., Bakker, R., Irving, H., Jaddoe, V. W., Malini, S., and Rehm, J. (2011). Dose–response relationship between alcohol consumption before and during pregnancy and the risks of low birthweight, preterm birth and small for gestational age (sga)—a systematic review and meta-analyses. *BJOG: An International Journal of Obstetrics & Gynaecology*, 118(12):1411–1421.

Pruett, D., Waterman, E. H., and Caughey, A. B. (2013). Fetal alcohol exposure: consequences, diagnosis, and treatment. *Obstetrical & gynecological survey*, 68(1):62–69.

Rhode Island General Laws (1989). Rhode island general laws. Web. Accessed: May 06, 2023. URL: http://webserver.rilin.state.ri.us/Statutes/title3/3-10/3-10-1.1.HTM.

Riley, E. P., Infante, M. A., and Warren, K. R. (2011). Fetal alcohol spectrum disorders: an overview. *Neuropsychology review*, 21:73–80.

Royal College of Physicians (1726). Annals. *Royal College of Physicians, London, England*, page 253.

Sant'Anna, P. H. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122.

Sullivan, W. (1899). A note on the influence of maternal inebriety on the offspring. *Journal of Mental Science*, 45(190):489–503.

Tax Policy Center (2023). State alcohol excise tax rates. Accessed on 1 May 2023. URL: https://www.taxpolicycenter.org/statistics/state-alcohol-excise-tax-rates.

Tenbrinck, M. S. and Buchin, S. Y. (1975). Fetal alcohol syndrome: Report of a case. *Jama*, 232(11):1144–1147.

Verkerk, P. H., van Noord-Zaadstra, B. M., Florey, C. d. V., De Jonge, G., and Verloove-Vanhorick, S. (1993). The effect of moderate maternal alcohol consumption on birth weight and gestational age in a low risk population. *Early human development*, 32(2-3):121–129.

Wagenaar, A. C., Salois, M. J., and Komro, K. A. (2009). Effects of beverage alcohol price and tax levels on drinking: a meta-analysis of 1003 estimates from 112 studies. *Addiction*, 104(2):179–190.

Wehby, G. L., Fletcher, J. M., Lehrer, S. F., Moreno, L. M., Murray, J. C., Wilcox, A., and Lie, R. T. (2011). A genetic instrumental variables analysis of the effects of prenatal smoking on birth weight: evidence from two samples. *Biodemography and social biology*, 57(1):3–32.

Wehby, G. L. and Scholder, S. v. H. K. (2013). Genetic instrumental variable studies of effects of prenatal risk factors. *Biodemography and social biology*, 59(1):4–36.

WHO (1950). *Official records (Third World Health Assembly 3.6)*. World Health Organization.

Windham, G. C., Fenster, L., Hopkins, B., and Swan, S. H. (1995). The association of moderate maternal and paternal alcohol consumption with birthweight and gestational age. *Epidemiology*, 6(6):591–597.

Wright, J., Barrison, I., Lewis, I., MacRae, K., Waterson, E., Toplis, P., Gordon, M., Morris, N., and Murray-Lyon, I. (1983). Alcohol consumption, pregnancy, and low birthweight. *The Lancet*, 321(8326):663–665.

Zhang, N. (2010). Alcohol taxes and birth outcomes. *International journal of environmental research and public health*, 7(5):1901–1912.