The Effect of Sibling Gender on Substance Use During Adolescence: Evidence from Dizygotic Twins*

Eunju Lee†

September 12, 2024

Abstract

This paper examines the impact of sibling gender on substance use during adolescence. I analyze a sample of dizygotic twins, leveraging the exogenous variation in their assigned sex at birth. This design helps me to address some methodological concerns in studying sibling gender effects and provide clean causal estimates. I find that among male adolescents, having a brother increases the probabilities of using cigarettes, alcohol, and marijuana. Regarding potential mechanisms, I find that the results are consistent with the channel of direct sibling influences, but not with the channels of differential parental investment, family structure, or epigenetic influences.

*I am very grateful to Marianne Bitler, Jenna Stearns, and Scott Carrell for their invaluable support and advice. I also thank Marianne Page, Ryan Matthew Finnigan, Benjamin Cowan, Alfonso Flores-Lagunes, Chang Jae (C.J.) Lee, Seojung Oh, Jou-Chun Lin, Estelle Shin, Baiyu Zhou, Kalyani Chaudhuri, Kevin Dinh, and seminar and conference participants at the 2022 Add Health Users Conference, the 2023 WEAI International Conference, the 2023 ASHEcon conference, UC Davis Social Mobility Tea, and UC Davis Applied Micro Student Talks for helpful comments and conversations. The opinions expressed in this paper are solely those of the author and do not necessarily reflect the viewpoints of any other organization. Any errors in the content are the author's responsibility.

†Department of Economics, Louisiana State University. Email: eunjulee@lsu.edu

1 Introduction

Which peers influence an adolescent's human capital development, and how? This question is of critical importance to parents, educators, policymakers, and researchers because peer dynamics can shape the optimal organization of schools, neighborhoods, and other social settings where adolescents interact. Peer characteristics, such as gender, may play a pivotal role in determining human capital outcomes. Adolescents tend to form stronger bonds with same-gender peers (Poulin and Pedersen, 2007), potentially intensifying peer influence within these groups. In contexts where there are large gender disparities—such as science, technology, engineering, and mathematics (STEM) participation (Brenøe and Zölitz, 2020)—greater exposure to peers of a particular gender may widen the existing gender gap in outcomes.[1] Understanding how peer gender shapes behavior is essential for developing effective strategies to enhance human capital production.

In this paper, I investigate how peer gender influences adolescent substance use. Despite the harmful effects of adolescent substance use on human capital formation and long-term outcomes,[2] it remains a widespread issue in the United States. In 2010, 75.6 percent of high school students had tried cigarettes, alcohol, or other drugs, and 46.1 percent were current users (National Center on Addiction and Substance Abuse, 2011). While previous studies have established that peers significantly influence substance use behaviors among adolescents (Gaviria and Raphael, 2001; Lundborg, 2006; Clark and Lohéac, 2007), the specific role of peer gender in this context remains unclear. This paper aims to fill that gap by investigating the causal effect of peer gender on adolescent substance use, a domain with a well-documented gender gap (Gruber, 2009).

Among various peer relationships, this study focuses on siblings, a particularly influential peer group. Siblings spend a substantial amount of time together (Claes, 1998), sometimes even more than with their friends.[3] This extensive exposure fosters substantial mutual influence (Black et al.,

[1]For instance, Brenøe and Zölitz (2020) suggests that having more female peers changes the gender gap in high school GPA in favor of males, which may lead women to perceive themselves as less equipped for STEM studies.

[2]For more details, see the followings: Academic performances (Ellickson et al., 2001; DeSimone, 2010; Balsa et al., 2011), educational attainments (Cook and Moore, 1993; Dee and Evans, 2003; Chatterji, 2006a,b), criminal activities (Sen et al., 2009; Chalfin et al., 2023) and labor market outcomes (MacDonald and Pudney, 2000; Van Ours, 2004).

[3]By surveying adolescents aged 11-18 years, the paper examines the daily time allocation with siblings and close friends. The findings underscore the important role of siblings among adolescents. In Italy, the time spent with siblings was substantial, with 4.80 hours for brothers, 4.65 hours for sisters, while close friends received 2.09 hours. Similarly in Belgium, adolescents spent 3.70 hours for brothers and 3.95 hours for sisters, but 3.00 hours for close friends.

2021; Dahl et al., 2014), unlike more distant peer groups, such as students in the same school-grade, who may share the same environment but do not necessarily interact closely. By examining the effect of sibling gender, this paper aims to uncover how peer gender—within the context of these close relationships—can shape adolescent behaviors, specifically substance use.

There are two key challenges in estimating the causal effects of sibling gender. First, sibling gender composition can be biased by selective parental fertility decisions. Parents often have specific preferences about their children's gender composition, which influences their subsequent fertility (Angrist and Evans, 1998; Dahl and Moretti, 2008). In some cases, this preference may even result in sex-selective abortions (Yamaguchi, 1989; Lin et al., 2014). These behaviors bias the observed sibling gender composition and, consequently, the estimated effect of sibling gender. Second, it is difficult to address the birth order effect and its possible interaction with the sibling gender effect. For example, as Averett et al. (2011) noted, later-born children receive less parental supervision than first-born children, and this effect is further pronounced when they have a same-sex older sibling. Failure to address such an interaction can result in biased estimates, but addressing all potential interactions is very challenging.

I address these two challenges by utilizing a sample of dizygotic twins, formed from the fertilization of two separate eggs by two separate sperm, in the National Longitudinal Study of Adolescent Health (Add Health). First, the sex of dizygotic twins at birth is randomly assigned in the absence of assisted reproductive technologies. As these methods were not widely utilized when my sample respondents were born, it provides a natural experiment for investigating sibling gender, free from parental fertility selection. Second, the design minimizes birth order effects and effectively circumvents the intricate interactions between birth order and sibling gender effects, as twins share the same age, birth timing, and environment at each age.

I examine sibling gender effects on adolescent substance use by own gender. In other words, I compare males with a twin brother to males with a twin sister, and compare females with a twin brother to females with a twin sister. This separate analysis is due to the fact that sibling gender effects may differ depending on an individual's own gender, and is in line with the literature.

I find that teen males with a twin brother are more likely to smoke cigarettes, drink alcohol, and use marijuana than males with a twin sister. The magnitudes of these effects are large. For example, males with a twin brother are more likely to be current smokers, alcohol drinkers, and marijuana users than males with a twin sister by 9.9 percentage points (30.2 percent compared to the mean), 15.5 percentage points (31.6 percent) and 5.5 percentage points (33.3 percent), respectively. By contrast, sibling gender has little impact on females. The negative effects of having a twin brother for males (while no significant effects for females) are consistent with prior findings on peer effects: Males are both more influential as peers and are more influenced by peers than are females.[4]

I then explore potential mechanisms underlying the sibling gender effects: Direct sibling interactions (Peter et al., 2018), parental investment (Cools and Patacchini, 2019; Brenøe, 2021), and biological influence (Tapp et al., 2011). While I find no evidence of differential parental investment or changes in family structure by sibling gender, I argue that the results align with the sibling interactions channel. Specifically, I propose two examples of sibling interactions with suggestive evidence: Sharing a friend network for substance access and serving as a reference point for substance use.

This paper contributes to the literature on sibling gender effects in three important ways. First, by utilizing dizygotic twins, it effectively addresses two key methodological concerns: Selective parental fertility decisions and complex interactions between birth order and sibling gender. Prior studies have traditionally included both older and younger siblings in the sample and compared those with sisters to those with brothers.[5] To address the two concerns, particularly selective parental fertility, more recent studies have focused on the effects of the younger sibling's gender on the older sibling, as the younger sibling's gender is random conditional on the their birth. Yet, this recent approach is less intuitive in the context of substance use where older siblings exert more influence on younger siblings than vice versa (Rowe and Gulley, 1992; Lee and Schnorr, Forthcoming). In the context of substance use, my twin research design offers a more suitable solution to these methodological concerns.

Second, this paper studies the causal effects of sibling gender on substance use, which, despite its far-reaching negative impacts on health and well-being, is not well-established in the literature. While sibling gender has been found to significantly impact various outcomes,[6] prior studies have

[4]Examples include: Non-cognitive skills among siblings (Cyron et al., 2017; Golsteyn and Magnée, 2017), academic performance (Hoxby, 2000; Lavy and Schlosser, 2011; Gottfried and Graves, 2014; Lee et al., 2014; Hu, 2015; Hill, 2017) and disruptive behaviors among school peers (Clark and Lohéac, 2007; Carrell and Hoekstra, 2010) , and crime among juvenile prisoners (Stevenson, 2017).

[5]For instance, see Butcher and Case (1994), Kaestner (1997), Hauser and Kuo (1998), Conley (2000), Brunello and De Paola (2013), Anelli and Peri (2015), Cyron et al. (2017), and Rao and Chatterjee (2018).

[6]The literature suggests causal effects of sibling gender on occupations (Brenøe, 2021), earnings (Peter et al., 2018;

paid little attention to its effects on substance use. To the best of my knowledge, there are only two relevant economic papers. Cools and Patacchini (2019) briefly explores sibling gender effects on some disruptive behaviors only among females with younger brothers in their mechanism section. Argys et al. (2006) compares the effect of having older sisters to that of having older brothers among younger siblings regardless of birth order and birth space in their heterogeneity section. However, nearly 25 percent of their participants had both older sisters and brothers and were included in both estimations. Moreover, since parents who stop after one child are excluded from their sample, while those pursuing a second child are included, parental preferences for child gender affects both the observed sibling composition and parenting styles, potentially biasing their estimates of the sibling gender effect. By providing the first causal evidence on how sibling gender affects substance use, this paper fills the gap in the literature on sibling gender.

Third, this paper sheds light on an underexplored mechanism of sibling gender effects—direct sibling interactions—by providing suggestive evidence. While prior studies have extensively examined channels of parental investment and family structure,[7] the channel of direct sibling interactions has been less explored due to data limitations. Utilizing extensive survey data on adolescent behaviors in Add Health, this paper provides suggestive evidence on how siblings directly influence each other, deepening our understanding of sibling gender effects.

This paper also contributes to the broader literature on peer gender effects. Investigating the causal effect of "friend" gender is very challenging due to the selective nature of friendships. To avoid this selection, prior studies have relied on measures such as gender composition within specific peer groups.[8] However, it remains uncertain whether these measures accurately capture the impact of friends with whom adolescents spend substantial time together, as these measures may include peers with infrequent interactions. In contrast, twins are more likely to view each other as their friends, compared to non-twin full siblings (Fraley and Tancredy, 2012) and school peers (McGuire and Segal, 2013). Twins (and their parents) in my sample are not able to choose each other's sex, free from selecting friends of a particular gender. Twins share the same family environment at each developmental stage, which eliminates selective friendship based on family

Cools and Patacchini, 2019), family formation (Peter et al., 2018), and personality (Golsteyn and Magnée, 2020).

[7]Examples include Butcher and Case (1994), Oguzoglu and Ozbeklik (2016), Cools and Patacchini (2019), Golsteyn and Magnée (2020), and Brenøe (2021).

[8]Such peer groups include individuals in the same classroom (Hoxby, 2000; Gottfried and Graves, 2014; Lee et al., 2014), school-grade (Lavy and Schlosser, 2011; Hu, 2015; Hill, 2017), and neighborhood (Hill, 2015).

environments. By utilizing twins, this study offers valuable insights into the effect of friend gender.

The remainder of this paper proceeds as follows. Section 2 introduces the identification strategy and describes the data. Section 3 presents the results and checks their robustness. Section 4 explores potential mechanisms behind the findings. Section 5 concludes.

2 Empirical Strategy and Data

2.1 Twin research design

This study utilizes dizygotic twins who, on average, share approximately 50 percent of their genes, similar to other full siblings. Dizygotic twins, resulting from the fertilization of two separate eggs by two separate sperm, have been widely recognized in biology as having randomly determined sexes in nature. For instance, Weinberg's differential method, a commonly used approach for estimating the number of dizygotic twins and different-sex twins, assumes that the sex of each twin in a dizygotic pair is *independently* determined from the other twin's sex. Numerous empirical studies have provided substantial support for this method and particularly, its underlying assumption of independence (Vlietinck et al., 1988; Husby et al., 1991; Fellman and Eriksson, 2006).

For my twin research design, it is crucial to determine the zygosity of twin pairs and focus on dizygotic twins. This is because monozygotic twins, formed when a single zygote splits, are always of the same sex. As a result, given one's sex, there is no variation in the co-twin's sex to explore, making it impossible to investigate the effect of having a same-sex co-twin versus a different-sex co-twin. Furthermore, monozygotic twin pairs share 100 percent of their genes, setting them apart genetically from dizygotic twins or non-twin full siblings who, on average, share 50 percent of their genes. This genetic dissimilarity poses challenges when comparing them with same-sex dizygotic twins or same-sex non-twin full siblings.[9] Therefore, it is essential to have accurate information on zygosity and limit the analysis to dizygotic twins (Peter et al., 2018).[10]

[9]For example, Table 1.A1 indicates that the correlation in substance use is considerably higher among monozygotic twins than among dizygotic same-sex twins across all measures. This finding underscores the influence of genetic factors in shaping substance use behaviors among twin pairs.

[10]For readers who may be interested in comparing monozygotic twins and same-sex dizygotic twins, Table 1.A2 reveals that monozygotic twins are notably less likely to engage in substance use compared to dizygotic same-sex twins, particularly among males. This finding is consistent with the prior studies that monozygotic twins are known to

The twin research design has several advantages. First, it addresses the parental selection bias on the gender composition of their children. The traditional approach is to compare those with a brother to those with a sister regardless of their birth order (see footnote 6). However, it has been found that parents in the United States may want to have at least one son and one daughter (Angrist and Evans, 1998) or at least one son (Dahl and Moretti, 2008). This implies that parents who plan another birth after having a specific children's gender composition may have different preferences and even different parenting styles. With my research design of using dizygotic twins, however, I can circumvent the parental selection bias: Twins are born at the same time, the sexes of twins are determined independently, and parents cannot choose one child's birth based on the sex of the other child.

Second, the design allows me to explore sibling gender effects separately from the effect of their birth order. Among non-twin siblings, there is always a birth order, and thus we need to consider the birth order effect and its interaction with the sibling gender effect. For example, Price (2008) finds that the differences in time spent with the father among siblings are the greatest when the father has the first-born son and second-born daughter. It is also found that parents supervise their younger child less if there is a same-sex older child, potentially leading to younger children being supervised by their same-sex older siblings (Averett et al., 2011). In contrast, twins share the same age, birth timing, and environment at each age, effectively circumventing the intricate interactions between birth order and sibling gender effects.

Third, in a broader sense, the twin research design offers an exceptional framework to explore the effects of "friend" gender. Twins are more likely to regard each other as their friends than non-twin full siblings (Fraley and Tancredy, 2012) or school peers (McGuire and Segal, 2013). Meanwhile, the twin design eliminates two selection biases in friendships. Twins and their parents cannot choose the gender composition of the twin pair, leading to a causal estimate that avoids selection biases in friend gender.[11] Also, they are of the same age and share the same family environment at each age, which eliminates selective friendship based on family environments.

While the twin design offers certain advantages, it is important to be cautious when generaliz-

exhibit better behaviors and outcomes compared to same-sex dizygotic twins or non-twin full siblings: They are more likely to have stable marriage relationship (Heller et al., 1988), higher educational level and net worth (Felson, 2014), and better health outcomes (Kanazawa and Segal, 2019) and lower mortality rate (Sharrow and Anderson, 2016).

[11] Non-twin full siblings cannot choose each other's gender but their parents may do, which results in the parental selective fertility decisions.

ing the findings of this paper to non-twin families. Twinning can be influenced by several factors. Although the respondents in my sample were born before the widespread use of in vitro fertilization (IVF), as discussed in the next section, natural multiple pregnancies can still be mildly affected by factors such as maternal age (MacGillivray et al., 1988) and race (Oleszczuk et al., 2001).

Additionally, parents of twins may differ in their parenting styles and investments in children. Some studies suggest that raising twins fosters unique experiences for parents, which could lead to heightened emotional attachment (Holditch-Davis et al., 1999; Leonard and Denton, 2006). However, it also presents challenges, such as increased time demands and stress (Beck, 2002; Damato and Burant, 2008; Bolch et al., 2012; Heinonen, 2016), potentially resulting in stricter parenting practices and less nurturing behaviors (Anthony et al., 2005).

2.2 Data

This study utilizes data from the National Longitudinal Study of Adolescent to Adult Health Health (Add Health), a comprehensive survey conducted in the United States. The survey employed a stratified sampling method by selecting a random sample of high schools across the country in 1994. The sample included adolescents in grades 7 through 12 during the 1994-1995 school year, aiming to create a nationally representative cohort. The survey has since followed up with participants through five waves: Wave I (1994-1995), Wave II (1996), Wave III (2001-2002), Wave IV (2008-2009), and Wave V (2015-2017).

The Add Health dataset is highly suitable for this analysis due to its unique features. First, Add Health over-sampled twins and siblings and has a large twin sample size of 784 twin pairs, making it a valuable resource for twin research. While administrative data covering the entire population of twins would be optimal for twin research due to the low twinning rate, such sources often lack detailed information on substance use, particularly during adolescence. In contrast, Add Health stands out as the largest and most comprehensive longitudinal survey of US adolescents' health behaviors. By deliberately oversampling twins, Add Health ensures a sufficient number of twin participants, enabling rigorous examination of substance use outcomes among twins.

Second, 98.63 percent of the Add Health respondents were born between 1976 and 1982 when in vitro fertilization (IVF) or prenatal sex discernment techniques were not available or commonly utilized. The first instance of a child conceived with IVF in the United States occurred in 1981 (Barnhart, 2013). Prenatal sex discernment became available in the mid-70s, but their prevalence was not substantial enough to influence the sex ratios among children born in the United States until 1990, even among individuals whose parents hailed from countries with a high prevalence of sex-selective abortion (Almond and Edlund, 2008). Citro et al. (2014) also found that the prohibition of sex-selective abortions in Illinois and Pennsylvania—implemented in 1984 and 1989, respectively—had no discernible impact on sex ratios. This temporal context strongly supports the suitability of my twin research design to investigate the sibling gender effect.

Third, it provides comprehensive information on health behaviors, specifically on substance use of the respondents during adolescence. The primary objective of the survey was to facilitate research on the determinants of health and health behaviors among adolescents (Harris, 2013). As a result, the survey has collected extensive data on substance use across all waves, making it a valuable resource for studying this topic.

Among the twin pairs in the Add Health data, however, a few pairs are not suitable for my analysis. To begin with, I restrict my sample to dizygotic twins, following the classification of zygosity by the Add Health data team.[12] Additionally, I include only those who listed their co-twin in the household roster submitted by each respondent in Wave I (using the age and sex information of household members), which leads to the exclusion of 48 dizygotic twin pairs. [13] This restriction is in place to ensure that they live in the same household and, therefore, predominately have the same family environment. Table 1 displays the count of dizygotic twins in the final sample by sex and co-twin's sex.

The Add Health dataset also includes a sample of non-twin sibling pairs who, like the twin

[12]All twins with an different-sex co-twin were classified as dizygotic, as they are always dizygotic by nature. Then the same-sex twins were classified as dizygotic or monozygotic, based on their responses to the confusability of their appearance with the co-twin. (Questions about confusability include whether they looked like two peas in a pod when young children and whether strangers, teachers, or family members were confused by them.) If the information on self-reported appearance confusability was missing, then the data team used the responses of their mother to the questions on confusability. If all these responses are still not enough to classify the twin pairs, then the zygosity was determined by DNA tests (performed at Wave III and IV). Despite these efforts, 5.7 percent of twin pairs were not certainly classified. I drop these unclassified twin pairs in my analyses.

[13]For pairs who do not satisfy this condition, there can be two possibilities: (1) The twins do not live in the same household, or (2) the twins do live in the same household, but either one or both of the twin pairs forgot to report the other. Among the 48 pairs, 36 pairs attended the same schools, indicating that many of them are likely to fall into the latter category. The probability of not satisfying this condition is found not to be correlated with my treatment variable, the sex of the co-twin (correlation = 0.0276 for males; 0.0366 for females).

pairs, were in grades 7 through 12 during the 1994-95 school year. For comparison, I create a sample of non-twin sibling pairs comparable to the sample of twin pairs. Specifically, I ensure that the non-twin pairs were full siblings (sharing the same biological parents), resided in the same household at Wave I (sharing the same family environment), were the closest in age among their siblings based on household rosters (as twins were), and had an age spacing of two years (as closely spaced as possible while not being negatively selected).[14] The number of non-twin full siblings in the final sample is provided in Table 1.A3.

2.3 Specification

Following the existing literature on sibling gender effects, I conduct separate analyses for males and females, taking into account the potential differential impact of sibling gender for each group. Specifically, I compare males (or females) with a twin brother to males (or females) with a twin sister by employing the following model:

$$Y_i = \beta_0 + \beta_1 SS_i + X_i\Gamma + e_i \tag{1}$$

where Y_i is a measure of substance use of the respondent i and SS_i is a dummy variable which is equal to 1 if the respondent i has a same-sex twin sibling. The control variables X_i include region dummies, race dummies, age, age squared, residential mother's age, residential mother's college degree status, a dummy variable for missing maternal educational attainment, residential father's college degree status, and a dummy variable for missing paternal educational attainment.[15] The results, however, remain robust even when control variables are excluded, as the sex of the co-twin is randomly assigned.

Standard errors are clustered at the school level, as substance use can be correlated within

[14]Non-twin siblings with an age spacing of one year or less (30.13% of sibling pairs in the Add Health dataset) are excluded, as their parents may have distinct characteristics and parenting styles. Additionally, short birth intervals are associated with increased risks of adverse fetal and infant outcomes (Schummers et al., 2018). Furthermore, to focus on siblings closely spaced enough to be comparable to twins, those with an age spacing of three years or more (24.79% of sibling pairs) are also excluded.

[15]In my analysis, I have chosen not to include family size as a control variable in the main specification. Family size has been shown to impact substance use among adolescents, indicating a potential effect through this channel. However, considering that the gender composition of older children can also influence parents' fertility decisions, family size becomes an endogenous variable in Equation 1. The findings from subsection 3.1 indicate that the family size indeed does not significantly differ by co-twin sex and thus does not impact my results.

school (Powell et al., 2005; Lundborg, 2006; Clark and Lohéac, 2007; Fletcher, 2012; Fletcher and Ross, 2018). Almost all twin pairs (92 percent) attended the same school and therefore will be in the same cluster. The results are very similar when standard errors are clustered at the family level instead.

Note that for each regression for males (or females), both males (or females) from a same-sex twin pair and one male (or female) from a mixed-sex twin pair are used. In principle, this should not bias the estimates as long as the inclusion into my sample is not correlated with the sex of the co-twin. In subsection 3.4, I show the results remain robust even when applying a weight of 1/2 to those with a same-sex twin sibling or when using only one observation per twin pair, suggesting that this assumption holds.

2.4 Measures of substance use

Add Health includes extensive information about substance use among adolescents. I specifically focus on three substances that exhibit the highest use prevalence among adolescents: Cigarette, alcohol, and marijuana (Volkow, 2011).[16]

To protect the confidentiality of the data and reduce self-reporting bias on sensitive topics including substance use, Add Health interviews were conducted through audio-computer assisted self-interview (ACASI) on laptop computers: Respondents wore headphones, listened to pre-recorded questions, and answered on the laptop by themselves. The questions and answers were not heard or observed by the interviewer or any other people. The ACASI method has been found to improve the quality of self-reporting of sensitive information, making it widely favored and extensively utilized in research (Turner et al., 1998; Kumar et al., 2016).

Despite the adoption of the ACASI method, self-reported measures for substance use may still be subject to measurement errors.[17] Nevertheless, if such a measurement error is not systematically associated with the sex of the co-twin, any potential effects of such error would be nullified

[16] Add Health also examined the usage of other substances, which includes cocaine (including powder, freebase, or crack cocaine) and illegal drugs (such as LSD, PCP, ecstasy, mushrooms, speed, ice, heroin, or pills). Nevertheless, the prevalence of these substances is very low; only 29 respondents (2.16 percent) and 82 respondents (6.12 percent) indicated having ever tried cocaine and illegal drugs, respectively.

[17] For example, some studies suggest that girls are more likely to accurately report their substance use compared to boys in the United States (Siddiqui et al., 1999; Shillington and Clapp, 2000; Johnson and Mott, 2001).

in Equation 1.[18] To investigate an association between misreporting and co-twin sex, I use the question "How honestly have you answered the questions?" (answers range from 1 = "completely honest" to 4 = "not honestly at all") and check the likelihood of answering "completely honest." Table 1.A4 suggests no evidence of differential reporting behaviors based on the sex of the co-twin, which implies that the concerns about measurement errors may not be pronounced in my setting.

My primary focus centers on high school students aged 14 to 18, encompassing Wave I and II of the Add Health dataset. This specific age range is selected for two key reasons. First, considering that siblings generally begin living separately after completing the 12th grade, individuals older than 18 years old are excluded from the sample. Second, as substance consumption tends to be positively associated with age, I exclude individuals who are too young.

The survey construction resulted in two distinct groups of respondents: Those who were aged 14–18 years old in either Wave I or Wave II, and those who were aged 14–18 years old in both waves. For the former group, I analyze data from the wave when they were 14–18. In the case of the latter group, I incorporate data from both waves, applying a weight of 1/2 to each observation.

Table 2 displays the summary statistics for these measures among dizygotic twins. The results indicate a significant gender disparity, with male adolescents demonstrating a higher likelihood of substance consumption compared to their female counterparts across all measures. This finding aligns with previous studies that utilized the Monitoring the Future (MTF), the National Longitudinal Survey of Youth (NLSY), and the Youth Risk Behavior Survey (YRBS) to examine adolescent behavior in the 1990s (Cook et al., 2001; Gruber and Zinman, 2001; Pacula et al., 2001).

3 Effects of Having a Same-Sex Co-Twin

3.1 Exogeneity of co-twin gender

Table 3 presents the randomness tests for the sex of the co-twin among dizygotic twins, using demographic characteristics, parental backgrounds, and school characteristics. Notably, it shows that these variables are hardly correlated with the sex of the co-twin.[19] This finding supports the

[18] See subsection 7.1 for more discussions about self-report measurement errors.

[19] The only exceptions are race indicators for males. The joint F-test for race, however, fails to reject the null hypothesis, with a p-value of 0.12.

identification assumption that the co-twin's sex can be treated as random. Conversely, Table 1.A5 indicates that the sex of a non-twin sibling appears less random compared to that of a co-twin, especially among females.

Furthermore, I test whether the co-twin sex affects the family environment in my samples. Particularly, I examine the influence of co-twin sex on the number of siblings and parental cohabitation, both of which are known to be associated with adolescent substance use (Black et al., 2005; Booth and Kee, 2009; Ermisch and Francesconi, 2001; Gustavsen et al., 2016; Steele et al., 2009). If co-twin sex is systematically related to these factors, certain families may be more likely to be included in my sample, potentially introducing bias. The analysis is performed at the household level, using one observation per household.

Table 4 shows the results for my sample of dizygotic twins. It suggests that the sex of the co-twin does not significantly affect the total number of siblings (column (1)), potentially due to a self-induced reduction in maternal fertility itself following twin births (Record et al., 1978). The table also shows no evidence of differential parental marriage stability based on the sex of the co-twin: No difference in the probability of living with both biological parents until age 5 (column (2)) and the number of marriages or marriage-like relationships of the respondent parent until the survey wave (column (3)). In contrast, Table 1.A6 suggests that in the non-twin full sibling sample, having a same-sex sibling significantly increases the number of siblings, especially for females.

To address concerns about relatively small sample size in Table 4, I further utilize the 1990 Census Public Use Microdata Sample (PUMS), which represents 5 percent of the US population from all states. Within this sample, I identify women who are likely to be mothers of twins[20] and explore their total number of children and marital stability in Table 1.A7, following Dahl and Moretti (2008). All coefficients in Table 1.A7 are statistically insignificant, indicating limited evidence of differential family structure by co-twin's sex.

[20]The mother sample is restricted to women who were between the ages of 18 and 40, had at least one child with the oldest being younger than 18 years old, and were living with all the children they ever reported having delivered. Twins are identified as children in a household who were born between 1976 and 1983 as my Add Health sample and have a sibling with the same birth year, birthplace, and age in years. The analysis excludes triplets and higher-order multiples for simplicity, but cannot distinguish monozygotic twins and dizygotic twins.

The condition of mothers living with all the children having ever delivered is essential to identify twins due to the data limitation. This condition aligns with the prevailing trend of child custody being predominantly awarded to mothers in the 1990s. However, it should be noted that this restriction excludes divorced women who awarded custody of some of their children to the father, which may limit the generalizability of the findings to divorced households.

By comparison, Table 1.A8 shows the results for mothers of non-twin siblings with an age space of 2 years.[21] This table suggests that among mothers of non-twin siblings, having same-sex children is positively associated with the total number of children for both mothers of male children and mothers of female children. It also suggests that having same-sex children is associated with marital instability, especially for mothers of female children. These results are consistent with prior studies on non-twin full siblings (Angrist and Evans, 1998; Dahl and Moretti, 2008).

The comparison between the twin sample and the non-twin full sibling sample suggests that the sex of the co-twin is more random than the sex of the non-twin full sibling and that the twin sample is less likely to suffer from differential family environments than the non-twin sample. It highlights that while both the twins and the non-twin full siblings share 50% of the genes on average, the twin sample is advantageous in overcoming biases stemming from environmental factors, compared to the non-twin full sibling sample.

3.2 Sibling gender effects on substance use

Figure 1 shows the unadjusted substance consumption of the dizygotic twins in my sample. It reveals that males with a twin brother are more likely to be engaged in substance use than males with a twin sister. In contrast, there are no differences observed among females for any of the measures.

Table 5 presents the main estimates on the effects of having a same-sex co-twin, represented by the dummy variable SS in the main specification. The results align with Figure 1. For males, having a brother is associated with a higher probability of consuming substances. For example, having a brother is associated with higher probability of smoking during the last 30 days by 9.9 percent points, drinking alcohol during the last 12 months by 15.5 percent points, and any binge drinking ($\geq$ 5 drinks per day) during the last 12 months by 7.6 percent points.These estimates indicate 30 percent to 33 percent difference compared to the means.

For comparison, I also investigated the sibling gender effect in the non-twin full sibling sample. Table 1.A9 and Table 1.A10 present the impact of a younger sibling's gender on the older sibling

[21] Similar to Table 1.A7, the mother sample is restricted to women between the ages of 18 and 40, with at least one child, the oldest being younger than 18 years old, and living with all the children they ever reported having delivered. Non-twin siblings are identified as children in a household with an age spacing of 2 years, born between 1976 and 1983 as my Add Health sample.

and the older sibling's gender effect on the younger sibling, respectively.

The first approach, examining the impact of a younger sibling's gender on the older sibling, is widely adopted in recent studies (Peter et al., 2018; Cools and Patacchini, 2019; Golsteyn and Magnée, 2020; Brenøe, 2021). This is because assuming the younger sibling's gender to be random given the presence of a younger sibling, the estimate can be interpreted as causal. However, as the impact of the younger sibling is typically marginal in the context of substance use, the estimates in Table 1.A9 are statistically insignificant and closer to zero compared to those in Table 5.

The second approach, studying the impact of an older sibling's gender on the younger sibling, is more straightforward, as older siblings typically have a greater impact on younger siblings regarding substance use. However, it faces challenges from parental fertility selection biases. As a result, the estimates in Table 1.A10 are very noisy, especially in terms of their sign.

3.3 Interpretation of the twin estimates

In my twin sample, the coefficients for males exhibit considerably larger magnitudes than those for females.[22] These negative effects of having a male co-twin on males, with little influence observed among females, are consistent with previous literature on peer effects. Studies have shown that adolescents are more negatively influenced by brothers (compared to sisters) (Cyron et al., 2017; Golsteyn and Magnée, 2020) or male school peers (compared to female peers) (Hoxby, 2000; Clark and Lohéac, 2007; Kremer and Levy, 2008; Carrell and Hoekstra, 2010; Lavy and Schlosser, 2011; Gottfried and Graves, 2014; Hu, 2015; Hill, 2017; Carrell et al., 2018).

Estimates for male twins are large enough to be economically meaningful. Comparisons with the family environment literature highlight the importance of sibling gender, which is on par with other environmental factors. For example, regarding birth order, second-born boys have a 9.0 percent higher likelihood of trying smoking compared to first-born boys (Argys et al., 2006), similar to the 8.4 percent estimate in Table 5.[23] Furthermore, according to See (2016), every extra weekly

[22] Table 1.A11 further checks whether the coefficients for SS among males differ statistically from those among females. The results indicate statistical significance for two measures: Having ever tried marijuana and drinking in the last 12 months. The remaining findings largely align with those in Table 5, whose format is more straightforward and more commonly used in previous studies.

[23] Comparing the estimates in Table 5 with birth order effects on alcohol and marijuana use reveals even larger magnitudes. Fifth-born (or higher) boys exhibit a 10.3 percent higher likelihood of ever drinking alcohol and a 13.4 percent higher likelihood of ever using marijuana compared to first-born boys, while the corresponding estimates in Table 5 are 12.3 percent and 16.0 percent.

hour of engaged activities with the father reduces the chance of current cigarette smoking by 2.0 percent. Translating these results, my estimates for smoking imply that spending approximately 5.0 additional weekly hours with the father would yield similar effects. In addition, drawing on Chalfin and Deza (2018), which identified that each additional year of parental education reduces child binge drinking by 4.1 percent, the estimates for binge drinking are equivalent to a nearly 1.8 years of parental education.

My estimates also reveal substantial impacts in magnitude when comparing them with the effects of other peer influence on substance use. For example, using classmates' parental educational attainment as an instrumental variable, Powell et al. (2005) suggests that moving a high school student from a school where no children smoke to a school where 25 percent of the youths smoke increases the probability of smoking by 14.5 percent points. This implies that my estimate for smoking among male adolescents, 9.9 percent, corresponds to a 17.1 percent points increase in smoking among peers if the effects are proportional. Similarly, by exploiting a variation in alcohol consumption between classes within schools and grades, Lundborg (2006) suggests that a 10 percent point increase in classmates' binge drinking raises own binge drinking by 2.3 percent points. This translates my estimate for binge drinking among male adolescents, which is 7.6 percent points, to a 33.0 percent point increase in classmates' binge drinking. These findings underscore the influential role of co-twins (or close friends), emphasizing their strong bond compared to their classmates or school peers.

The findings presented in Table 5 reveal considerable magnitudes, even when compared to successful policy interventions. For instance, in terms of prices, my estimates for current cigarette smoking and alcohol consumption among male adolescents are approximately equivalent to a $0.98 (1994 dollars) increase per pack of cigarettes or a $1.83 (1993 dollars) increase in the price per six-pack of alcohol, based on the results of Gruber and Zinman (2001) and Cook et al. (2001). Another example is related to an education program, aimed at reducing adolescent substance use, where my estimate for binge drinking corresponds to approximately 40 percent of the effect observed in the intensive education program implemented in a school as studied by Botvin et al. (2001),[24] which resulted in a 57 percent reduction in binge drinking. Furthermore, when comparing my findings

[24]This intervention involves instructing skills to resist alcohol and drug consumption, promoting norms discouraging such behaviors, and enhancing essential personal and social skills.

with the impact of community anti-smoking legislation, my estimate for having ever tried smoking is comparable to approximately 26 percent of the effect achieved through the legislation (Jason et al., 1991).

3.4 Additional analysis

I run a series of robustness checks to confirm the consistency of the results across various specifications. Table 6 presents the estimated coefficients of the dummy variable SS in different specifications for males and females. Columns (1) and (2) provide the estimates without and with control variables, respectively, and the results are similar. Clustering the standard errors at the family level in column (3) of Table 5 shows minimal changes in the significance of the coefficients. To address over-representation of same-sex sibling pairs, column (4) further applies a weight of 1/2 to those with a same-sex twin sibling, and column (5) includes only one respondent per household. The coefficients and their significance remain similar. Additionally, column (6) examines the marginal effects using a probit model, yielding comparable results in terms of magnitude and significance.

To address the family-wise type I over-rejection error associated with testing multiple hypotheses, I employ summary standardized indices that aggregate information across multiple outcome variables, enhancing statistical power (Kling et al., 2007; Anderson, 2008). The summary indices are calculated by averaging the standardized z-scores for each outcome and then re-standardizing the average. Four summary indices are used: An all-substances summary index (comprising all outcome variables in Table Table 5), the cigarette summary index, the marijuana summary index, and the alcohol summary index. The results presented in Table 1.A12 indicate that among males, having a same-sex co-twin is associated with higher scores of substance use in all indices, leading to an increase ranging from 0.23 to 0.28 standard deviations. However, this effect is not observed among females, except in the case of marijuana.

4 Potential Mechanisms

Several mechanisms may underlie the results in subsection 3.2. In this section, I explore three potential mechanisms for sibling gender effects that have been suggested by prior studies: Social interactions between twins, parental investment, and epigenetic influences between twins.

4.1 Direct sibling influences

In this subsection, I explore if and how siblings can directly affect each other. Manski (2000) identifies that economic peer effects can work through one of three mechanisms: Constraints, preferences, or expectations. Here I propose two examples of such effects with suggestive evidence: Providing a network for access to substances (constraints channel) or serving as a reference point for substance consumption (preferences/expectations channel).

4.1.1 Sibling's network

Prior studies suggest that siblings share networks for job search (Peter et al., 2018; Rao and Chatterjee, 2018) and school choice (Dustan, 2018). Similarly, siblings can utilize each other's network for access to substances, especially during adolescence. Sibling gender may play a role in this network if brothers have more friends who use substances than sisters, and if it is easier for an adolescent to access the same-sex sibling's network compared to the different-sex sibling's network.

The predictions for this network channel are clear for males but unclear for females. For males, twin brothers' networks would have a higher number of substance users and be more likely to be utilized than twin sisters' networks. This suggests that twin brothers may contribute to an increase in males' substance consumption. However, for females, the twin sister's network may be easier to utilize but have fewer substance users compared to the twin brother's network. Hence, overall impacts on females' substance use is less clear.

Ideally, examining where or from whom adolescents obtain substances and whether these patterns differ by sibling gender would provide a more direct understanding of this dynamic. Unfortunately however, such a question is not available in Add Health. Instead, I provide suggestive evidence concerning the two elements of this channel: The higher prevalence of substance-using friends among brothers than sisters, and the ease of accessing same-sex sibling networks relative to different-sex sibling networks.

First, Table 7 utilizes three survey questions "Of your three best friends, how many smoke at least 1 cigarette a day?" Of your three best friends, how many drink alcohol at least once a month?" and Of your three best friends, how many use marijuana at least once a month?". The results show that female adolescents report having fewer best friends who engage in these behaviors. Additionally, females are less likely to have male friends and are more likely to have female friends, which suggests that their friends would also have fewer substance users in their network.

Second, I provide suggestive evidence on whether adolescents would find it easier to utilize the same-sex co-twin's network than the different-sex co-twin's. The first column of Table 8 examines the likelihood of having a friend in common, based on self-reported lists of five best friends. This variable is coded as 1 if any common friend is reported within a twin pair and 0 otherwise. Siblings could not be listed as friends, and if a respondent mistakenly included their sibling, I excluded the sibling from the friend list.[25] The results indicate that for both males and females, having a same-sex co-twin is associated with a higher probability of sharing a friend. Additionally, Table 8 suggests that twins spend more time with their co-twin (column (2)) and with their co-twin and their common friends together (column (3)) [26] if they are the same-sex than if they are different-sex.

Third, I investigate whether the gender of the co-twin not only affects the co-twin's friend network but also one's *own* friend network. The first five columns in Table 9 suggest that the gender of the co-twin is related to the respondent's friend network as well. For males, having a same-sex co-twin is linked to a 16.9 percentage point higher likelihood of having friends who smoke and 17.9 percentage points for friends who drink alcohol. For females, the gender of the co-twin is not directly related to own number of substance-using best friends but is associated with the likelihood of having at least one male friend.

The findings for the non-twin full sibling sample sharply contrast with those from the twin sample. Same-sex non-twin siblings are more likely to share common friends and spend time together than opposite-sex siblings, but these differences are much smaller than in the twin sample.

[25]However, this measure may not be the most reliable for studying friendships within the twin sample due to its relatively small size. Friends attending non-sampled schools could not be identified, significantly reducing the sample size. Despite this limitation, the measure still offers suggestive evidence on the accessibility of a sibling's social network.

[26]The outcome variable in column (2) is a dummy for whether the respondent chose the answer 'a lot' to the question "How much time do you and {Sibling} spend together?" among the choices 'a lot', 'some', 'little', and 'none.' Similarly, the outcome variable in column (3) is a dummy for whether the respondent chose the answer 'a lot' to the question "How much time do you and {Sibling} spend with the same friend or group of friends?" among the choices 'a lot', 'some', 'little', and 'none.'

For example, among females, same-sex non-twin siblings are 18.8 percentage points more likely to share friends among older siblings (Table 1.A13) and 18.9 percentage points among younger siblings (Table 1.A14), which is less than half the 45.5 percentage point difference seen in twins (Table 8). Moreover, the sample means for these variables in the non-twin sample are notably lower than in the twin sample. Among females, the likelihood of having a common friend is 28.1% among older siblings and 31.3% among younger siblings, about half of the 55.6% observed among twins, indicating less network utilization among non-twin siblings.

In line with the lower network utilization among non-twin full siblings, the sibling gender effect on the own friend network also presents a notable contrast to the twin sample. For male non-twins, the gender of a younger sibling no longer significantly affects the older sibling's friend network (Table 1.A15), nor does the older sibling's gender affect the younger sibling's friend network (Table 1.A16). For female non-twins, a sister's network might be as accessible or easier to access, and includes fewer substance users compared to a brother's network. Therefore, having a sister is either negatively correlated (Table 1.A15) or shows no correlation (Table 1.A16) with the number of substance-using friends. These contrasts between the twin sample and the non-twin sample further support the channel of sibling network.

4.1.2 Siblings as reference points

Another way that siblings can directly influence each other is by serving as reference points, comparing themselves to each other and shaping each other's preferences and behaviors (Schmitt, 1972).[27] Peter et al. (2018) propose two reasons why sibling gender matters as a reference point. First, for outcomes with a notable gender gap, brothers and sisters represent different reference points. Second, same-sex siblings serve as more salient reference points compared to different-sex siblings, in particular, if they consist of boys (Conley, 2000; Grose, 2021). This argument is also consistent with the literature, which reports stronger spillover effects among same-sex siblings than among different-sex siblings in various outcome variables (Eriksson et al., 2016; Joensen and Nielsen, 2018; Nicoletti and Rabe, 2019; Dahl et al., 2020; Gurantz et al., 2020; Bingley et al., 2021).

[27] For example, studies have found that a higher income than siblings is associated with higher life satisfaction (Kuegler, 2009), educational spillovers are driven by sibling rivalry (Joensen and Nielsen, 2018), and food consumption habits may be influenced by siblings (Farrell and Shields, 2001)

In the context of substance use, observing a sibling engaging in substance use may reduce hesitancy or feelings of guilt among adolescents.[28] The predictions align with those detailed in the network channel. For males, they are more willing to use substances when their twin brothers do compared to when twin sisters do, and indeed, twin brothers are more likely to use substances than twin sisters. As a result, the presence of twin brothers might contribute to an increase in substance consumption among males. Conversely, for females, they are less reluctant to use substances when their twin sisters do than when their twin brothers do, but twin sisters are less likely to use substances than twin brothers. Consequently, the overall impact is less clear for females.

A direct approach to assess this channel would involve investigating the extent to which an adolescent compares themself with their twin siblings, particularly concerning substance use, and whether this tendency varies based on the gender of the sibling. However, such detailed information is hardly available in datasets. Instead, two exercises can provide indicative evidence addressing the two dimensions of this channel.

First, brothers tend to be a more significant reference point for substance consumption due to the observed higher likelihood of substance use among males, not only in Table 2 but also in previous research (Cook et al., 2001; Gruber and Zinman, 2001; Pacula et al., 2001). Second, under the assumption that the amount of time spent together is associated with the saliency of the reference point, the final two columns of Table Table 8 indicate that same-sex twin siblings hold more influence as reference points. (Similarly, for non-twin full siblings, same-sex siblings serve as a reference point, which is as salient as or more salient than different-sex siblings, based on Table 1.A13 and Table 1.A14.) These predictions consistently align with my findings for twins in Table 5 (and for non-twins in Table 1.A9 and Table 1.A10), as well as prior studies such as Peter et al. (2018) and Joensen and Nielsen (2018).[29]

[28] Conversely, adolescents might also be inclined to cease substance consumption if their siblings discontinue such behavior. However, as highlighted by Harris and López-Valcárcel (2008), social influences tend to be asymmetric. Specifically, each smoking sibling in a household is linked to a 7.6 percent increase in the probability of smoking, while each non-smoking sibling is associated with a 3.5 percent reduction in that probability.

[29] Peter et al. (2018) find that men earn more when they have a brother (while the effect of having a brother on earnings is small or insignificant for women) and women give birth earlier when they have a sister. They argue that these findings are driven by stronger competition between same-sex siblings and differential level of reference points by sibling gender. Joensen and Nielsen (2018) also find larger peer effects on education for brothers, supporting the notion of heightened competition between male siblings.

4.2 Parental investment

Sibling gender may indirectly affect adolescents through differential parental treatment (Lundberg et al., 2007; Price, 2008; Mammen, 2011).

I examine three categories of parental investment—time spent with children, allowances, and preventive care—following Cools and Patacchini (2019). Columns (1) and (2) investigate parental time investment, which reflects the overall extent of parental involvement and supervision.[30] Column (3) focuses on the weekly allowance in dollars given to adolescents, a factor that could impact adolescent substance use due to their limited budget for purchasing substances. Columns (4) and (5) explore two preventive care measures from the previous year: Physical and dental check-ups. These measures, albeit not comprehensive, may offer insights into parental concern for child health and possibly their attitudes towards children's substance use. The findings suggest that sibling gender has minimal impact on parental investment, at least in terms of the measures in Table 10.

Table 1.A17 and Table 1.A18 perform the same analysis for non-twin siblings. Unlike prior studies (Cools and Patacchini, 2019; Brenøe, 2021), nearly all estimates in these tables are statistically insignificant, possibly due to the small sample size. Nevertheless, most of these estimates are larger than those in Table 10, implying that potential differences in parental investment by co-twin sex among dizygotic twins, if any, would not exceed the differences by sibling sex among non-twin full siblings. This indicates that considering the pronounced differences in sibling gender effects between twins and non-twins, parental investment does not seem to be the driving mechanism underlying the substantial sibling gender effects observed among male twins.

4.3 Epigenetic influence

Compared to non-twin siblings who only encounter each other after birth, twins are exposed to one another both in utero and after birth. Therefore, it is possible that any biological influences shared between twins in utero could impact their future outcomes.

[30] The outcome variables are defined as the total number of activities that a respondent had with their father and mother, respectively, in the past 4 weeks, ranging from 0 to 10 (going shopping; playing a sport; going to a religious service or church-related event; talking about someone the respondent is dating or a party the respondent went to; going to a movie, play, museum, or concert, or sports event; talking about a personal problem the respondent was having; having a serious argument about the behavior of the respondent; talking about the respondent's school work or grades; working on a project for school; talking about other things the respondent is doing in school)

One possibility, primarily supported by animal studies, is the twin testosterone transfer hypothesis (Clemens, 1974; Vom Saal, 1989). This hypothesis suggests that testosterone, crucial for the sexual differentiation of the male fetus, might be transferred between twin fetuses within the same pregnancy and consequently, females with male co-twins could be exposed to elevated testosterone levels than females with female co-twins (Tapp et al., 2011). While evidence from human twin studies is still limited and inconsistent, some findings suggest that such prenatal testosterone exposure may lead to more masculine cognitive and behavioral traits among females from different-sex pairs compared to those from same-sex pairs (Bütikofer et al., 2019; Cronqvist et al., 2016).

However, testosterone transfer in utero does not appear to be a relevant mechanism in my findings, even if such masculine traits might include substance use during adolescence. If it were the underlying mechanism, we would expect no significant differences in substance use between males with male co-twins and those with female co-twins. Conversely, a higher likelihood of substance use would be anticipated among females with male co-twins compared to those with female co-twins. Such expectations are inconsistent with the findings in Table 5 for both genders in terms of both the direction and statistical significance.

5 Conclusion

This study explores the effects of having a same-sex sibling on substance use during adolescence by utilizing a dataset of dizygotic twins in the United States. The twin design offers several advantages for studying the causal impact of having a same-sex sibling on substance use. First, the sex of a co-twin is, in nature, randomly assigned. Second, this approach overcomes potential selection biases stemming from parental fertility decisions influenced by older children's gender. Third, twins share birth age, timing, and environment at each stage, minimizing the sibling birth order effects and circumventing complex interactions between birth order and sibling gender effects. Lastly, since twins cannot choose each other as peers, yet they view each other as their friends, this setup enables studying substance consumption among "friends" without being influenced by selection biases.

Among males, I observe that having a twin brother increases the probability of smoking, drinking alcohol, and using marijuana. In contrast, among females, there is no significant difference in substance use between those with twin brothers and those with twin sisters. An examination of potential mechanisms indicates that these findings align with direct sibling interactions and are hardly driven by parental investment and family structure. I further propose two examples of how siblings directly influence each other: Sharing a network for substance access and serving as a reference point for substance use.

The magnitudes of the estimates, especially for male adolescents, are sizable. The comparison with the estimates in the prior studies suggests that siblings are one of the most influential factors in adolescent substance use. The results underscore the crucial role of siblings and influential peers in shaping substance-related beliefs and consumption behaviors.

References

ALMOND, D. AND L. EDLUND (2008): “Son-biased sex ratios in the 2000 United States Census,” *Proceedings of the National Academy of Sciences*, 105, 5681–5682.

ANDERSON, M. L. (2008): “Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects,” *Journal of the American Statistical Association*, 103, 1481–1495.

ANELLI, M. AND G. PERI (2015): “Gender of siblings and choice of college major,” *CESifo Economic Studies*, 61, 53–71.

ANGRIST, J. D. AND W. N. EVANS (1998): “Children and their parents’ labor supply: evidence from exogenous variation in family size,” *American Economic Review*, 88, 450–477.

ANTHONY, L. G., B. J. ANTHONY, D. N. GLANVILLE, D. Q. NAIMAN, C. WAANDERS, AND S. SHAFFER (2005): “The relationships between parenting stress, parenting behaviour and preschoolers’ social competence and behaviour problems in the classroom,” *Infant and Child Development: An International Journal of Research and Practice*, 14, 133–154.

ARGYS, L. M., D. I. REES, S. L. AVERETT, AND B. WITOONCHART (2006): “Birth order and risky adolescent behavior,” *Economic Inquiry*, 44, 215–233.

AVERETT, S. L., L. M. ARGYS, AND D. I. REES (2011): “Older siblings and adolescent risky behavior: does parenting play a role?” *Journal of Population Economics*, 24, 957–978.

BALSA, A. I., L. M. GIULIANO, AND M. T. FRENCH (2011): “The effects of alcohol use on academic achievement in high school,” *Economics of Education Review*, 30, 1–15.

BARNHART, K. T. (2013): “Assisted reproductive technologies and perinatal morbidity: interrogating the association,” *Fertility and Sterility*, 99, 299–302.

BECK, C. T. (2002): “Releasing the pause button: mothering twins during the first year of life,” *Qualitative Health Research*, 12, 593–608.

BINGLEY, P., P. LUNDBORG, AND S. V. LYK-JENSEN (2021): “Brothers in arms spillovers from a draft lottery,” *Journal of Human Resources*, 56, 225–268.

BLACK, S. E., S. BREINING, D. N. FIGLIO, J. GURYAN, K. KARBOWNIK, H. S. NIELSEN, J. ROTH, AND M. SIMONSEN (2021): “Sibling spillovers,” *The Economic Journal*, 131, 101–128.

BLACK, S. E., P. J. DEVEREUX, AND K. G. SALVANES (2005): “The more the merrier? The effect of family size and birth order on children’s education,” *The Quarterly Journal of Economics*, 120, 669–700.

BOLCH, C. E., P. G. DAVIS, M. P. UMSTAD, AND J. R. FISHER (2012): “Multiple birth families with children with special needs: A qualitative investigation of mothers’ experiences,” *Twin Research and Human Genetics*, 15, 503–515.

Booth, A. L. and H. J. Kee (2009): "Birth order matters: the effect of family size and birth order on educational attainment," *Journal of Population Economics*, 22, 367–397.

Botvin, G. J., K. W. Griffin, T. Diaz, and M. Ifill-Williams (2001): "Preventing binge drinking during early adolescence: one-and two-year follow-up of a school-based preventive intervention." *Psychology of Addictive Behaviors*, 15, 360.

Brenøe, A. A. (2021): "Brothers increase women's gender conformity," *Journal of Population Economics*, 1–38.

Brenøe, A. A. and U. Zölitz (2020): "Exposure to more female peers widens the gender gap in STEM participation," *Journal of Labor Economics*, 38, 1009–1054.

Brunello, G. and M. De Paola (2013): "Leadership at school: does the gender of siblings matter?" *Economics Letters*, 120, 61–64.

Butcher, K. F. and A. Case (1994): "The effect of sibling sex composition on women's education and earnings," *Quarterly Journal of Economics*, 109, 531–563.

Bütikofer, A., D. N. Figlio, K. Karbownik, C. W. Kuzawa, and K. G. Salvanes (2019): "Evidence that prenatal testosterone transfer from male twins reduces the fertility and socioeconomic success of their female co-twins," *Proceedings of the National Academy of Sciences*, 116, 6749–6753.

Carrell, S. E., M. Hoekstra, and E. Kuka (2018): "The long-run effects of disruptive peers," *American Economic Review*, 108, 3377–3415.

Carrell, S. E. and M. L. Hoekstra (2010): "Externalities in the classroom: How children exposed to domestic violence affect everyone's kids," *American Economic Journal: Applied Economics*, 2, 211–28.

Chalfin, A. and M. Deza (2018): "The effect of parental education on children's drug and alcohol use," *AEA Papers and Proceedings*, 108, 373–378.

Chalfin, A., B. Hansen, and R. Ryley (2023): "The minimum legal drinking age and crime victimization," *Journal of Human Resources*, 58, 1141–1177.

Chatterji, P. (2006a): "Does alcohol use during high school affect educational attainment?: evidence from the National Education Longitudinal Study," *Economics of Education Review*, 25, 482–497.

——— (2006b): "Illicit drug use and educational attainment," *Health Economics*, 15, 489–511.

Citro, B., J. Gilson, S. Kalantry, K. Stricker, et al. (2014): "Replacing myths with facts: sex-selective abortion laws in the United States," *Cornell Law Faculty Publications*, 1399.

Claes, M. (1998): "Adolescents' closeness with parents, siblings, and friends in three countries: Canada, Belgium, and Italy," *Journal of Youth and Adolescence*, 27, 165–184.

CLARK, A. E. AND Y. LOHÉAC (2007): ""It wasn't me, it was them!" Social influence in risky behavior by adolescents," *Journal of Health Economics*, 26, 763–784.

CLEMENS, L. G. (1974): "Neurohormonal control of male sexual behavior," in *Reproductive Behavior*, Springer, 23–53.

CONLEY, D. (2000): "Sibship sex composition: effects on educational attainment," *Social Science Research*, 29, 441–457.

COOK, P. J. AND M. J. MOORE (1993): "Drinking and schooling," *Journal of Health Economics*, 12, 411–429.

COOK, P. J., M. J. MOORE, ET AL. (2001): "Environment and persistence in youthful drinking patterns," in *Risky Behavior Among Youths: An Economic Analysis*, ed. by J. Gruber, University of Chicago Press Chicago, chap. 10, 375–437.

COOLS, A. AND E. PATACCHINI (2019): "The brother earnings penalty," *Labour Economics*, 58, 37–51.

CRONQVIST, H., A. PREVITERO, S. SIEGEL, AND R. E. WHITE (2016): "The fetal origins hypothesis in finance: Prenatal environment, the gender gap, and investor behavior," *Review of Financial Studies*, 29, 739–786.

CYRON, L., G. SCHWERDT, AND M. VIARENGO (2017): "The effect of opposite sex siblings on cognitive and noncognitive skills in early childhood," *Applied Economics Letters*, 24, 1369–1373.

DAHL, G. B., K. V. LØKEN, AND M. MOGSTAD (2014): "Peer effects in program participation," *American Economic Review*, 104, 2049–2074.

DAHL, G. B. AND E. MORETTI (2008): "The demand for sons," *Review of Economic Studies*, 75, 1085–1120.

DAHL, G. B., D.-O. ROOTH, AND A. STENBERG (2020): "Family spillovers in field of study," NBER Working Paper 27618.

DAMATO, E. G. AND C. BURANT (2008): "Sleep patterns and fatigue in parents of twins," *Journal of Obstetric, Gynecologic & Neonatal Nursing*, 37, 738–749.

DEE, T. S. AND W. N. EVANS (2003): "Teen drinking and educational attainment: evidence from two-sample instrumental variables estimates," *Journal of Labor Economics*, 21, 178–209.

DESIMONE, J. (2010): "Drinking and academic performance in high school," *Applied Economics*, 42, 1481–1497.

DUSTAN, A. (2018): "Family networks and school choice," *Journal of Development Economics*, 134, 372–391.

ELLICKSON, P. L., J. S. TUCKER, AND D. J. KLEIN (2001): "High-risk behaviors associated with early smoking: results from a 5-year follow-up," *Journal of Adolescent Health*, 28, 465–473.

ERIKSSON, K. H., R. HJALMARSSON, M. J. LINDQUIST, AND A. SANDBERG (2016): "The importance of family background and neighborhood effects as determinants of crime," *Journal of Population Economics*, 29, 219–262.

ERMISCH, J. F. AND M. FRANCESCONI (2001): "Family structure and children's achievements," *Journal of Population Economics*, 14, 249–270.

FARRELL, L. AND M. A. SHIELDS (2001): "Child expenditure: The role of working mothers, lone parents, sibling composition and household provision," *IZA Discussion Paper 388*.

FELLMAN, J. AND A. W. ERIKSSON (2006): "Weinberg's differential rule reconsidered," *Human Biology*, 78, 253–275.

FELSON, J. (2014): "What can we learn from twin studies? A comprehensive evaluation of the equal environments assumption," *Social Science Research*, 43, 184–199.

FLETCHER, J. M. (2012): "Peer influences on adolescent alcohol consumption: evidence using an instrumental variables/fixed effect approach," *Journal of Population Economics*, 25, 1265–1286.

FLETCHER, J. M. AND S. L. ROSS (2018): "Estimating the effects of friends on health behaviors of adolescents," *Health Economics*, 27, 1450–1483.

FRALEY, R. C. AND C. M. TANCREDY (2012): "Twin and sibling attachment in a nationally representative sample," *Personality and Social Psychology Bulletin*, 38, 308–316.

GAVIRIA, A. AND S. RAPHAEL (2001): "School-based peer effects and juvenile behavior," *Review of Economics and Statistics*, 83, 257–268.

GOLSTEYN, B. H. AND C. A. MAGNÉE (2017): "Does birth spacing affect personality?" *Journal of Economic Psychology*, 60, 92–108.

——— (2020): "Does sibling gender affect personality traits?" *Economics of Education Review*, 77, 102016.

GOTTFRIED, M. A. AND J. GRAVES (2014): "Peer effects and policy: The relationship between classroom gender composition and student achievement in early elementary school," *The BE Journal of Economic Analysis & Policy*, 14, 937–977.

GROSE, M. (2021): *Why first-borns rule the world and later-borns want to change it*, Random House Australia.

GRUBER, J. (2009): *Risky Behavior Among Youths: An Economic Analysis*, University of Chicago Press.

GRUBER, J. AND J. ZINMAN (2001): "Youth smoking in the United States: evidence and implications," in *Risky Behavior Among Youths: An Economic Analysis*, ed. by J. Gruber, University of Chicago Press Chicago, chap. 2, 69–120.

GURANTZ, O., M. HURWITZ, AND J. SMITH (2020): "Sibling effects on high school exam taking and performance," *Journal of Economic Behavior & Organization*, 178, 534–549.

GUSTAVSEN, G. W., R. M. NAYGA, AND X. WU (2016): "Effects of parental divorce on teenage children's risk behaviors: incidence and persistence," *Journal of Family and Economic Issues*, 37, 474–487.

HARRIS, J. E. AND B. G. LÓPEZ-VALCÁRCEL (2008): "Asymmetric peer effects in the analysis of cigarette smoking among young people in the United States, 1992–1999," *Journal of Health Economics*, 27, 249–264.

HARRIS, K. M. (2013): *The Add Health study: Design and Accomplishments*, Carolina Population Center.

HAUSER, R. M. AND H.-H. D. KUO (1998): "Does the gender composition of sibships affect women's educational attainment?" *Journal of Human Resources*, 644–657.

HEINONEN, K. (2016): "Supporting multiple birth families at home," *International Journal of Caring Sciences*, 9, 422–431.

HELLER, R., D. O'CONNELL, D. ROBERTS, J. ALLEN, J. KNAPP, P. STEELE, D. SILOVE, D. RAO, AND G. VOGLER (1988): "Lifestyle factors in monozygotic and dizygotic twins," *Genetic Epidemiology*, 5, 311–321.

HILL, A. J. (2015): "The girl next door: The effect of opposite gender friends on high school achievement," *American Economic Journal: Applied Economics*, 7, 147–177.

——— (2017): "The positive influence of female college students on their male peers," *Labour Economics*, 44, 151–160.

HOLDITCH-DAVIS, D., D. ROBERTS, AND M. SANDELOWSKI (1999): "Early parental interactions with and perceptions of multiple birth infants," *Journal of Advanced Nursing*, 30, 200–210.

HOXBY, C. M. (2000): "Peer effects in the classroom: Learning from gender and race variation," NBER Working Paper 7867.

HU, F. (2015): "Do girl peers improve your academic performance?" *Economics Letters*, 137, 54–58.

HUSBY, H., N. HOLM, A. GERNOW, S. THOMSEN, K. KOCK, AND H. GÜRTLER (1991): "Zygosity, placental membranes and Weinberg's rule in a Danish consecutive twin series," *Acta Geneticae Medicae et Gemellologiae: Twin Research*, 40, 147–152.

JASON, L. A., P. Y. JI, M. D. ANES, AND S. H. BIRKHEAD (1991): "Active enforcement of cigarette control laws in the prevention of cigarette sales to minors," *Journal of the American Medical Association*, 266, 3159–3161.

JOENSEN, J. S. AND H. S. NIELSEN (2018): "Spillovers in education choice," *Journal of Public Economics*, 157, 158–183.

JOHNSON, T. P. AND J. A. MOTT (2001): "The reliability of self-reported age of onset of tobacco, alcohol and illicit drug use," *Addiction*, 96, 1187–1198.

KAESTNER, R. (1997): "Are brothers really better? Sibling sex composition and educational achievement revisited." *Journal of Human Resources*, 32, 250–84.

KANAZAWA, S. AND N. L. SEGAL (2019): "Do monozygotic twins have higher genetic quality than dizygotic twins and singletons? Hints from attractiveness ratings and self-reported health," *Evolutionary Biology*, 46, 164–169.

KLING, J. R., J. B. LIEBMAN, AND L. F. KATZ (2007): "Experimental analysis of neighborhood effects," *Econometrica*, 75, 83–119.

KREMER, M. AND D. LEVY (2008): "Peer effects and alcohol use among college students," *Journal of Economic Perspectives*, 22, 189–206.

KUEGLER, A. (2009): "A curse of comparison? Evidence on reference groups for relative income concerns," *World Bank Policy Research Working Paper 4820*.

KUMAR, P. C., C. M. CLELAND, M. N. GOUREVITCH, J. ROTROSEN, S. STRAUSS, L. RUSSELL, AND J. MCNEELY (2016): "Accuracy of the audio computer assisted self interview version of the alcohol, smoking and substance involvement screening test (ACASI ASSIST) for identifying unhealthy substance use and substance use disorders in primary care patients," *Drug and Alcohol Dependence*, 165, 38–44.

LAVY, V. AND A. SCHLOSSER (2011): "Mechanisms and impacts of gender peer effects at school," *American Economic Journal: Applied Economics*, 3, 1–33.

LEE, E. AND G. SCHNORR (Forthcoming): "Am I My Brother's Barkeeper? Sibling Spillovers in Alcohol Consumption at the Minimum Legal Drinking Age," *American Journal of Health Economics*.

LEE, S., L. J. TURNER, S. WOO, AND K. KIM (2014): "All or nothing? The impact of school and classroom gender composition on effort and academic achievement," NBER Working Paper 20722.

LEONARD, L. G. AND J. DENTON (2006): "Preparation for parenting multiple birth children," *Early Human Development*, 82, 371–378.

LIN, M.-J., J.-T. LIU, AND N. QIAN (2014): "More missing women, fewer dying girls: The impact of sex-selective abortion on sex at birth and relative female mortality in Taiwan," *Journal of the European Economic Association*, 12, 899–926.

LUNDBERG, S., S. W. PABILONIA, AND J. WARD-BATTS (2007): "Time allocation of parents and investments in sons and daughters," https://www.researchgate.net/publication/346550748.

LUNDBORG, P. (2006): "Having the wrong friends? Peer effects in adolescent substance use," *Journal of Health Economics*, 25, 214–233.

MACDONALD, Z. AND S. PUDNEY (2000): "The wages of sin? Illegal drug use and the labour market," *Labour*, 14, 657–673.

MACGILLIVRAY, I., M. SAMPHIER, J. LITTLE, M. IAN, M. DORRIS, AND T. BARBARA (1988): "Factors affecting twinning," *Twinning and Twins*, 1, 67–97.

MAMMEN, K. (2011): "Fathers' time investments in children: do sons get more?" *Journal of Population Economics*, 24, 839–871.

MANSKI, C. F. (2000): "Economic analysis of social interactions," *Journal of Economic Perspectives*, 14, 115–136.

MCGUIRE, S. AND N. L. SEGAL (2013): "Peer network overlap in twin, sibling, and friend dyads," *Child Development*, 84, 500–511.

NATIONAL CENTER ON ADDICTION AND SUBSTANCE ABUSE (2011): *Adolescent Substance Use: America's# 1 Public Health Problem*, CASA New York.

NICOLETTI, C. AND B. RABE (2019): "Sibling spillover effects in school achievement," *Journal of Applied Econometrics*, 34, 482–501.

OGUZOGLU, U. AND S. OZBEKLIK (2016): "Like father, like daughter (unless there is a son): Sibling sex composition and women's stem major choice in college," IZA Discussion Paper 10052.

OLESZCZUK, J. J., A. CERVANTES, J. L. KIELY, D. M. KEITH, AND L. G. KEITH (2001): "Maternal race/ethnicity and twinning rates in the United States, 1989-1991." *Journal of Reproductive Medicine*, 46, 550–557.

PACULA, R. L., M. GROSSMAN, F. J. CHALOUPKA, P. M. O'MALLEY, L. D. JOHNSTON, M. C. FARRELLY, ET AL. (2001): "Marijuana and youth," in *Risky Behavior Among Youths: An Economic Analysis*, ed. by J. Gruber, University of Chicago Press Chicago, chap. 6, 271–326.

PETER, N., P. LUNDBORG, S. MIKKELSEN, AND D. WEBBINK (2018): "The effect of a sibling's gender on earnings and family formation," *Labour Economics*, 54, 61–78.

POULIN, F. AND S. PEDERSEN (2007): "Developmental changes in gender composition of friendship networks in adolescent girls and boys." *Developmental Psychology*, 43, 1484.

POWELL, L. M., J. A. TAURAS, AND H. ROSS (2005): "The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior," *Journal of Health Economics*, 24, 950–968.

PRICE, J. (2008): "Parent-child quality time does birth order matter?" *Journal of Human Resources*, 43, 240–265.

RAO, N. AND T. CHATTERJEE (2018): “Sibling gender and wage differences,” *Applied Economics*, 50, 1725–1745.

RECORD, R., E. ARMSTRONG, AND R. LANCASHIRE (1978): “A study of the fertility of mothers of twins.” *Journal of Epidemiology & Community Health*, 32, 183–189.

ROWE, D. C. AND B. L. GULLEY (1992): “Sibling effects on substance use and delinquency,” *Criminology*, 30, 217–234.

SCHMITT, R. L. (1972): *The Reference Other Orientation: An Extension of the Reference Group Concept*, Carbondale: Southern Illinois University Press.

SCHUMMERS, L., J. A. HUTCHEON, S. HERNANDEZ-DIAZ, P. L. WILLIAMS, M. R. HACKER, T. J. VANDERWEELE, AND W. V. NORMAN (2018): “Association of short interpregnancy interval with pregnancy outcomes according to maternal age,” *JAMA internal medicine*, 178, 1661–1670.

SEE, S. G. (2016): “Parental supervision and adolescent risky behaviors,” *Review of Economics of the Household*, 14, 185–206.

SEN, B., S. AVERETT, L. ARGYS, AND D. I. REES (2009): “The effect of substance use on the delinquent behaviour of adolescents,” *Applied Economics Letters*, 16, 1721–1729.

SHARROW, D. J. AND J. J. ANDERSON (2016): “A twin protection effect? Explaining twin survival advantages with a two-process mortality model,” *PloS One*, 11, e0154774.

SHILLINGTON, A. M. AND J. D. CLAPP (2000): “Self-report stability of adolescent substance use: are there differences for gender, ethnicity and age?” *Drug and Alcohol Dependence*, 60, 19–27.

SIDDIQUI, O., J. A. MOTT, T. L. ANDERSON, AND B. R. FLAY (1999): “Characteristics of inconsistent respondents who have “ever used” drugs in a school-based sample,” *Substance Use & Misuse*, 34, 269–295.

STEELE, F., W. SIGLE-RUSHTON, AND Ø. KRAVDAL (2009): “Consequences of family disruption on children’s educational outcomes in Norway,” *Demography*, 46, 553–574.

STEVENSON, M. (2017): “Breaking bad: Mechanisms of social influence and the path to criminality in juvenile jails,” *Review of Economics and Statistics*, 99, 824–838.

TAPP, A. L., M. T. MAYBERY, AND A. J. WHITEHOUSE (2011): “Evaluating the twin testosterone transfer hypothesis: A review of the empirical evidence,” *Hormones and Behavior*, 60, 713–722.

TURNER, C. F., B. H. FORSYTH, J. M. O’REILLY, P. C. COOLEY, T. K. SMITH, S. M. ROGERS, AND H. G. MILLER (1998): “Automated self-interviewing and the survey measurement of sensitive behaviors,” *Computer Assisted Survey Information Collection*, 455–473.

VAN OURS, J. C. (2004): “A pint a day raises a man’s pay; but smoking blows that gain away,” *Journal of Health Economics*, 23, 863–886.

Vlietinck, R., C. Derom, R. Derom, H. Van den Berghe, and M. Thiery (1988): "The validity of Weinberg's rule in the East Flanders prospective twin survey (EFPTS)," *Acta Geneticae Medicae et Gemellologiae: Twin Research*, 37, 137–141.

Volkow, N. D. (2011): *Principles of Drug Addiction Treatment: A Research-Based Guide*, vol. 12, DIANE Publishing.

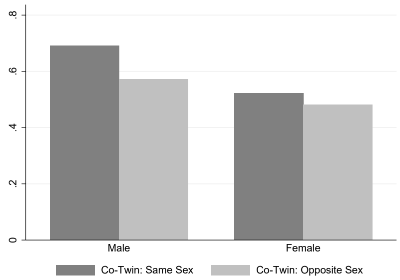
Vom Saal, F. (1989): "Sexual differentiation in litter-bearing mammals: Influence of sex of adjacent fetuses in utero," *Journal of Animal Science*, 67, 1824–1840.

Yamaguchi, K. (1989): "A formal theory for male-preferring stopping rules of childbearing: Sex differences in birth order and in the number of siblings," *Demography*, 26, 451–465.

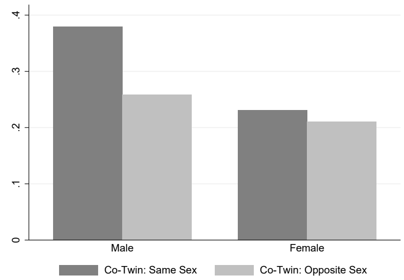
6 Figures and Tables

Figure 1. Substance Use by Sibling Gender Composition Among Dizygotic Twins

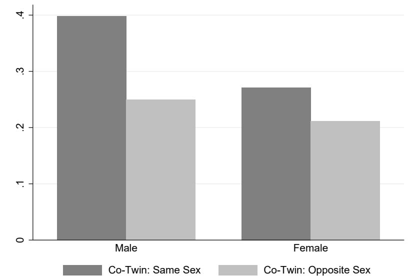
(a) Ever tried smoking



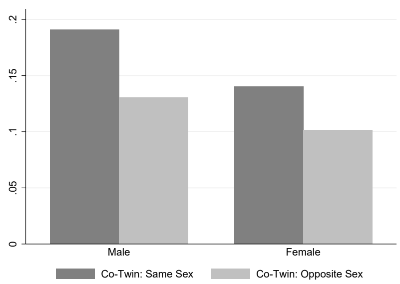
(b) Smoked during the last 30 days



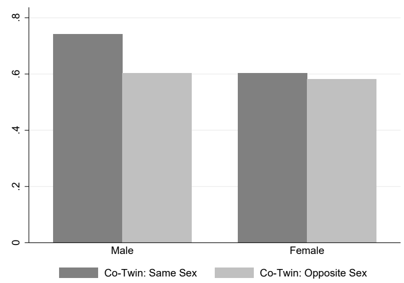
(c) Ever tried marijuana



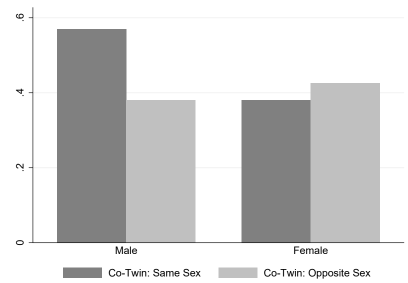
(d) Used marijuana in the last 30 days



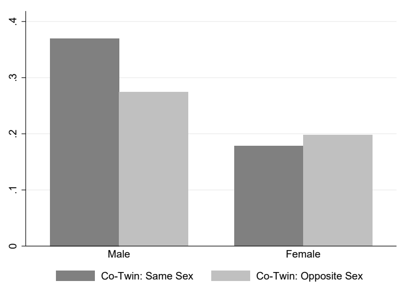
(e) Ever tried alcohol



(f) Drunk during the last 12 months



(g) Any binge drinking during the last 12 months



Notes: Each bar represents the average value of the corresponding substance use by own gender and sibling gender in my sample of dizygotic twins.

Table 1. Number of Dizygotic Twin Respondents

	Male	Female	**Total**
Have a different-sex twin	180	180	360
Have a same-sex twin	240	208	448
Total	420	388	808

Notes: The twin sample consists of all respondents who were identified as dizygotic twins by Add Health and reported each other in their household roster at Wave I.

Table 2. Summary Statistics of Substance Use Among Dizygotic Twins

	Male	Female	Male-Female
Variables	Mean (SD)	Mean (SD)	Diff (SE)
Cigarette			
Ever tried smoking	0.641	0.503	0.137***
	(0.274)	(0.314)	(0.037)
Smoked during the last 30 days	0.328	0.221	0.107***
	(0.249)	(0.225)	(0.030)
Marijuana			
Ever tried marijuana	0.335	0.244	0.091***
	(0.230)	(0.258)	(0.030)
Used marijuana during the last 30 days	0.165	0.123	0.043*
	(0.177)	(0.201)	(0.024)
Alcohol			
Ever tried alcohol	0.683	0.594	0.089**
	(0.299)	(0.339)	(0.040)
Drank during the last 12 months	0.490	0.401	0.089**
	(0.294)	(0.276)	(0.034)
Any binge drinking during the last 12 months	0.330	0.187	0.142***
	(0.274)	(0.217)	(0.031)
N	686	656	1342

Notes: This table shows descriptive statistics for my sample of dizygotic twins in the Add Health dataset by gender and t-tests of differences in means by gender (columns (3)). Means are weighted by the reciprocal of the observation count for each respondent. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3. Randomness Test for the Sex of the Co-Twin

Sample	Male			Female		
	(1) Same Sex	(2) Different Sex	(3) (1)-(2)	(4) Same Sex	(5) Different Sex	(6) (1)-(2)
Age in years	16.558	16.461	0.096	16.325	16.463	-0.137
			(0.156)			(0.172)
Race:						
Black	0.200	0.294	-0.094**	0.327	0.296	0.031
			(0.045)			(0.057)
Hispanic	0.125	0.122	0.003	0.106	0.123	-0.017
			(0.039)			(0.039)
Others	0.675	0.583	0.092*	0.567	0.581	-0.014
			(0.051)			(0.060)
Region:						
West	0.260	0.205	0.055	0.185	0.206	-0.021
			(0.046)			(0.051)
Midwest	0.166	0.193	-0.027	0.283	0.200	0.083
			(0.042)			(0.066)
South	0.444	0.494	-0.050	0.387	0.487	-0.100
			(0.058)			(0.070)
Northeast	0.130	0.108	0.022	0.145	0.106	0.038
			(0.039)			(0.042)
Residential mother's age	42.960	42.488	0.471	41.980	42.509	-0.529
			(0.673)			(0.685)
Residential mother: college degree or more	0.283	0.267	0.017	0.221	0.251	-0.030
			(0.051)			(0.051)
Residential father's age	45.659	45.322	0.337	44.492	45.240	-0.747
			(1.148)			(0.978)
Residential father: college degree or more	0.317	0.272	0.044	0.202	0.263	-0.061
			(0.050)			(0.049)

Variables	Male			Female		
	(1) Same Sex	(2) Different Sex	(3) (1)-(2)	(4) Same Sex	(5) Different Sex	(6) (4)-(5)
Living with both biological parents at age 5	0.792	0.761	0.031	0.721	0.788	-0.067*
			(0.047)			(0.040)
Household income (unit: $1,000)	43.622	46.489	-2.868	45.617	43.640	1.977
			(4.995)			(4.564)
Urbanicity:						
Urban	0.339	0.368	-0.029	0.344	0.367	-0.023
			(0.056)			(0.067)
Suburban	0.575	0.534	0.041	0.574	0.538	0.036
			(0.057)			(0.069)
Rural	0.086	0.098	-0.012	0.082	0.095	-0.013
			(0.030)			(0.036)
School not selected from the sampling frame	0.029	0.033	-0.004	0.063	0.056	0.007
			(0.005)			(0.010)
School size:						
Small (1-400 students)	0.069	0.075	-0.006	0.067	0.071	-0.004
			(0.030)			(0.032)
Medium (401-1000 students)	0.403	0.437	-0.033	0.405	0.450	-0.045
			(0.055)			(0.067)
Large (1001-4000 students)	0.528	0.489	0.039	0.528	0.479	0.049
			(0.057)			(0.069)
Percentage of White students in the school	80.924	78.114	2.811	79.128	78.006	1.122
			(2.193)			(3.797)
Average class size in the school	26.911	26.790	0.121	27.078	26.722	0.356
			(0.534)			(0.701)
N	393	298	691	365	294	659

Notes: This table shows descriptive statistics for my sample of dizygotic twins in the Add Health dataset by sex composition and t-tests of differences in means by sex composition (columns (3) and (6)). Clustered standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4. Co-Twin Sex and Family Structure Among Dizygotic Twins

	Number of Siblings	Living With Both Biological Parents at Age 5	Respondent Parent's Number of Marriages
Panel 1: Male			
SS	-0.048	-0.003	0.104
	(0.131)	(0.043)	(0.077)
Mean	2.130	0.770	1.184
N	283	283	253
Panel 2: Female			
SS	-0.122	-0.022	0.012
	(0.135)	(0.044)	(0.087)
Mean	2.138	0.767	1.154
N	271	271	237

Notes: Each cell represents a separate regression based on Equation 1 among dizygotic twins. The analysis is performed at the household level, using only one observation per household. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Clustered standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Effects of Having a Same-Sex Co-Twin on Substance Use During Adolescence

	Cigarette		Marijuana	
	Ever Tried Smoking	Smoked During the Last 30 Days	Ever Tried Marijuana	Used Marijuana During the Last 30 Days
Panel 1: Male				
SS	0.084*	0.099**	0.160***	0.055
	(0.046)	(0.041)	(0.047)	(0.037)
Mean	0.641	0.328	0.335	0.165
N	644	636	625	634
Panel 2: Female				
SS	0.048	0.028	0.062	0.048*
	(0.052)	(0.045)	(0.044)	(0.027)
Mean	0.503	0.221	0.244	0.123
N	625	623	607	616

	Alcohol		
	Ever Tried Alcohol	Drank During the Last 12 Months	Any Binge Drinking During the Last 12 Months
Panel 1: Male			
SS	0.123**	0.155***	0.076*
	(0.050)	(0.043)	(0.040)
Mean	0.683	0.490	0.330
N	644	641	639
Panel 2: Female			
SS	0.052	-0.016	-0.011
	(0.066)	(0.050)	(0.040)
Mean	0.594	0.401	0.187
N	628	624	623

Notes: Each cell represents a separate regression of the corresponding measure of substance use based on Equation 1 among dizygotic twins. Mean represents the sample average of the corresponding measure for the corresponding sex, which is also reported in Table 2. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6. Robustness to Alternative Specifications

	Male					
	(1)	(2)	(3)	(4)	(5)	(6)
Cigarette						
Ever tried smoking						
SS	0.119**	0.084*	0.084*	0.086*	0.101**	0.082*
	(0.046)	(0.046)	(0.050)	(0.047)	(0.049)	(0.044)
N	687	644	644	644	463	642
Smoked during the last 30 days						
SS	0.122***	0.099**	0.099**	0.102**	0.133***	0.094**
	(0.041)	(0.041)	(0.045)	(0.041)	(0.046)	(0.039)
N	679	636	636	636	460	634
Marijuana						
Ever tried marijuana						
SS	0.149***	0.160***	0.160***	0.163***	0.181***	0.164***
	(0.050)	(0.047)	(0.047)	(0.046)	(0.053)	(0.045)
N	664	625	625	625	456	623
Used marijuana during the last 30 days						
SS	0.060	0.055	0.055	0.054	0.055	0.054
	(0.038)	(0.037)	(0.035)	(0.036)	(0.039)	(0.037)
N	673	634	634	634	457	611
Alcohol						
Ever tried alcohol						
SS	0.138**	0.123**	0.123**	0.117**	0.131*	0.110***
	(0.054)	(0.050)	(0.057)	(0.049)	(0.070)	(0.043)
N	688	644	644	644	462	642
Drank during the last 12 months						
SS	0.189***	0.155***	0.155***	0.154***	0.162***	0.152***
	(0.044)	(0.043)	(0.044)	(0.043)	(0.050)	(0.040)
N	683	641	641	641	459	641
Any binge drinking during the last 12 months						
SS	0.096**	0.076*	0.076*	0.077*	0.087**	0.075*
	(0.040)	(0.040)	(0.045)	(0.040)	(0.042)	(0.038)
N	681	639	639	639	458	639
Controls	**No**	Yes	Yes	Yes	Yes	Yes
Cluster	School	School	**Family**	School	School	School
SS Sample	All / H	All / H	All / H	**1/2 Wt**	**1 Obs / H**	All / H
Specification	Linear	Linear	Linear	Linear	Linear	**Probit**

	Female					
	(1)	(2)	(3)	(4)	(5)	(6)
Cigarette						
Ever tried smoking						
SS	0.042	0.048	0.048	0.053	0.047	0.039
	(0.055)	(0.052)	(0.056)	(0.053)	(0.059)	(0.053)
N	656	625	625	625	454	610
Smoked during the last 30 days						
SS	0.021	0.028	0.028	0.034	0.019	0.034
	(0.047)	(0.045)	(0.042)	(0.044)	(0.046)	(0.044)
N	654	623	623	623	453	618
Marijuana						
Ever tried marijuana						
SS	0.060	0.062	0.062	0.063	0.077	0.059
	(0.044)	(0.044)	(0.047)	(0.044)	(0.052)	(0.042)
N	636	607	607	607	443	602
Used marijuana during the last 30 days						
SS	0.039	0.048*	0.048	0.050*	0.032	0.042
	(0.027)	(0.027)	(0.031)	(0.028)	(0.030)	(0.027)
N	644	616	616	616	448	611
Alcohol						
Ever tried alcohol						
SS	0.022	0.052	0.052	0.063	0.041	0.014
	(0.070)	(0.066)	(0.072)	(0.068)	(0.083)	(0.049)
N	659	628	628	628	456	623
Drank during the last 12 months						
SS	-0.046	-0.016	-0.016	-0.008	-0.034	-0.016
	(0.052)	(0.050)	(0.047)	(0.050)	(0.053)	(0.049)
N	655	624	624	624	453	619
Any binge drinking during the last 12 months						
SS	-0.020	-0.011	-0.011	-0.008	-0.012	-0.007
	(0.043)	(0.040)	(0.036)	(0.040)	(0.043)	(0.040)
N	654	623	623	623	453	618
Controls	**No**	Yes	Yes	Yes	Yes	Yes
Cluster	School	School	**Family**	School	School	School
SS Sample	All / H	All / H	All / H	**1/2 Wt**	**1 Obs / H**	All / H
Specification	Linear	Linear	Linear	Linear	Linear	**Probit**

Notes: Each cell represents a separate regression of the corresponding measure of substance use among dizygotic twins. Controls include region dummies, race dummies, age and age squared variables, residential mothers' age, residential mother's college degree (dummy), missing maternal educational attainment (dummy), residential father's college degree (dummy), and missing paternal educational attainment (dummy). Estimates are weighted by the reciprocal of the observation count for each respondent, except for column (4), where the model additionally weights those with a same-sex twin sibling by multiplying the weight by 1/2. Clustered standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7. Exploring the Peer Network: Number of Substance-Using Friends by Sex

	Among the Three Best Friends, Number of Those Who			Having at Least One	
	Smoke	Use Marijuana	Drink Alcohol	Male Friend	Female Friend
I am Female	-0.268***	-0.179***	-0.229***	-0.119***	0.127***
	(0.071)	(0.068)	(0.078)	(0.017)	(0.018)
Mean	0.838	0.706	1.108	0.918	0.916
N	1251	1252	1250	1273	1273

Notes: Each cell represents a separate regression based on Equation 1 among all dizygotic twins, where SS_i is substituted with $I_a m_F emale_i$. This is a dummy variable that indicates whether the respondent is female. Mean represents the sample average of the corresponding measure among all respondents. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8. Exploring the Channel of the Peer Network: Common Friends and Time Spent Together by Co-Twin's Sex

	Whether Having at Least One	Whether Spending a Lot of Time with	
	Common Friend Outside of the Family	Co-Twin	Co-Twin and Common Friends
Panel 1: Male			
SS	0.219**	0.220***	0.257***
	(0.092)	(0.043)	(0.041)
Mean	0.515	0.465	0.323
N	221	601	600
Panel 2: Female			
SS	0.455***	0.241***	0.268***
	(0.087)	(0.042)	(0.048)
Mean	0.556	0.585	0.376
N	203	583	583

Notes: Each cell represents a separate regression based on Equation 1 among dizygotic twins. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9. Exploring the Channel of the Peer Network: Substance-Using Friends by Co-twin's Sex

	Among the Three Best Friends, Number of Those Who			Having at Least One	
	Smoke	Use Marijuana	Drink Alcohol	Male Friend	Female Friend
Panel 1: Male					
SS	0.169*	0.144	0.179*	0.012	0.025
	(0.093)	(0.094)	(0.108)	(0.012)	(0.032)
Mean	0.959	0.780	1.224	0.972	0.857
N	633	634	635	646	646
Panel 2: Female					
SS	0.101	0.046	0.102	-0.060**	-0.010
	(0.105)	(0.093)	(0.105)	(0.027)	(0.012)
Mean	0.706	0.625	0.979	0.860	0.980
N	618	618	615	627	627

Notes: Each cell represents a separate regression based on Equation 1 among dizygotic twins. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10. Exploring the Channel of the Parental Investment

	Time		Allowance	Preventive Care	
	Activities With Father	Activities With Mother	Weekly Amount	Physical Check-up	Dental Check-up
Panel 1: Male					
SS	0.227	-0.111	0.831	0.017	-0.013
	(0.211)	(0.180)	(0.862)	(0.051)	(0.052)
Mean	2.804	3.457	5.755	0.683	0.617
N	468	637	644	642	646
Panel 2: Female					
SS	0.220	0.161	1.552	-0.019	-0.019
	(0.198)	(0.192)	(1.243)	(0.053)	(0.052)
Mean	2.426	4.010	6.005	0.695	0.652
N	417	618	627	620	619

Notes: Each cell represents a separate regression based on Equation 1 among dizygotic twins. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7 Appendices

7.1 Self-report Measurement Errors

Using data from the 1999-2000 National Health and Nutrition Examination Survey, I also explore the discrepancy between self-reported current smoking status and an objective measure of smoking obtained from saliva cotinine levels (defined as a smoker if cotinine level $\geq$ 15ng/ml).

In the sample of 6,758 youths aged 12–19, 5.3% are self-reported (SR) smoker but cotinine-validated (CV) non-smoker, 27.5% are both SR and CV smoker, 65.7% are both SR and CV non-smoker, and 1.5% are SR non-smoker but CV smoker. This analysis suggests that in the most extreme case where (1) there is no real difference in the likelihood of current smoking among males based on the sex of their co-twin, (2) all males with a same-sex co-twin over-report their smoking status (i.e., 34.3% (= 5.3% + 27.5% + 1.5%) report smoking and the remaining 65.7% report not smoking), and (3) all males with a different-sex co-twin under-report smoking (i.e., 72.5% (= 5.3% + 65.7% + 1.5%) report not smoking and the remaining 27.5% report smoking), we can still observe a 6.8 percentage point difference in the likelihood of current smoking. When compared to the 9.9 percentage point difference estimated in the twin sample (Table 5), it indicates that at most, misreporting behaviors could account for up to 68.7% of the observed estimate.

7.2 Appendix Tables

Table 1.A1. Correlation in Substance Use Among Same-Sex Twins

	Males		Females	
	MZ	DZ	MZ	DZ
Cigarette				
Ever tried smoking				
Correlation	0.518***	0.324***	0.600***	0.452***
	(0.077)	(0.105)	(0.080)	(0.084)
N	211	193	217	181
Smoked during the last 30 days				
Correlation	0.585***	0.290***	0.525***	0.372***
	(0.062)	(0.080)	(0.090)	(0.090)
N	209	185	213	180
Marijuana				
Ever tried marijuana during your life				
Correlation	0.523***	0.275***	0.591***	0.380***
	(0.090)	(0.080)	(0.080)	(0.092)
N	202	177	209	169
Used marijuana during the last 30 days				
Correlation	0.614***	0.288***	0.287**	0.253***
	(0.095)	(0.090)	(0.133)	(0.094)
N	208	182	213	174
Alcohol				
Ever tried drinking alcohol				
Correlation	0.373***	0.065	0.480***	0.729***
	(0.084)	(0.150)	(0.066)	(0.143)
N	211	192	218	182
Drank during the last 12 months				
Correlation	0.422***	0.215***	0.532***	0.278***
	(0.079)	(0.074)	(0.065)	(0.080)
N	208	191	218	181
Any binge drinking during the last 12 months				
Correlation	0.538***	0.366***	0.521***	0.219**
	(0.087)	(0.070)	(0.089)	(0.100)
N	208	190	216	180

Notes: Each model is estimated among twins with a same-sex twin sibling. MZ and DZ represent monozygotic and dizygotic twins. Coefficients are unadjusted correlation coefficients for the corresponding measure of substance use. Estimates are weighted by the reciprocal of the observation count for each respondent. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A2. Comparisons Between Monozygotic Twins and Dizygotic Twins with a Same-Sex Co-Twin

	Cigarette		Marijuana	
	Ever Tried Smoking	Smoked During the Last 30 Days	Ever Tried Marijuana	Used Marijuana During the Last 30 Days
Panel 1: Male				
MZ	-0.061	-0.090**	-0.108**	-0.044
	(0.050)	(0.037)	(0.050)	(0.036)
Mean	0.638	0.318	0.345	0.172
N	792	782	770	782
Panel 2: Female				
MZ	0.102*	0.054	-0.005	-0.012
	(0.058)	(0.044)	(0.044)	(0.032)
Mean	0.554	0.249	0.250	0.127
N	801	797	778	789

	Alcohol		
	Ever Tried Alcohol	Drank During the Last 12 Months	Any Binge Drinking During the Last 12 Months
Panel 1: Male			
MZ	-0.171***	-0.167***	-0.104***
	(0.049)	(0.040)	(0.038)
Mean	0.656	0.471	0.304
N	791	790	788
Panel 2: Female			
MZ	-0.019	0.037	0.022
	(0.063)	(0.048)	(0.041)
Mean	0.586	0.411	0.206
N	804	802	799

Notes: Each model is estimated among all twins with a same-sex twin sibling. It follows Equation 1, substituting SS with MZ, a dummy variable for respondent i being a monozygotic twin. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A3. Number of Non-Twin Full Sibling Respondents

	Male	Female	**Total**
Having a different-sex sibling	127	127	254
Having a same-sex sibling	206	192	398
Total	333	398	652

Notes: The sample of non-twin full siblings consists of the respondents identified by Add Health to have a non-twin full sibling who was also in grades 7-12 during the 1994-95 school year and agreed to be a respondent of the Add Health survey. To be consistent with the twin sample, I require that each individual should report his/her sibling in the household roster at Wave I, that the sibling is his/her closest sibling in terms of age, and that the age difference with the sibling is two years.

Table 1.A4. Likelihood of Reporting Completely Honestly Among Dizygotic Twins

	Males	Females
SS	0.017	-0.043
	(0.054)	(0.047)
Mean	0.457	0.589
N	646	628

Notes: Each model is estimated among dizygotic twins using Equation 1. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A5. Randomness Test for the Sex of the Non-Twin Sibling

Sample	Male			Female		
	(1) Same Sex	(2) Different Sex	(3) (1)-(2)	(4) Same Sex	(5) Different Sex	(6) (1)-(2)
Age in years	16.709	16.601	0.108	16.764	16.494	0.271*
			(0.179)			(0.146)
Race:						
Black	0.136	0.110	0.026	0.083	0.110	-0.027
			(0.044)			(0.034)
Hispanic	0.150	0.102	0.048	0.099	0.110	-0.011
			(0.035)			(0.046)
Others	0.714	0.787	-0.074	0.818	0.780	0.038
			(0.057)			(0.057)
Region:						
West	0.236	0.228	0.008	0.237	0.228	0.008
			(0.039)			(0.051)
Midwest	0.330	0.425	-0.095*	0.382	0.425	-0.043
			(0.056)			(0.053)
South	0.305	0.228	0.077	0.258	0.228	0.030
			(0.054)			(0.050)
Northeast	0.128	0.118	0.010	0.124	0.118	0.006
			(0.033)			(0.033)
Residential mother's age	40.708	40.462	0.246	40.378	40.603	-0.226
			(0.586)			(0.473)
Residential mother: college degree or more	0.306	0.252	0.054	0.266	0.276	-0.010
			(0.048)			(0.055)
Residential father's age	43.546	43.126	0.420	43.027	43.010	0.018
			(0.776)			(0.675)
Residential father: college degree or more	0.252	0.205	0.048	0.276	0.197	0.079*
			(0.039)			(0.047)

Sample	Male			Female		
	(1) Same Sex	(2) Different Sex	(3) (1)-(2)	(4) Same Sex	(5) Different Sex	(6) (1)-(2)
Living with both biological parents at age 5	0.825	0.827	-0.002	0.828	0.874	-0.046
			(0.050)			(0.041)
Household income (unit: $1,000)	44.031	54.556	-10.524	44.214	55.904	-11.690
			(9.676)			(9.802)
Urbanicity:						
Urban	0.212	0.134	0.078	0.274	0.134	0.140**
			(0.047)			(0.060)
Suburban	0.488	0.465	0.023	0.371	0.465	-0.094
			(0.057)			(0.056)
Rural	0.300	0.402	-0.101*	0.355	0.402	-0.047
			(0.052)			(0.049)
School not selected from the sampling frame	0.015	0.000	0.015	0.031	0.000	0.031
			(0.015)			(0.031)
School size:						
Small (1-400 students)	0.212	0.220	-0.009	0.177	0.236	-0.059
			(0.038)			(0.051)
Medium (401-1000 students)	0.256	0.362	-0.106*	0.242	0.323	-0.081
			(0.059)			(0.052)
Large (1001-4000 students)	0.532	0.417	0.115*	0.581	0.441	0.140**
			(0.064)			(0.060)
Percentage of White students in the school	83.358	86.152	-2.794	88.413	85.960	2.453
			(2.163)			(2.054)
Average class size in the school	25.627	25.376	0.251	25.571	25.488	0.083
			(0.476)			(0.590)
N	338	204	542	324	213	537

Notes: This table shows descriptive statistics for my sample of non-twin full siblings by sex composition and t-tests of differences in means by sex composition (columns (3) and (6)). Clustered standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A6. Sibling Gender and Family Structure Among Non-Twin Full Siblings

	Number of Siblings	Living With Both Biological Parents at Age 5	Respondent Parent's Number of Marriages
Panel 1: Male			
SS	0.146	0.006	0.029
	(0.149)	(0.052)	(0.107)
Mean	1.717	0.826	1.288
N	215	215	195
Panel 2: Female			
SS	0.330**	-0.043	-0.017
	(0.163)	(0.038)	(0.097)
Mean	1.776	0.857	1.289
N	211	211	184

Notes: Each cell represents a separate regression based on Equation 1 in the sample of non-twin full siblings. The analysis is performed at the household level, using only one observation per household. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A7. Family Structure Among Mothers of Twin in the 1990 Census

	Fertility		Marital Status			
	Number of Own Children	Number of Deliveries Ever	Living Without Any Spouse	Living Without Child's Bio Dad	Never Married	Current Divorce or Seperation
Panel 1: Male						
SS	-0.041	-0.040	-0.000	0.004	0.002	-0.001
	(0.033)	(0.034)	(0.011)	(0.012)	(0.006)	(0.011)
Mean	3.232	3.232	0.225	0.278	0.058	0.159
N	4,321	4,321	4,321	4,321	4,321	4,072
Panel 2: Female						
SS	0.051	0.053	-0.001	0.002	0.002	-0.005
	(0.034)	(0.034)	(0.011)	(0.012)	(0.006)	(0.011)
Mean	3.280	3.282	0.226	0.278	0.058	0.158
N	4,368	4,368	4,368	4,368	4,368	4,114

Notes: Each cell represents a separate regression based on Equation 1 among women in the 1990 Census data, aged 18 to 40, with at least one child (the oldest being younger than 18 years old), living with all the children they ever reported having delivered, and having two children with the same birth year, birthplace, and age in years (whose birth year are between 1976 and 1983 as my Add Health sample). Standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A8. Family Structure Among Mothers of Non-Twins in the 1990 Census

	Fertility		Marital Status			
	Number of Own Children	Number of Deliveries Ever	Living Without Any Spouse	Living Without Child's Bio Dad	Never Married	Current Divorce or Seperation
Panel 1: Male Child						
SS	0.087***	0.088***	0.001	-0.001	0.003**	-0.000
	(0.007)	(0.008)	(0.003)	(0.003)	(0.001)	(0.002)
Mean	3.232	3.232	0.225	0.278	0.058	0.159
N	70,724	70,724	70,724	70,724	70,724	68,342
Panel 2: Female Child						
SS	0.121***	0.121***	0.006**	0.009***	0.001	0.005*
	(0.008)	(0.008)	(0.003)	(0.003)	(0.001)	(0.003)
Mean	3.280	3.282	0.226	0.278	0.058	0.158
N	68,234	68,234	68,234	68,234	68,234	65,950

Notes: Each cell represents a separate regression based on Equation 1 among women in the 1990 Census data, aged 18 to 40, with at least one child (the oldest being younger than 18 years old), living with all the children they ever reported having delivered, and having two children with age space of 2 years (whose birth year are between 1976 and 1983 as my Add Health sample). Standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A9. Effects of Younger Sibling's Gender on Older Sibling's Substance Use

	Cigarette		Marijuana	
	Ever Tried Smoking	Smoked During the Last 30 Days	Ever Tried Marijuana	Used Marijuana During the Last 30 Days
Panel 1: Male				
SS	0.011	-0.046	0.070	0.007
	(0.110)	(0.066)	(0.059)	(0.049)
Mean	0.598	0.301	0.321	0.177
N	258	258	250	255
Panel 2: Female				
SS	0.030	-0.054	-0.096	-0.018
	(0.072)	(0.062)	(0.068)	(0.046)
Mean	0.561	0.272	0.238	0.131
N	232	233	224	232

	Alcohol		
	Ever Tried Alcohol	Drank During the Last 12 Months	Any Binge Drinking During the Last 12 Months
Panel 1: Male			
SS	-0.001	0.060	-0.040
	(0.075)	(0.069)	(0.090)
Mean	0.652	0.523	0.369
N	260	259	257
Panel 2: Female			
SS	0.068	-0.045	-0.012
	(0.109)	(0.086)	(0.057)
Mean	0.648	0.514	0.252
N	233	233	233

Notes: Each cell represents a separate regression based on Equation 1 among non-twin full siblings. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A10. Effects of Older Sibling's Gender on Younger Sibling's Substance Use

	Cigarette		Marijuana	
	Ever Tried Smoking	Smoked During the Last 30 Days	Ever Tried Marijuana	Used Marijuana During the Last 30 Days
Panel 1: Male				
SS	-0.089	0.015	0.005	0.010
	(0.093)	(0.065)	(0.067)	(0.045)
Mean	0.641	0.301	0.255	0.130
N	253	252	244	250
Panel 2: Female				
SS	-0.064	-0.015	0.013	0.029
	(0.072)	(0.065)	(0.054)	(0.044)
Mean	0.647	0.326	0.259	0.132
N	269	268	268	268

	Alcohol		
	Ever Tried Alcohol	Drank During the Last 12 Months	Any Binge Drinking During the Last 12 Months
Panel 1: Male			
SS	-0.006	-0.044	-0.047
	(0.088)	(0.064)	(0.050)
Mean	0.586	0.414	0.292
N	252	251	251
Panel 2: Female			
SS	0.088	0.049	-0.067
	(0.084)	(0.072)	(0.071)
Mean	0.623	0.447	0.232
N	270	270	268

Notes: Each cell represents a separate regression based on Equation 1 among non-twin full siblings. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A11. Comparing the Effects of Having a Same-Sex Co-Twin Between Males and Females Among Dizygotic Twins

	Cigarette		Marijuana	
	Ever Tried Smoking	Smoked During the Last 30 Days	Ever Tried Marijuana	Used Marijuana During the Last 30 Days
SS	0.098**	0.102**	0.157***	0.053
	(0.048)	(0.041)	(0.047)	(0.036)
Female	-0.109**	-0.060*	-0.055	-0.033
	(0.049)	(0.034)	(0.038)	(0.026)
SS × Female	-0.052	-0.071	-0.098*	-0.013
	(0.067)	(0.062)	(0.059)	(0.039)
N	1269	1259	1232	1250

	Alcohol		
	Ever Tried Alcohol	Drank During the Last 12 Months	Any Binge Drinking During the Last 12 Months
SS	0.128**	0.166***	0.085**
	(0.053)	(0.044)	(0.038)
Female	-0.026	0.040	-0.082**
	(0.056)	(0.036)	(0.041)
SS × Female	-0.072	-0.178***	-0.093
	(0.080)	(0.061)	(0.057)
N	1272	1265	1262

Notes: Each model is estimated based on Equation 1 among all dizygotic twins. Standard errors clustered at the school level are indicated within parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A12. Effects of Having a Same-Sex Co-Twin Using Summary Indices

	All	By Substance		
	All	Cigarette	Marijuana	Alcohol
Panel 1: Male				
SS	0.273***	0.230**	0.285***	0.234***
	(0.076)	(0.091)	(0.105)	(0.085)
N	646	644	635	646
Panel 2: Female				
SS	0.157	0.090	0.149*	0.045
	(0.117)	(0.101)	(0.086)	(0.110)
N	628	625	616	628

Notes: Each cell represents a separate regression based on Equation 1 among dizygotic twins. Summary indices are derived by averaging the standardized z-scores for each outcome and then re-standardizing the average. These summary indices encompass the all substances summary index (comprising all outcome variables in Table Table 5), the cigarette summary index, the marijuana summary index, and the alcohol summary index. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A13. The Impact of a Younger Sibling's Sex on Common Friends and Time Spent Together Among Older Siblings

	Having at Least One	Spending a Lot of Time with	
	Common Friend	Co-Twin	Co-Twin and Common Friends
Panel 1: Male			
SS	0.091	0.123	0.112**
	(0.133)	(0.076)	(0.054)
Mean	0.290	0.336	0.199
N	98	229	229
Panel 2: Female			
SS	0.188*	0.161**	0.092
	(0.106)	(0.080)	(0.060)
Mean	0.281	0.494	0.179
N	95	205	205

Notes: Each cell represents a separate regression based on Equation 1 among non-twin full siblings. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A14. The Impact of an Older Sibling's Sex on Common Friends and Time Spent Together Among Younger Siblings

	Having at Least One	Spending a Lot of Time with	
	Common Friend	Co-Twin	Co-Twin and Common Friends
Panel 1: Male			
SS	0.171	0.101	0.016
	(0.168)	(0.088)	(0.063)
Mean	0.250	0.405	0.227
N	85	204	204
Panel 2: Female			
SS	0.189**	0.229***	-0.007
	(0.088)	(0.073)	(0.071)
Mean	0.313	0.383	0.212
N	118	239	240

Notes: Each cell represents a separate regression based on Equation 1 among non-twin full siblings. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A15. The Impact of a Younger Sibling's Sex on the Older Sibling's Friend Network

	Among the Three Best Friends, Number of Those Who			Having at Least One	
	Smoke	Use Marijuana	Drink Alcohol	Male Friend	Female Friend
Panel 1: Male					
SS	-0.108	0.029	-0.223	0.013	0.061
	(0.163)	(0.116)	(0.148)	(0.020)	(0.039)
Mean	0.889	0.714	1.292	0.979	0.893
N	257	257	257	260	260
Panel 2: Female					
SS	-0.337***	-0.170	-0.372***	-0.114***	0.006
	(0.112)	(0.152)	(0.128)	(0.037)	(0.022)
Mean	0.865	0.594	1.141	0.924	0.976
N	232	231	233	234	235

Notes: Each cell represents a separate regression based on Equation 1 among non-twin full siblings. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A16. The Impact of an Older Sibling's Sex on the Younger Sibling's Friend Network

	Among the Three Best Friends, Number of Those Who			Having at Least One	
	Smoke	Use Marijuana	Drink Alcohol	Male Friend	Female Friend
Panel 1: Male					
SS	-0.009	-0.022	-0.117	-0.013	0.082
	(0.154)	(0.150)	(0.140)	(0.010)	(0.063)
Mean	0.800	0.520	1.014	0.983	0.842
N	251	250	249	253	253
Panel 2: Female					
SS	0.011	0.128	-0.075	-0.016	-0.001
	(0.126)	(0.165)	(0.157)	(0.050)	(0.019)
Mean	0.887	0.569	1.081	0.882	0.971
N	269	269	268	270	270

Notes: Each cell represents a separate regression based on Equation 1 among non-twin full siblings. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A17. Effects of Younger Sibling's Gender on Parental Investment in Older Siblings

	Time		Allowance	Preventive Care	
	Activities With Father	Activities With Mother	Weekly Amount	Physical Check-up	Dental Check-up
Panel 1: Male					
SS	-0.463	-0.075	-1.668	-0.027	0.018
	(0.331)	(0.271)	(1.343)	(0.076)	(0.076)
Mean	2.842	3.544	5.566	0.701	0.684
N	218	256	259	260	260
Panel 2: Female					
SS	0.287	0.304	-0.209	-0.083	-0.092
	(0.340)	(0.320)	(1.232)	(0.074)	(0.079)
Mean	3.162	4.627	5.292	0.737	0.717
N	171	225	232	235	235

Notes: Each cell represents a separate regression based on Equation 1 among non-twin full siblings. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A18. Effects of Older Sibling's Gender on Parental Investment in Younger Siblings

	Time		Allowance	Preventive Care	
	Activities With Father	Activities With Mother	Weekly Amount	Physical Check-up	Dental Check-up
Panel 1: Male					
SS	0.353	0.195	1.144	0.011	-0.013
	(0.353)	(0.285)	(1.206)	(0.081)	(0.086)
Mean	3.038	3.429	6.257	0.742	0.656
N	198	248	251	253	249
Panel 2: Female					
SS	-0.093	0.022	1.031	-0.148*	-0.090
	(0.312)	(0.302)	(1.144)	(0.077)	(0.069)
Mean	2.965	4.321	6.505	0.689	0.719
N	225	267	267	267	271

Notes: Each cell represents a separate regression based on Equation 1 among non-twin full siblings. Mean represents the sample average of the corresponding measure for the corresponding sex. Standard errors clustered at the school level are indicated within parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$