Title
Peer Influence, Genetic Propensity, and Binge Drinking: A Natural Experiment and a Replication.

Journal
AJS; American journal of sociology, 121(3)

ISSN
0002-9602

Authors
Guo, Guang
Li, Yi
Wang, Hongyu
et al.

Publication Date
2015-11-01

DOI
10.1086/683224

Peer reviewed

Peer Influence, Genetic Propensity, and Binge Drinking: A Natural Experiment and a Replication[1]

Guang Guo and Yi Li
University of North Carolina at Chapel Hill

Hongyu Wang and Tianji Cai
University of Macau

Greg J. Duncan
University of California, Irvine

The authors draw data from the College Roommate Study (ROOM) and the National Longitudinal Study of Adolescent Health to investigate gene-environment interaction effects on youth binge drinking. In ROOM, the environmental influence was measured by the precollege drinking behavior of randomly assigned roommates. Random assignment safeguards against friend selection and removes the threat of gene-environment correlation that makes gene-environment interaction effects difficult to interpret. On average, being randomly assigned a drinking peer as opposed to a nondrinking peer increased college binge drinking by 0.5–1.0 episodes per month, or 20%–40% the average amount of binge drinking. However, this peer influence was found only among youths with a medium level of genetic propensity for alcohol use; those with either a low or high genetic propensity were not influenced by peer drinking. A replication of the findings is provided in data drawn from Add Health. The study shows that gene-environment interaction analysis can uncover social-contextual effects likely to be missed by traditional sociological approaches.

[1] This research uses data from the College Roommate Study, funded by a grant from the William T. Grant Foundation, and also data from Add Health, funded by the National Institute of Child Health and Human Development (P01-HD3921), with cooperative

0002-9602/2015/12103-00XX$10.00

INTRODUCTION

We investigate how peer influence on binge drinking (defined as five or more drinks in a row for males and four or more drinks in a row for females) is moderated by genetic propensity for alcohol use among youth in the United States. The study draws on social science and genotype data from the College Roommate Study (ROOM) and the National Longitudinal Study of Adolescent Health (Add Health; Harris et al. 2003). q1

Our study breaks new ground by using a natural experimental design to investigate gene-by-peer influence interactions on youth risky behavior. Experimental variation is generated from roommates who were randomly assigned to one another on a college campus. Peers are often believed to influence risky behavior among youth, but peer influence remains difficult to investigate because observational data cannot separate peer influence from friend selection (see the review by Kandel 1978; Manski 1993; Moffitt 2001; Mouw 2006). Randomized roommates have been used previously to estimate peer influences but have not been used in a gene-environment ($G \times E$) interaction analysis. Our experimental study design avoids confounds of peer selection and removes the threat of gene-environment correlation (rGE) with respect to peer influence: rGE occurs when an apparent environmental influence can be partially attributed to genetic factors (Jaffee and Price 2007; Conley and Rauscher 2013; Fletcher and Conley 2013; Wagner et al. 2013).

In this study, we used a modified theory of social learning (Bandura 1971) to guide empirical analysis. The modification allows a peer effect on binge drinking to condition on genetic propensity for alcohol use. Genetic propensity could be conceptualized as an intrinsic component of self-control (Gottfredson and Hirschi 1990). We measure self-control by genetic propensity based on molecular genetic data to address complications arising from measuring self-control via self-behavior. For this $G \times E$ interaction analysis, we propose a swing theory that amounts to an extended classic diathesis-stress model.

We develop a methodological strategy for $G \times E$ interaction analysis when a relatively large number of genetic variables are involved. Our focus on peer influence as a single environmental exposure alleviates the burden

funding from 17 other agencies (http://www.cpc.unc.edu/addhealth/contract.html). We gratefully acknowledge support to the first author from the National Institutes of Health (R03 HD042490-02, R03 HD053385-01, RC1 DA029425-01) and the National Science Foundation (SES-0210389). Special acknowledgment is owed to Rick Bradley of the Housing Department, the Odum Institute, and Kirk Wilhelmsen of the Genetics Department at the University of North Carolina, Chapel Hill. Direct correspondence to Guang Guo, Department of Sociology, CB 3210, University of North Carolina, Chapel Hill, North Carolina 27599. E-mail: guang_guo@unc.edu

of multiple testing in a G × E interaction analysis. We replicate the G × E interaction findings from ROOM using nonexperimental data on friends from Add Health (Harris et al. 2003).

The G × E interaction studies can advance sociological reasoning in several important ways. First, G × E interaction analysis can reveal social-contextual effects that would remain hidden in a traditional sociological study. Many social-contextual effects such as peer effects might be moderated by genetic effects if, for example, social-contextual effects are much weaker or even completely absent in some group of individuals with a particular range of genetic propensity and much stronger in another group with different ranges of genetic propensities. Traditional sociological studies focusing on social-contextual effects estimate average effects across various levels of genetic propensity. Consistent with this theoretical argument, empirical evidence from the current study shows that youth with a middle level of genetic propensity for alcohol use tend to be more affected by peer drinking than youth with a low or high level of genetic propensity.

Second, many sociological outcomes such as binge drinking, substance abuse, and educational and occupational achievement are likely to be influenced by numerous interacting genetic and social-contextual factors, each asserting a modest effect. Studies of G × E interaction may help to better understand the complicated interplay among these social-contextual and genetic factors. Third, G × E interaction analysis can help predict which groups of individuals are most sensitive to a particular social-contextual exposure. Finally, G × E interaction findings can be used to design intervention strategies to remove or adjust social-contextual exposures of those groups with a genetic vulnerability to these exposures. Genetic variations are fixed; social-contextual exposures might be alterable. For example, in the current case, an effective intervention strategy might be developed to alter peer networks for individuals with a particularly vulnerable genetic propensity for binge drinking.

BACKGROUND

Excessive Alcohol Consumption and Peer Influence among Youth

Excessive drinking, including binge drinking among youth, can have severe negative consequences for both those who drink and those around them. It may lead, for example, to physical aggression and violence, sexual aggression, sexual victimization, academic difficulties, problems with friends, health issues, and fatal traffic accidents (Wechsler et al. 1994; Panel on Contexts and Consequences 2002; Roudsari, Leahy, and Walters 2009; Schilling et al. 2009; Singleton and Wolfson 2009; Palmer et al. 2010). Research shows that alcohol use is especially damaging for brain function-

ing during adolescence (Bellis et al. 2000; Tapert, Caldwell, and Burke 2005).

Excessive drinking is also associated with violence in marital relationships and other social circumstances (Kantor and Straus 1990; Leonard, Collins, and Quigley 2003; Leonard, Quigley, and Collins 2003). Alcohol use is associated with more than 50% of sexual assaults among youth. Alcohol use is linked with serious personal injuries and accidents. Approximately one-third of 18–24-year-olds admitted to emergency rooms for severe injuries are under the influence of alcohol. Heavy drinking is related to homicides, suicides, and drownings. Alcohol is involved in about 50% of all fatal traffic accidents. Consequences of heavy drinking also have "secondhand effects" similar to the effects of secondhand smoking, including noise, property damage, vomit, and litter.

Alcohol use is widespread on U.S. college campuses. It is a major part of the college culture and present at many social occasions and peer interaction functions (Thombs 1999). College students tend to consider alcohol use an acceptable behavior (Johnson 1989; Eastman 2002). Many view college years as a period during which they can use alcohol excessively before taking on the responsibilities of adulthood.

Borsari and Carey (2001) outline a number of mechanisms through which peers influence college drinking. Direct peer influence can take the form of friendly gestures (e.g., buying a drink or a round of drinks) or overt pressures to drink (e.g., pressuring peers to play drinking games). Peers may also influence college drinking indirectly by acting as role models. Through their own drinking behavior, peers indicate what behaviors are accepted and appropriate.

Evidence for a Genetic Propensity for Alcohol Use

Twin and adoption studies have demonstrated an important genetic basis for alcohol-related disorders or alcoholism, with large twin studies showing that more than one-half of the variation in alcoholism is due to genetic factors (Goldman and Bergen 1998). Advances in genomic studies in recent years have improved our understanding of the molecular genetic origin of addictive behavior, including alcoholism. A large number of animal studies have been performed with ethanol using methods of gene deletion, gene overexpression, and gene knock-in (Crabbe et al. 2006). These studies implicate genes related to serotonin, gamma-aminobutyric acid, opioids, dopamine, and protein kinase C. The studies suggest diverse genetic pathways leading to addiction, which implies that genetic studies of addiction in humans must deal with the possibility of genetic heterogeneity (different cases of alcoholism may be related to different genetic loci) and polygenicity (multiple genes may contribute to an individual's alcoholism).

Theoretical Framework

Our modified social learning theory is illustrated in figures 1 and 2. Figure 1*a* shows the standard model of social learning in which peer drinking has a direct effect on binge drinking. Figure 1*b* describes a modified model of social learning, or a G × E interaction model. This model incorporates the idea that individuals have different genetic vulnerabilities or propensities toward alcoholism. As a result, the effect of social learning depends on the level of the propensity. Figure 2 further describes this G × E interaction model, or swing model, which, as we explain below, amounts to an extended diathesis-stress model. From two different angles, figures 2*a* and 2*b* illustrate our main theoretical argument that peer influence on binge drinking tends to be the greatest on individuals having medium genetic propensity for binge drinking.

Social differential theory and social learning theory.—Social differential theory and social learning theory are often used to guide the interpretation of peer influences. Sutherland's (1947) differential association theory maintains that delinquent behavior is acquired through close association with peers in which the attitudes and norms of delinquent peers are learned.

Social learning theory expands Sutherland's theory to account for an individual's critical analysis of external influences. In contrast to the differential association theory, the social learning perspective believes that humans are capable of developing hypotheses about which behavior may be beneficial and acceptable. Bandura maintains (1971, pp. 2–3, 13–15, 26–27) that in social learning, human behavior is a result of an interaction between an individual's internal forces and external conditions. Social learning theory distinguishes two concepts: imitation and operant conditioning (Akers 1985, 2001). When first initiated, delinquent behavior is often imitated or learned by observing similar behavior in others. Operant conditioning is also a form of learning in which an individual's behavior is shaped

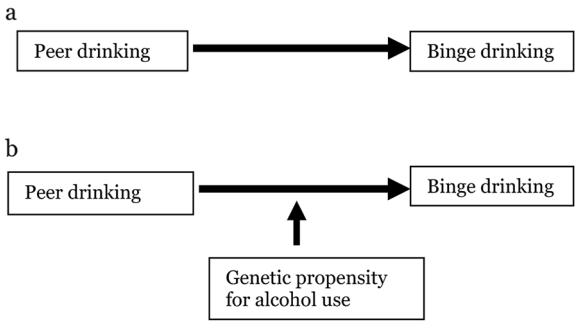

FIG. 1.—*a*, Standard theory of social learning; *b*, modified theory of social learning, or gene-environment interaction.

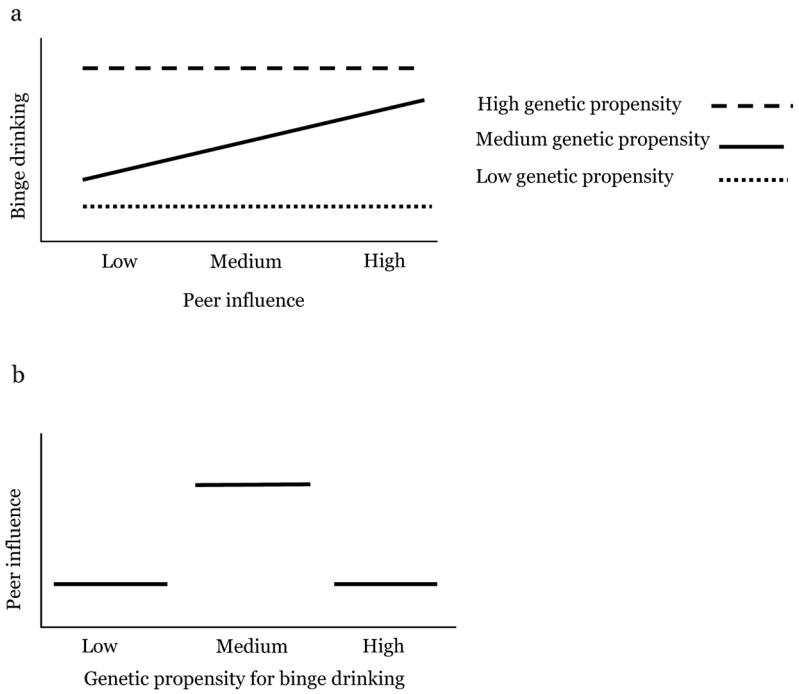

FIG. 2.—*a*, G × E interaction: swing model, or an extended diathesis-stress model, with binge drinking plotted against peer influence. *b*, same model, but with peer influence plotted against genetic propensity.

by its consequences. A positive consequence of a delinquent act encourages delinquency, while a negative consequence discourages the behavior.

For our analysis, the social learning perspective has two specific theoretical implications. First, binge drinking on the part of peers helps create the sense that binge drinking is acceptable or normal on a college campus, which suggests that binge drinking among youth could be learned via intimate social interactions with drinking peers. Second, when internally assessing the positive versus negative consequences of binge drinking, an individual's intrinsic propensity for alcoholism could condition the assessment. For example, the consequences of binge drinking could be underestimated or overestimated because of one's genetic propensity for alcohol use. The underestimation or overestimation will, in turn, affect the strength of peer influence.

Genetic propensity and self-control.—In this study, we allow social learning to depend on genetic propensity for alcohol use. Generally speaking, indi-

viduals possessing a higher level of genetic propensity tend to overestimate the positive and underestimate the negative consequences of binge drinking.

The idea of conditioning on genetic propensity is closely related to the theory of self-control (Gottfredson and Hirschi 1990). The theory describes delinquency-prone individuals as lacking self-control over their own desires and being incapable of resisting the immediate gratification provided by crime or analogous behaviors. Gottfredson and Hirschi propose to measure low self-control by a number of individual characteristics and behaviors: the urge to gratify desires immediately; lack of diligence and persistence in a course of action; lack of commitment to job, marriage, and children; lack of skills and planning; tendency to drink excessively; use of illegal drugs; and gambling. Self-control theory suggests that individuals respond to peer influence differentially, with those with weaker self-control more susceptible to negative peer influence.

When self-control is measured by lagged dependent variables such as previous self-behaviors and other observed self-characteristics, its effect can be difficult to interpret because prior measures of self-control are the product of a complicated interplay among genetic makeup and the social contexts that had preceded the outcome behavior. In this study, we measure self-control for alcohol use by genetic propensity for alcohol use. Since genetic propensity is formed at conception, which precedes much of social-contextual influences, measuring self-control via genetic propensity enables us to isolate genetic (and therefore more stable) parts of self-control and to investigate the interaction of this more stable part of self-control with peer effects.

Swing theory.—As can be seen in a discussion of the intellectual development of the model in Ellis et al. (2011), swing theory is an extension of the classic G × E interaction model of diathesis stress. The G × E interaction refers to the interdependence of genetic effects and environmental effects. In a G × E interaction, the size or even the presence of an environmental effect depends on a genetic effect, or vice versa. Diathesis refers to a hereditary disposition. A key element of a diathesis-stress model is that some individuals possess a biological vulnerability or a vulnerable allele that renders them more susceptible to an environmental risk. For example, Caspi and colleagues (2002) reported that maltreated male children in New Zealand with the 3R or 5R of VNTR in the MAOA gene were more vulnerable and more likely to engage in violent behavior than are maltreated children with other alleles of the VNTR.

The diathesis-stress model is quite intuitive. Similar models must have developed independently and repeatedly. An example is the long tradition of the idea of "frailty" in the study of human health and mortality (Clayton 1978; Vaupel, Manton, and Stallard 1979; Guo 1993). Frailty is analogous

to diathesis and suggests that some individuals are inherently more frail than others, even though frailty is generally unobserved. In the current study, individuals who have a genetic propensity for alcohol use may be more susceptible to the influence of a binge-drinking peer.

The swing model extends the diathesis-stress model in one key way, by hypothesizing that individuals with a medium genetic propensity are more likely to be swayed by an adverse environmental influence to develop a risk behavior than individuals at the extremes with a low and high propensity. Thus, swing theory overlaps diathesis-stress theory in the range of low and medium genetic propensities. The difference is that swing theory hypothesizes that some adverse environmental influences tend to be muted when genetic propensity becomes extremely high.

In the current study, swing theory predicts that individuals with a medium propensity for binge drinking are more likely to be influenced by a binge-drinking roommate than are individuals with a low or high propensity. Individuals with a very low genetic propensity tend not be affected because they are inherently disinterested in alcohol use. Individuals with a high genetic propensity tend not be influenced by peers because of already developed drinking habits.

A similar interaction pattern has been considered theoretically and empirically in diverse fields of study. Politics and political science pay major attention to swing voters who have not shown any party identification, whose political position may swing, and who may swing the outcome of an election. In democratic elections, concentrating campaign resources on swing voters is a common practice. For example, political scientists investigated whether an incumbent government often attempts to influence an election by spending funds in regions where a considerable portion of the population are swing voters, rather than investing resources in base voters of either side (Lindbeck and Weibull 1993; Dixit and Londregan 1996; Dahlberg and Johansson 2002).

A number of studies reported that animals with an intermediate level of an intrinsic characteristic reacted more strongly to a stimulus than those with a low or high level of that characteristic. In a rat study, three groups of nonaggressive, low-to-intermediate aggressive, and highly aggressive male rats were confronted with strange male intruders. When given low doses of ethanol, the nonaggressive and aggressive rats did not become aggressive, while the intermediate group of rats demonstrated a marked increase in the frequency of attack activities (Blanchard et al. 1987). In mice studies, animals with either a low or high level of the pituitary-adrenocortical hormones reacted nonaggressively to or avoided agonistic stimuli, whereas animals with an intermediate level of the hormones reacted aggressively to the same stimuli (Brain, Nowell, and Wouters 1971; Leshner et al. 1973; Candland and Leshner 1974; Leshner and Moyer 1975; Leshner, Moyer, and

Walker 1975). In a study of social rank and aggressive behavior in response to feeding among rhesus monkeys that lived in a large enclosed area, researchers observed increased levels of aggression among medium-ranking members but not among low-ranking or high-ranking members during feeding competition (Belzung and Anderson 1986).

Our proposed swing theory differs from extant G × E interaction theories such as the social push model (Raine 2002; Boardman, Daw, and Freese 2013) and the model of genetic differential sensitivity (Belsky 1997; Boyce and Ellis 2005; Belsky and Pluess 2009; Mitchell et al. 2011, 2013; Simons et al. 2011). Raine (2002) developed the social push model from the observation that the resting heart rate was especially low among antisocial individuals from privileged middle classes but not low social classes. The social push model hypothesizes that while a relative lack of adverse social factors in the middle classes makes biological effects "shine through," a much higher level of adverse social causes of antisocial behaviors essentially "camouflage" biological effects in low social classes. In a sense, the social push model could also be conceptualized as an extension of the diathesis-stress model. However, unlike the swing model, which focuses on the sensitivity of an environmental effect size, the social push model emphasizes the sensitivity of a biological effects size. In a social push model, a biological or genetic effect could shine through or be overwhelmed by social factors.

The model of genetic differential sensitivity makes two opposite predictions for the same individuals with a genetic vulnerability. In an unfavorable environment, the model is identical to that of the diathesis-stress model. However, in a favorable environment, these same individuals are more likely to experience positive outcomes. Thus, individuals with the same genetic makeup who are adversely affected by an unfavorable environment would benefit more from a favorable environment than those without that genetic makeup. A swing model does not consider these crossover effects.

Establishing Causality of Peer Effects

The empirical study of peer influence has two well-known complications. The first is endogeneity, which in this case means that observational studies cannot separate peer influence from friend selection (Kandel 1978; Manski 1993; Moffitt 2001; Mouw 2006). Studies of peer influence may find a similar level of binge drinking among friends, but this finding alone cannot differentiate between a scenario in which the similarity is due to peer influence and the scenario in which individuals are selected as friends because they exhibit similar behaviors and attitudes. In other words, these studies are unable to determine whether the similarity among friends is because one takes on the color of one's company or because birds of a feather flock

together. The second issue stems from the simultaneity of peer influence (Manski 1993; Moffitt 2001). Friends in a friendship dyad simultaneously influence each other. If the study design does not take the simultaneity into account, friend effects may be overestimated.

Random assignment of college roommates provides a rare opportunity to isolate peer influence. In the case of our college as well as many other colleges, roommate randomization is conditioned on students' housing preferences (e.g., dormitory/apartment location, number of roommates). The conditional random assignment provided by a housing lottery ensures that roommates are no more correlated than by chance in terms of their precollege drinking behaviors and other characteristics. To avoid simultaneity, we predict college outcomes of our subjects by their roommates' precollege behaviors before they began sharing a room in a residential hall.

Data from randomly assigned roommates in college have been used in studies of academic achievement (Sacerdote 2001; Zimmerman 2003; Foster 2006; Kremer and Levy 2008), fraternity membership (Sacerdote 2001), and drinking (Duncan et al. 2005). None of these roommate studies has considered genetic propensities and how outcomes are influenced by the interaction between peers' and self genetic propensities.

DATA, MEASURES, AND METHODS

Data Sources

Data sources consist of a discovery data set, ROOM, and a replication data set, Add Health. In ROOM, subjects were freshmen, sophomores, and juniors in a large U.S. public university in the spring semester of 2008. All subjects in this study had been randomly assigned roommates when they first entered the university. Students who were not randomly assigned such as those who requested a specific roommate or who participated in a themed housing program (e.g., foreign languages, health sciences, substance free) were excluded. About 15% of all students opted for a themed housing program.

When randomly assigning roommates, the university housing office placed data from applications into a large database, which was loaded into the software program RMS for random matching. Every student was then randomly assigned a unique RMS-ID number. After the first student had been placed in a room, the RMS program assigned his or her roommate as the next student in the chronological RMS-ID order with a compatible gender, smoking status, and type of requested room.

To assess to what extent our analytical sample in ROOM is representative of the undergraduate student population at the university, we compared the characteristics of ROOM with those of the entire undergraduate student population. The mean incomes are 4.62 and 4.75, respectively, for

the survey respondents from ROOM and the respondents of the Cooperative Institutional Research Program (CIRP). The nine income categories in both surveys were coded as follows: 1 ≤ $25,000, 2 = $25,000–$50,000, 3 = $50,000–$75,000, 4 = $75,000–$100,000, 5 = $100,000–$150,000, 6 = $150,000–$200,000, 7 = $200,000–$250,000, 8 = $250,000–$500,000, and 9 ≥ $500,000. Levels of father's education are 5.49 and 5.21, respectively, for ROOM and CIRP. The seven categories of father's education are coded as follows: 1 = middle school or less, 2 = some high school, 3 = high school graduate, 4 = some college, 5 = postsecondary school other than college, 6 = college degree, and 7 = graduate or professional coursework or degree. The proportions of males, Asian, Hispanics, African-Americans, whites, and others are 0.39, 0.07, 0.058, 0.12, 0.66, and 0.091 for ROOM, respectively, and 0.42, 0.073, 0.093, 0.092, 0.66, and 0.086 for the student population according to online information on Student Headcount by Level, Race/Ethnicity, and Sex, Fall 2011. The slightly larger proportion of African-American students in ROOM is due to our deliberate oversampling of these individuals for ROOM. q3

The information on alcohol use and socioeconomic background in ROOM was obtained via a web-based survey, which was completed by 2,664 (79.5%) of the eligible students. Of those who completed the web survey, 2,080 (78.7%) came to a campus office and provided a saliva sample. Students who did not live on campus, who were too young (under 18) to be included in the alcohol study, or who were in a study-abroad program in a foreign country for the semester were considered ineligible. Our final analysis sample included 2,006 students, or 1,003 pairs of randomly assigned roommates in which both roommates participated in the study. Of the 1,003 pairs, 694 had genotype data for both roommates and 309 had genotype data for only one of the two roommates. These 309 pairs of roommates were included in the analysis since only self genotype is necessary for our analyses.

Data were also drawn from the CIRP. These data were used to verify random assignment of roommates in ROOM. Each year, a large number of universities administer this freshman survey to entering students during orientation or registration. The survey gathers information from a range of student characteristics including a number of health behaviors. This part of our analysis only includes individuals in CIRP who are also roommates in ROOM. These two independent studies targeted the same student body.

Add Health is a longitudinal study of health behaviors among a school-based national sample of adolescents in grades 7–12 in 1994–95 in the United States (Harris et al. 2003). Eighty high schools were randomly selected from a stratified nationally representative sample of all public and private high schools in the United States. These strata were based on region of the country, urbanicity, school type (public, private, and parochial), and

racial composition. For each of the 80 high schools, the largest feeder school (typically a middle school or junior high) was targeted for recruiting. The final sample consisted of 134 schools. In wave 1, a self-administered in-school questionnaire was given to all seventh through twelfth graders attending these schools on a chosen day in 1994–95. About 90,000 students, or 77%, responded. The survey included questions on students' risky behaviors. Students were asked to nominate up to five male and five female friends from rosters of the high school and the feeder school. These nominated friends were themselves subjects of Add Health. Thus, information on binge drinking was reported by friends themselves rather than from ego projection. In addition to the in-school survey, an in-home interview was conducted in 1994–95, 1995–96, and 2002 (waves 1–3). The binge-drinking outcomes in Add Health in this analysis were from the in-home interview in wave 3. Add Health consists of 2,270 individuals whose saliva was gathered at Add Health wave 3 in 2002 and who have valid genotype and survey data. These individuals represent 87% of the 2,612 individuals whose saliva DNA was collected at wave 3.

In ROOM, high school drinking by a roommate was used as the peer influence on ego's binge drinking in college. To make the structure of analysis in Add Health as similar as possible to that in ROOM, friends' self-reported drinking at wave 1 in 1994–95 was used to predict ego's binge drinking at wave 3 in 2002. Because the study subjects of Add Health were age 12–18 at wave 1, we excluded those who were younger than 15 at wave 1 from Add Health analysis. Our final analysis sample of Add Health consisted of 1,612 respondents who were 15 years or older at wave 1.

Measures of Genotype

In ROOM, DNA was extracted according to the manufacturer's instructions from 2 milliliters of saliva, which contains buccal epithelial and white blood cells, collected from participants in an Oragene DNA collection kit. Our median DNA yield was 27.3 micrograms, with a minimum of 0 micrograms for six individuals and a maximum of 71.3 micrograms. DNA was plated for Illumina genotyping at 30 microliters at >50 nanograms/microliter. For ROOM, we designed an Illumina GoldenGate assay for 384 candidate single nucleotide polymorphisms (SNPs), including a set of 186 ancestral informative markers (AIMs; Enoch et al. 2006). Apart from the 162 AIMs, which were successfully genotyped out of the 186 targeted, another 186 SNPs in 28 genes were successfully genotyped, and these SNPs were selected because of their implications for behaviors such as alcohol use, smoking, risky sexual behavior, and aggression.

In Add Health, genomic DNA gathered at wave 3 in 2002 was isolated from buccal cells with an average yield of 58 ± 1 micrograms. The geno-

type data used in this analysis were based on an Illumina GoldenGate assay for 1,536 candidate SNPs including the same 186 AIMs (Enoch et al. 2006) targeted in ROOM. Excluding the 121 successfully genotyped AIMs of the 186 targeted, a total of 1,019 SNPs in 130 genes were available for analysis. The GoldenGate array of 1,536 was designed to include primarily genetic variants related to risky behaviors such as aggression, alcohol use, smoking, and illegal drug use. These variants include those in the 55 genes assembled by Maxson and colleagues to keep track of genes that have been shown to have an effect on aggression in mice studies (Maxson and Canastar 2003; Maxson 2009). Between ROOM and Add Health, 101 SNPs are common after excluding AIMs.

In both ROOM and Add Health, the bioancestry scores of Africans, Europeans, and East Asians were estimated by the Structure procedure using the AIMs (Pritchard, Stephens, and Donnelly 2000); the three scores for each individual in the three broad race/ethnic groups sum to one. These bioancestry measures are included in regression models to control for population stratification.

Measures of Binge Drinking and Controls

In the ROOM G × E interaction analysis, three binge-drinking measures were used as the outcome variables: a monthly count of binge-drinking episodes (1) during the fall semester of the first year of college, (2) during the past fall semester, and (3) during the past two weeks. The responses of "never," "less than once a month," "once or twice a month," "about once a week," "2–4 times a week," and "every day or almost every day" were coded 0, 0.5, 1.5, 4.3, 12.9, and 25, respectively. Since ROOM respondents were interviewed in the spring of the first, second, or third college years, the recall period for responses to the first question varies from a few months to over two years. The two binge-drinking measures used as outcome variables in the Add Health analysis were from wave 3: monthly binge-drinking episodes in the past two weeks and monthly count of binge-drinking days over the past year. These Add Health outcomes were coded in the same fashion as those in the ROOM analysis.

In ROOM, "roommate drank" was coded 1 for roommates who reported using alcohol in high school and 0 otherwise. Precollege drinking rather than college drinking was used to measure peer drinking in order to avoid the simultaneity of the roommates' drinking in college. Peer drinking measures were coded dichotomously to simplify G × E interaction analysis. In Add Health, peer drinking was based on self-reported drinking by nominated friends themselves at wave 1 in 1994–95 rather than drinking projected by egos. If peer drinking from nominated friends was missing, ego-reported friend drinking at wave 1 was used instead. "Friend drank" at

wave 1 was measured by an indicator variable coded 1 for those friends who drank beer, wine, or liquor or who got drunk over the past 12 months, and 0 otherwise.

In the ROOM analysis and as described below, the response variable in the false-discovery-rate (FDR) procedure and the stepwise regression was a binary indicator of alcohol use in college. The response variable in the model of genetic propensity in Add Health was also a binary variable based on alcohol use at wave 3.

The G × E interaction analysis for ROOM controlled for gender, father's education, mother's education, roommate's father's education, roommate's mother's education, family income, roommate's family income, weekly church attendance or religiosity, roommate's religiosity, total SAT score, roommate's total SAT score, GPA in college, roommate's GPA in college, having a nonwhite roommate, and fixed effects constructed from the housing preferences on the housing form. The G × E interaction models for Add Health controlled for gender, age at wave 3, PVT test score, parental education, family income, parental unemployment status, presence of two biological parents, household size, and religiosity. Both ROOM and Add Health controlled for bioancestry scores. The descriptive statistics are given in table 1. q4

Analytical Strategy

Our empirical challenges are to (1) estimate credible gene-environment interaction effects on youth binge drinking using 101 SNPs in 21 genes in ROOM and (2) determine whether these findings replicate in an independent data source (Add Health). We developed an approach for gene-environment interaction analysis in which genotype is measured by a relatively large number of genetic variables. Our overall strategy consisted of three stages. Stage 1 yields a polygenic propensity score ranging from 0 to 1. In stage 2, the propensity score is interacted with the precollege drinking behavior of the roommate to predict self binge drinking in college. The stage 1 analysis and the stage 2 analysis are applied to ROOM, the discovery data set. In stage 3, data from Add Health are used to replicate the G × E interaction effects from ROOM.

Stage 1 used a procedure of FDR (Benjamini et al. 2001; Benjamini and Yekutieli 2005) that selected SNPs out of the 186 available in ROOM. The FDR is a recently developed method of addressing multiple testing problems when a large amount of statistical tests are conducted simultaneously. The method was proposed as an alternative to more conservative approaches such as Bonferroni correction. The selected SNPs were individually (in separate regressions) predictive of binge drinking at the significance level of $P < .10$. Then, all FDR-selected SNPs were simultaneously

entered into a stepwise regression, which had the same binary binge-drinking outcome as in the FDR procedure. Only genetic variants with a *P*-value of .05 or smaller in the stepwise regression were retained and used to calculate the genetic propensity score for each individual, which was the predicted probability based on the final stepwise logistic regression. The FDR was first used to reduce the number of genetic variables in preparation for the stepwise regression. The process estimated a genetic propensity score for each individual that ranged from 0 to 1.

In the division of our sample into the low, medium, and high propensity groups, we had to balance competing considerations. On one hand, the low and high propensity groups should include individuals with quite low or quite high genetic propensities for alcohol use, in order to differentiate them from individuals with midrange genetic propensities. On the other hand, the low and high groups should be large enough to support reliable statistical inference. The final sets of cutoff points for ROOM are 0.2 and 0.8 and for Add Health, 0.3 and 0.8, where 0.2, 0.3, and 0.8 are the 20th, 30th, and 80th percentile of the propensity score, respectively.

Selection criteria were much more stringent in the stepwise regression than in the FDR procedure. While the FDR procedure estimated the effect of each genetic variant in a separate regression independent of the other genetic variants, the stepwise regression chose a set of genetic variants that were included in a regression simultaneously and that remained statistically significant simultaneously in a single regression. For this reason, the number of genetic variants that survived the stepwise regression was much smaller than those that survived the FDR procedure.

In stage 2, the gene-environment interaction analysis consisted of two sets of regression models. The first set (eq. [1]) compared the effect of peer drinking across groups with low, medium, and high levels of genetic propensity for alcohol use in three separate regression models. Then the G × E interaction was estimated in a single regression (eq. [2]).

Equation (1) provides an initial test of the swing theory, and the equation was estimated separately for the low, medium, or high propensity groups:

$$\text{self binge drinking}_{ij} = \beta_0 + \beta_1 \text{peer drank}_{ij} + \beta_2 \text{controls}_{ij} + \nu_j + e_{ij}, \quad (1)$$

where i indexes individuals, j indexes fixed effects cells, "peer drank" represents pairing with a roommate who drank in high school, and ν_j are fixed effects of the cells based on housing preferences. The inclusion of these fixed effects ensures that estimated peer influences are based solely on variation induced by random assignment. The swing theory predicts a larger peer effect in the medium propensity group than either the low or the high propensity group. The findings from equation (1) provide justification for proceeding to the analysis based on equation (2).

TABLE 1
DESCRIPTION OF THE VARIABLES IN ROOM AND ADD HEALTH

	ROOM			ADD HEALTH		
VARIABLE	Mean	SD	*N*	Mean	SD	*N*
Outcomes and measures of peer influences:						
Binge drinking first semester monthly episodes	2.28	4.01	1,696			
Binge drinking past semester monthly episodes	2.25	3.94	1,697			
Binge drinking past two weeks monthly episodes	2.33	3.91	1,696			
Roommate drank in high school (0 or 1)	.39		1,693			
Binge drinking past two weeks monthly episodes				2.04	4.18	1,612
Binge drinking days past year monthly count				1.56	3.90	1,604
Friend drinking at wave 1 (0 or 1)				.66		1,612
Control:						
European ancestry score	.78	.36	1,702	.69	.40	1,612
African ancestry score	.15	.32	1,702	.17	.34	1,612
Male	.39		1,703	.49		1,612
Age at wave 3				22.7	1.22	1,612
Parent unemployed				.045		1,409
Father's education	16.06	2.17	1,660			
Roommate's father's education	16.04	2.15	1,657			
Mother's education	15.91	1.98	1,695			
Roommate's mother's education	15.88	2.0	1,690			
Parental education, higher of the two parents						1,566
Less than high school				.126		197
High school				.284		445
More than high school				.590		924

Family income in $10,000	13.81	11.48	1,635		
Roommate's family income in $10,000	13.84	11.65	1,624		
Family income					1,275
0–20,000				.180	229
20,000–60,000				.550	701
>60,000				.271	345
Two bioparent presence				.613	1,612
Household size					1,612
≤2				.012	18
2–7				.845	1,362
>7				.144	232
Church attendance weekly	.193		1,685	.409	1,612
Church attendance weekly roommate	.206		1,683		
SAT	1,322	157	1,457		
Roommate's SAT score	1,328	147	1,443		
GPA	4.21	.48	1,671		
Roommate's GPA	4.21	.48	1,664		
Having nonwhite roommate	.33		1,703		
PVT test score					1,581
<90				.238	377
90–110				.471	744
110–150				.291	460

Equation (2) tests the swing theory in a G × E interaction analysis in a single regression model in which the low and high genetic propensity groups are combined into one group:

$$\text{self binge drinking}_{ij} = \beta_0 + \beta_1\left(\frac{\text{self medium}}{\text{peer drank}}\right)_{ij} + \beta_2\left(\frac{\text{self low high}}{\text{peer drank}}\right)_{ij} + \beta_3\left(\frac{\text{self low high}}{\text{peer nondrank}}\right)_{ij} + \beta_4\text{controls}_{ij} + \nu_j + e_{ij}, \quad (2)$$

where "peer nondrank" represents pairing with a roommate who did not drink in high school; "self medium" and "self low high" stand for "self in the medium genetic propensity group" and "self in the low or high genetic propensity group," respectively; the combination of "self medium/peer nondrank" was the omitted reference group; subscripts i and j are again indexed for individuals and housing preference cells, respectively; and ν_j are fixed effects of the cells. Both equations are estimated by ordinary least squares regression.

The G × E interaction model (2) includes dummy variables for three of the four combinations of self genetic propensity and roommate's precollege drinking behavior. Drawing on the initial G × E interaction findings from equation (1), individuals in the low and high propensity groups were combined into a single category in equation (2). The swing theory hypothesizes that β_1 is positive and statistically significant or that college students with a medium genetic propensity binge drank more when paired with roommates who drank in high school than did college students with a medium genetic propensity paired with a roommate who did not drink in high school.

The coefficient β_3 for "self low high / peer drank" in equation (2) can be estimated after omitting the combination of "self low high / peer nondrank." This model tests the hypothesis that a college student with a low or high genetic propensity would not increase binge drinking when paired with a roommate who drank in high school. Evidence for the test lends further support to the swing theory. q5

In 694 of the 1,003 roommate pairs, both roommates have DNA measures; in the rest of the roommate pairs (309 = 1,003 − 694), only one roommate has DNA measures. Within each of the 694 pairs, both individuals can be used as ego or used to construct the dependent variable. To avoid the arbitrariness of which of the two in a pair is used as an ego, we performed 500 analyses with each analysis randomly selecting one of the two members in a roommate pair to construct the response variable. This randomizing procedure was applied to the 694 of the 1,003 roommate pairs in which both roommates have DNA measures. The final re-

gression coefficients and *t*-statistics were averages over the coefficients and *t*-statistics from the 500 analyses.

In the analysis of the Add Health data, the correlation within sibling clusters was addressed by a generalized estimation equation model (Liang, Zeger, and Qaqish 1992). The missing values of nongenetic variables were imputed by the multiple imputation technique in both ROOM and Add Health (Rubin 1987). The five multiple completed data sets were then analyzed separately by statistical software SAS before the results were combined to produce the overall inference. Missing values in genotype data were imputed via MACH (Li et al. 2009; Marchini and Howie 2010). To address population stratification, all stepwise regressions and G × E interaction models controlled for ancestry scores of Africa and Europe (Pritchard et al. 2000).

RESULTS

To verify random assignment of roommates, we calculated within-dorm intraclass correlations via the linear and ordinal-mixed models (Searle, Casella, and McCulloch 1992) for 15 precollege responses obtained by CIRP. None of the CIRP responses between roommates correlated at $P < .05$ (table 2). We also calculated within-dorm intraclass correlation for two precollege measures of alcohol use in ROOM. Neither drinking nor binge drinking shows a statistically significant within-dorm correlation. These findings indicate that the precollege drinking measures (which are used as peer influence in G × E interaction analysis) as well as other precollege health behaviors are, indeed, uncorrelated among college roommates.

The FDR screening via a binary model of the 186 SNPs in 28 genes yielded 73 SNPs in 21 genes that significantly predicted alcohol use at the 10% level in ROOM (table A1). In the ROOM stepwise regression analysis, 10 of the 73 SNPs remained simultaneously significant at the 5% level. The 10 SNPs were from six genes: DRD2, MAOA, LMO3, TPH2, DBH, and DRD4 (table 3). Only five of these 10 SNPs were genotyped in Add Health: RS4245145 (DRD2), RS2242592 (DRD2), RS1125394 (DRD2), RS3027405 (MAOA), and RS7975434 (LMO3). Consequently, in the replication analysis based on Add Health, only these five were used in a logistic regression estimating genetic propensity for drinking. In Add Health, four of the five SNPs were simultaneously significant when the five were included in a single logistic regression model. These four SNPs' coefficients had the same sign and similar size as the same four SNPs in ROOM. We estimated the genetic propensity models of ROOM using the same set of five SNPs that were available in Add Health. All five main effects from the 5-SNP ROOM analysis were very similar to the same five main effects in the 10-SNP ROOM analysis (table 3). A second genetic propensity for drink-

TABLE 2
Within-Dormitory Intraclass Correlation

Variable	Intraclass Correlation Coefficient	$P > \|t\|$	N	Mixed Model
Precollege behaviors, CIRP:				
Had drank beer	.088	.32	1,122	Ordered logit
Had drank wine/liquor	.020	.80	1,119	Ordered logit
Smoked cigarettes	.000	. . .	1,127	Ordered logit
Physical exercise	.087	.19	1,100	Linear
Partying	.000	. . .	1,097	Linear
Religious service attendance	.000	1.00	1,127	Ordered logit
Felt depressed	.090	.30	1,125	Ordered logit
Frequency of volunteering	.000	1.00	1,122	Ordered logit
Frequency of community service	.013	.86	1,120	Ordered logit
Political view	.044	.52	1,091	Ordered logit
Hours socializing with friends	.000	. . .	1,100	Linear
Hours volunteering	.025	.75	1,088	Linear
Hours watching TV	.022	.72	1,094	Linear
Hours reading	.039	.58	1,092	Linear
Hours playing video games	.000	. . .	1,096	Linear
Precollege alcohol use, ROOM:				
High school drinking	.010	.87	1,298	Ordered logit
High school binge drinking	.000	. . .	1,345	Binary

Note.—Correlations are estimated by the linear and binary mixed model controlling for the fixed effects for 15 precollege responses in CIRP and two high school measures on alcohol use in ROOM for the purpose of checking random assignment of roommates in ROOM. The CIRP results are within-dorm correlations based on precollege CIRP responses; these CIRP subjects are also roommates in ROOM. The ROOM results are within-dorm correlations based on ROOM high school responses; these subjects are in ROOM.

ing based on the same procedure was constructed using the much larger genetic SNP set from Add Health, and this propensity score was constructed from 27 SNPs that were simultaneously significant at the 5% level (table A2).

Table 4 presents initial evidence on peer-by-genetic-propensity interaction from ROOM based on equation (1) that estimated a peer influence within each of the three levels of genetic propensity. The genetic propensity was based on the five SNPs identified in the ROOM stepwise regression. All peer effects were estimated after adjusting for a full set of controls including bioancestry scores. Only college students with genetic propensity scores in the middle range appeared to increase their binge drinking in response to roommate assignment. Pairing these students with a roommate who drank in high school increased binge-drinking episodes per month in the first semester of college, the past semester, and over the past two weeks by 0.95 ($P = .021$), 0.73 ($P = .069$), and 0.88 ($P = .022$), respectively. These are equivalent to 42%, 32%, and 38% of the overall average amount of

q8

TABLE 3
COEFFICIENTS AND *P*-VALUES OF THE SNPS SELECTED BY LOGISTIC STEPWISE REGRESSION FROM ROOM AND ADD HEALTH

SNP	Gene	ROOM (10 SNPs)	Add Health (5 SNPs)	ROOM (5 SNPs)
rs4245145	DRD2	.348 (.023)	.70 (<.0001)	.416 (.0048)
rs3027405	MAOA	.350 (.0029)	.058 (.54)	.369 (.0013)
rs2242592	DRD2	.281 (.0018)	.20 (.0023)	.340 (<.0001)
rs7975434	LMO3	−.241 (.013)	−.32 (<.0001)	−.269 (.0048)
rs1125394	DRD2	.443 (<.0001)	.24 (.0012)	.470 (<.0001)
RS7967586 . . .	TPH2	.841 (.012)		
RS12283680 . . .	DRD2	1.59 (.0043)		
RS1541332 . . .	DBH	.215 (.0059)		
RS3758653 . . .	DRD4	.238 (.014)		
RS12364283 . . .	DRD2	−.404 (.032)		
N		2,060	2,249	2,060

NOTE.—Out of the 10 SNPs selected from the discovery data set ROOM, only five were genotyped in the replication data set Add Health; replication was attempted only on these five SNPs. These five SNPs were tested again in ROOM. All FDR-selected SNPs were entered into the stepwise regression. The stepwise regression controlled for bioancestry scores to address population admixture. The larger sample size in this ROOM analysis than that in the ROOM G × E interaction analysis is because the latter analysis requires subjects in pairs of roommates. The larger sample size in this Add Health analysis than the Add Health G × E interaction analysis is because the Add Health analysis excluded those who dominated friends at wave 1 when they were younger than 15. Neither restriction is present in the estimation of main genetic effects. *P*-values in parentheses.

binge drinking reported in the sample. There was no evidence that individuals with either low or high genetic propensity were affected by roommate drinking history.

Peer influences in the Add Health data were estimated by relating wave 3 reports of binge drinking among a representative national sample of 21–26-year-olds to patterns of alcohol use among their friends reported in wave 1 when they were in middle or high school, seven years before wave 3. Regressions in table 5 repeated the table 4 analysis using data from Add Health and showed that key results from the ROOM data were replicated. When a genetic propensity for drinking was estimated from the five SNPs, having a drinking friend at wave 1 increased binge-drinking episodes by 1.1 per month over the two weeks before the report and 0.55 binge-drinking days per month over the past year at wave 3. The results from the 27-SNP genetic propensity were similar to the 5-SNP results: nominating a drinking friend at wave 1 was associated with 0.76 more binge-drinking episodes per month and 0.72 more binge-drinking days per month over the past year at wave 3. All of these increases happened to youth with a medium genetic propensity for drinking. Youth in the low or high propensity groups were not influenced by friends' drinking habits.

q9

TABLE 4

FULL MODELS OF PEER INFLUENCE BY GENETIC PROPENSITY INTERACTION ON BINGE DRINKING: ROOM ANALYSIS

	First Semester Binge			Past Semester Binge			Past Two Weeks Binge		
	Low (0–.2)	Medium (.2–.8)	High (.8–1)	Low (0–.2)	Medium (.2–.8)	High (.8–1)	Low (0–.2)	Medium (.2–.8)	High (.8–1)
Roommate drank in high school	.415	*.952**	−.296	.271	*.734+*	−.514	−.540	*.882**	−.629
P-value	(.43)	(*.021*)	(.771)	(.63)	(*.069*)	(.508)	(.399)	(*.022*)	(.42)
Respondent characteristics:									
Female	−.052	−1.38**	−1.42	−.233	−1.52***	−.956	−.936	−1.56***	−1.19
Father's education	−.109	.048	−.047	−.094	.058	−.126	−.070	.056	−.157
Mother's education	.010	.030	.004	.003	−.004	−.061	.079	.034	.016
Family income $10,000	.015	.037*	.020	.038	.042*	.046	.009	.045*	.043
GPA	−.502	−.494	−.152	−.324	−.557	−.239	−.154	−.524	.216
SAT/100	.170	−.038	−.027	.194	−.061	−.007	.146	−.034	−.132
Nonwhite roommate	−.244	.107	−.552	−.127	.091	−1.09	.843	.317	−.776
Church attendance weekly	−1.21 *	−.1.64**	−1.36	−1.38*	−1.16**	−1.05	−.552	−1.44**	−1.44
Roommate characteristics:									
Fathers' education	−.153	−.007	−.046	−.099	−.009	−.163	−.119	−.041	−.153
Mother's education	.012	−.061	−.125	.005	.069	−.033	−.006	.044	.129
Family income $10,000	−.012	.001	.074*	−.003	−.015	.063*	−.002	−.006	.046
GPA	−.135	.362	.366	−.370	.439	.567	−.428	.432	.510
SAT/100	.101	.010	.212	.043	−.097	.163	.186	.051	.306
Church attendance weekly	−.067	.353	−.026	−.076	.091	−.459	−.147	−.054	−.560
Bioancestry (African)	−.245	.542	−2.05	−.014	.485	1.32	−.365	.126	−.977
Bioancestry (European)	1.09	1.41+	1.65	1.16	1.48+	5.12	1.48	1.35+	1.85
N	202	674	127	202	674	127	202	674	127

NOTE.—Peer effects are estimated separately at low, medium, or high genetic propensity for alcohol use. Genetic propensity for binge drinking was measured by five SNPs from the results from a stepwise logistic regression (table 3). The predicted genetic propensity score ranging 0–1 is used to divide the entire sample into three groups with two cutoff points at .2 and .8, which are 20th and 80th percentiles of the propensity score, respectively. Effects of primary interest italicized. *P*-values are provided only for these coefficients.

\+ $P < .10$.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

The findings based on equation (1) provide suggestive evidence for the swing theory. To increase statistical power for testing the swing theory, we combined the sample of low genetic propensity and the sample of high genetic propensity into a single category. Equation (2) tested the swing theory in a single equation.

In table 6, the G × E interaction analysis based on ROOM revealed that when paired with a roommate with a drinking history in high school, college students with a medium propensity reported 0.76 ($P = .036$), 0.59 ($P = .101$), and 0.82 ($P = .019$) more binge-drinking episodes per month for the first semester, the past semester, and the past two weeks, respectively, relative to students with roommates who did not drink in high school. Separate regression models testing the peer influence among individuals with a low or high propensity showed three much smaller coefficients of 0.17, 0.08, and −0.29 and much larger *P*-values of .69, .84, and .48, respectively. A bootstrapping analysis was performed to test whether a pair of roommate effects is statistically different from one another—a roommate effect associated with a medium genetic propensity (e.g., 0.764 for the first semester in table 6) versus a roommate effect associated with a low or high genetic propensity (e.g., 0.167 for the first semester in table 6). Respectively, the analysis yields 95% confidence intervals of (−0.037, 1.11), (.0004, 1.089), and (0.27, 1.35) for binge drinking in the first semester, past semester, and past month, supporting the conclusion in two of the three cases that those with a medium genetic propensity are more susceptible to peer influence than those with a low or high propensity. **q10** **q11**

Table 7 shows that the findings in ROOM described in table 6 largely replicate in Add Health. For individuals with a medium genetic propensity, having a drinking friend at wave 1 was associated with 1.03 ($P = .0007$) and 0.65 ($P = .038$) more binge-drinking episodes per month for the past two weeks for the 5-SNP propensity and the 27-SNP propensity, respectively, than those who did not report a drinking friend. For those with a low or high propensity, the two estimated peer effects for the 5-SNP propensity and the 27-SNP propensity were much smaller (0.22 with $P = .42$ and 0.48 with $P = .08$, respectively). The findings for the second binge-drinking measure of days over the past year in Add Health are similar, with positive and statistically significant peer effects (0.54 with $P = .055$ and 0.62 with $P = .03$) only found among those with a medium genetic propensity. The peer effects for those with a low or high propensity were smaller and nonsignificant (.43 with $P = .12$ and .34 with $P = .22$). The findings in tables 6 and 7 are summarized in figure 2. **q12**

To further test the robustness of our findings, we repeated the ROOM and Add Health analyses while setting the *P*-value to .1 and 1.0, respectively, when selecting SNPs for stepwise regression. One set consists of all available SNPs after pruning, or the deletion of highly correlated SNPs;

TABLE 5

Full Models of Peer Influence by Genetic Propensity Interaction on Binge Drinking: Add Health Wave 3

	Binge Drinking Past Two Weeks Monthly Episodes						Binge Drinking Past Year Monthly Days					
	Low (0–.3)		Medium (.3–.8)		High (.8–1)		Low (0–.3)		Medium (.3–.8)		High (.8–1)	
	5 SNPs	27 SNPs	5 SNPs	27 SNPs	5 SNPs	27 SNPs	5 SNPs	27 SNPs	5 SNPs	27 SNPs	5 SNPs	27 SNPs
Friend drank	−.12	.52	*1.1***	*.76**	.41	.35	.21	.06	*.55**	*.72**	.71	.65
P-value	(.75)	(.074)	(*.0007*)	(*.017*)	(.30)	(.53)	(.55)	(.84)	(*.059*)	(*.015*)	(.13)	(.24)
Age	−.032	.203	−.135	−.38***	−.49**	−.29	.067	−.026	−.22*	−.21*	−.51**	−.42*
Male	1.8***	1.4***	1.8***	1.8***	1.6***	2.0***	1.6***	1.2***	1.4***	1.8***	2.2***	2.2***
European ancestry	.180**	.068	.037	.088	.131	−.062	.155**	.131**	.092*	.063	.002	−.081
African ancestry	.009	−.008	.056	.015	−.2440	.217	.074	.025	.073	.114	.149	−.224
Cognitive score (90–110):												
<90	.150	−.251	.295	.853	.383	−.631	−.215	−.533	.465	.553	−.594	−.498
110–150	−.716	−.036	.443	−.376	−.360	.373	−.116	.272	.207	−.312	−.460	−.016
Parental unemployment	−.230	.165	−.452	−.288	1.070	.560	.368	.516	−.695*	−.380	.769	.209

Parental education (high school):												
<High school	−.100	−.556	−1.017*	.351	1.424	−.702	.753	−.215	−.883*	.162	.771	.119
>High school	.316	.127	−.110	.101	.568	.436	.409	−.241	−.307	.139	−.173	.166
Family income (20,000–60,000):												
0–20,000	−.602	−.559	−.594	−.481	−.079	−.271	−.608	−.48	−.685	−.607	−.785	−1.220
>60,000	.151	.015	.014	.143	.060	−.375	−.058	−.059	.132	.028	−.370	−.536
Two biological parents	−.514	−1.14**	−.154	.072	.140	.652	−.049	−.60	−.298	.134	−.157	−.489
Household size (3–6):												
1–2	.940	−.879	1.114	1.081	.050	1.327	1.546	.124	−1.098	.202	−.959	−2.98**
>7	−.022	.390	.461	−.515	−.951	−.23	−.188	−.02	.244	.093	−.459	−1.12**
Church weekly	.215	−.259	−.392	−.358	−.96*	−.547	.329	.271	−.252	−.268	−.330	−.503
N	485	478	706	830	421	304	482	474	704	827	418	303

NOTE.—Peer effects are estimated separately at low, medium, or high genetic propensity for alcohol use. Genetic propensity for binge drinking was measured by the same five SNPs from the ROOM stepwise regression (table 3). The predicted genetic propensity score ranging 0–1 is used to divide the entire sample into three groups with two cutoff points at .3 and .8, which are 30th and 80th percentiles of the propensity score, respectively. Effects of primary interest italicized. *P*-values are provided only for these coefficients.

+ $P < .10$.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

TABLE 6
FULL MODELS OF PEER INFLUENCE BY GENETIC PROPENSITY INTERACTION ON BINGE DRINKING WITH PEER EFFECT ESTIMATED IN A SINGLE REGRESSION MODEL: ROOM

	First Semester Binge		Past Semester Binge		Past Two Weeks Binge	
	Medium	Low or High	Medium	Low or High	Medium	Low or High
Roommate drank/self medium	*.764**	.746+	*.585+*	.641	*.821**	.445
P-value	(*.036*)		(*.101*)		(*.019*)	
Roommate nondrank/self low or high	.018	. . .	−.057	. . .	.376	. . .
Roommate nondrank/self medium	. . .	−.018	. . .	.057	. . .	−.376
Roommate drank/self low or high	.184	*.167*	.026	*.083*	.089	*−.288*
P-value		(*.69*)		(*.84*)		(*.48*)
Respondent characteristics:						
Female	−1.09***	−1.09***	−1.02***	−1.02***	−1.29***	−1.29***
Father's education	.005	.005	.011	.011	−.017	−.017
Mother's education	.025	.025	−.014	−.014	.010	.010
Family income $10,000	.029*	.029*	.039**	.039**	.040***	.040***
GPA	−.328	−.328	−.385	−.385	−.305	−.305
SAT/100	.018	.018	.023	.023	.001	.001
Nonwhite roommate	−.175	−.175	−.211	−.211	.078	.078
Church attendance weekly	−1.39***	−1.39***	−1.34***	−1.34***	−1.22***	−1.22***

Roommate characteristics:						
Fathers' education	−.032	−.032	−.033	−.033	−.099	−.099
Mother's education	−.052	−.052	.033	.033	.071	.071
Family income $10,000	.011	.011	.003	.003	.005	.005
GPA	.205	.205	.266	.266	.306	.306
SAT/100	.059	.059	−.033	−.033	.095	.095
Church attendance weekly	.161	.161	−.038	−.038	−.128	−.128
Bioancestry (African)	−.042	−.042	.071	.071	−.432	−.432
Bioancestry (European)	1.39*	1.39*	1.42*	1.42*	1.31	1.31

NOTE.—Each column presents the coefficients from a single regression model. The three "medium" models test the effect of pairing with a roommate who drank in high school relative to pairing with a roommate who did not drink in high school given self medium genetic propensity. In contrast, the three "low or high" models test the same effect given self low or high genetic propensity. Genetic propensity for alcohol use was measured by five SNPs from the results for a stepwise logistic regression (table 3). Effects of primary interest italicized. *P*-values are provided only for these coefficients. $N = 1{,}003$.

+ $P < .10$.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

TABLE 7
Full Models of Peer Influence by Genetic Propensity Interaction on Binge Drinking with Peer Effect Estimated in a Single Regression Model: Add Health Wave 3

	Binge Drinking Past Two Weeks Monthly Episodes				Binge Drinking Past Year Monthly Days			
	Medium		Low or High		Medium		Low or High	
	5 SNPs	27 SNPs	5 SNPs	27 SNPs	5 SNPs	27 SNPs	5 SNPs	27 SNPs
Friend drank/self medium	*1.026***	*.648**	.405	.833**	*.540**	*.619**	.352	.670*
P-value	(*.0007*)	(*.038*)			(*.055*)	(*.03*)		
Friend nondrank/self low or high	.621*	−.184	. . .	. . .	.188	−.052	. . .	. . .
Friend nondrank/self medium	. . .	. . .	−.621*	.184	. . .	. . .	−.188	.052
Friend drank/self low or high	.840**	.299	*.219*	*.484*	.619*	.285	*.431*	*.337*
P-value			(*.42*)	(*.08*)			(*.12*)	(*.22*)
Age	−.191*	−.193**	−.191*	−.193**	−.206*	−.210**	−.206*	−.210**
Male	1.79***	1.78***	1.79***	1.78***	1.69***	1.69***	1.69***	1.69***
European ancestry	.102**	.096**	.102**	.096**	.094**	.088**	.094**	.088**
African ancestry	−.007	.002	−.007	.002	.020	.027	.020	.027
Cognitive score (90–110):								
<90	.243	.266	.243	.266	−.001	.012	−.001	.012
110–150	−.120	−.100	−.120	−.100	−.086	−.078	−.086	−.078
Parental unemployment	−.080	−.055	−.080	−.055	−.026	−.011	−.026	−.011

Parental education (high school):								
<High school	−.145	−.168	−.145	−.168	.040	.031	.040	.031
>High school	.179	.174	.179	.174	−.007	−.006	−.007	−.006
Family income (20,000–60,000):								
0–20,000	−.456	−.438	−.456	−.438	−.672	−.659*	−.672	−.659*
>60,000	.066	.071	.066	.071	−.050	−.056	−.050	−.056
Two biological parents	−.208	−.218	−.208	−.218	−.189	−.208	−.189	−.208
Household size (3–6):								
1–2	.303	.281	.303	.281	−.228	−.207	−.228	−.207
>7	−.084	−.066	−.084	−.066	−.095	−.094	−.095	−.094
Church weekly	−.347	−.364	−.347	−.364	−.148	−.153	−.148	−.153
N	1,612	1,612	1,612	1,612	1,604	1,604	1,604	1,604

NOTE.—Each column presents the coefficients from a single regression model. The three "medium" models test the effect of pairing with a friend who drank at wave 1 relative to pairing with a roommate who did not drink at wave 1 given self medium genetic propensity. In contrast, the three "low or high" models test the same effect given self low or high genetic propensity. Genetic propensity for alcohol use was measured by five SNPs at first to replicate the findings from ROOM and then by 27 SNPs. Effects of primary interest italicized. *P*-values are provided only for these coefficients.

+ $P < .10$.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

this procedure amounts to setting $P < 1.0$. ROOM and Add Health have 138 and 54 SNPs for analysis, respectively, when $P < 1$. The second set was obtained by setting $P < .1$, and this set consists of 71 and 26 SNPs in ROOM and Add Health, respectively. These four sets of G × E interaction findings are consistent with the two sets of G × E interaction findings that are based on $P < .05$ (tables A3 and A4). **q13**

DISCUSSION

In this study, we tested the swing theory for a particular form of G × E interaction. On average, having a drinking peer increased binge drinking by 0.5–1.0 episodes per month as compared to having a nondrinking peer, which amounts to an increase of 20%–40% in the average number of binge-drinking episodes. Consistent with the swing theory, the peer influence was found only among youth with a medium genetic propensity for alcohol use; youth with low or high propensities were not influenced by peer drinking.

This G × E interaction finding for youth binge drinking was replicated across ROOM and Add Health. Within each study, the same G × E interaction finding was replicated across more than one binge-drinking measure. Within Add Health, the G × E interaction finding was replicated between analyses based on the 5-SNP and the 27-SNP propensities. The G × E interaction findings were tested for robustness by using two larger sets of SNPs in ROOM and Add Health in which the *P*-value is set to .1 and 1.0, respectively, when selecting in the stepwise regression. Consistently, only individuals with a medium genetic propensity for alcohol use are significantly influenced by peer drinking.

The empirical support of the swing theory has an intuitive explanation. Youth with a low genetic propensity are intrinsically uninterested in and unaffected by drinking around them, and youth with a high genetic propensity for alcohol use are likely to have already developed the habit of consuming alcohol excessively with or without peer drinking. Youth with a medium genetic propensity are thus the group that is most susceptible to a peer effect.

The MAOA gene is known to be implicated in risky behaviors; however, the MAOA variants are difficult to investigate in a combined sample of males and females for two reasons. First, males have one X chromosome and females have two. For any SNP on the X chromosome, a male is hemizygous and has zero or one allele, and a female has zero, one, or two alleles. These differences make it difficult to code males and females consistently. The second difficulty arises from the process of X inactivation among females. In about three-quarters of loci on the X chromosome, only one allele out of each pair of alleles is expressed; the other allele is inactivated (Ober, Loisel, and

Gilad 2008). Recently, a number of approaches have been proposed to address the difficulties (Clayton 2008; Konig et al. 2014). In the current study, we used an approach based on the assumption that the effect of the single allele in males is equivalent to that of the two alleles in females (Clayton 2008). In a SNP analysis, this approach amounts to coding the value of 1 as 2 in males. All our findings related the X chromosome are based on this approach. We also repeated the analysis after removing the MAOA variant. The results remain similar.

The reliability of binge-drinking measures remains an issue. There is always a possibility that self-report of binge drinking is tinted by factors such as social pressure. However, in the current study, we obtained a drinking measure for 1,094 individuals who were subjects in both ROOM and CIRP, and the within-same-person correlation on the drinking measure is 0.86 and highly statistically significant. ROOM and CIRP are two independent studies sponsored and executed by two different sets of researchers at different times. This finding indicates that the study subjects report their drinking with a remarkable level of consistency.

We conducted sensitivity analyses that control for self precollege drinking in our models reported in tables 6 and 7; the main findings remain unaffected with the additional control. Among other things, the precollege drinking controls for factors such as social pressure that is constant across different measures of alcohol use.

Three out of the four SNPs that are replicated between ROOM and Add Health are in the DRD2 gene, and the fourth SNP is in the LMO3 gene. The three SNPs in DRD2 are independently associated with alcohol use because the three effects are adjusted for one another in a single regression model (table 3). The dopamine D2 receptor (DRD2) gene located on chromosome 11 q22–q23 encodes the dopamine D2 receptor. Because of its key role in the dopaminergic system, DRD2 is a prime suspect in investigations of genetic links with risky behaviors including alcoholism. The DRD2 antagonist haloperidol has long been used to treat aggressive behavior in psychotic patients. Animal models implicated DRD2 in ethanol preference (Crabbe et al. 1999). A recent major multistage genome-wide association study of 36,989 cases and 113,075 controls prominently highlights the role of DRD2 in schizophrenia (Ripke et al. 2014).

LMO3 belongs to the LIM-only protein family with a function to modulate transcription by using its two tandem LIM domains to bind to DNA-binding proteins (Kadrmas and Beckerle 2004). The relationship between LMO3 and alcoholism has been studied for many years in animals. The fruit fly Drosophila melanogaster has been used to identify novel genes that affect behavioral responses to ethanol. Several studies found that reduced dLMO expression led to increased sensitivity to the sedating effect of eth-

anol and a decreased level of ethanol consumption, whereas the increased dLMO expression had the opposite effect both in flies and mice, suggesting that LMO3 may play an important role in alcohol preference in invertebrate systems and in mammals (Tsai et al. 2004; Lasek et al. 2011). It is speculated that LMO3 may affect behavioral responses to ethanol in humans through its ability to regulate transcription, which, in turn, can affect the patterning of certain brain structures such as the cortex or amygdala (Bulchand, Subramanlan, and Tole 2003; Remedios, Subramanian, and Tole 2004). The subtle changes in brain structures may later affect behavior responses to ethanol (Lasek et al. 2011).

Our study has important and specific implications for sociological research that focuses on understanding sociological influences. Empirical evidence for social learning is much more discoverable when a swing theory is incorporated into the social learning theory. A swing theory allows social learning to depend on genetic propensity (figs. 1 and 2). The swing theory is based on the idea that a social environmental influence such as peer influence has a different effect on different individuals, depending on an individual's genetic propensity for the trait. Ignoring the genetic propensity leads to estimating a peer effect averaged over individuals with various levels of genetic propensity. Such an average effect is generally much smaller or statistically nonsignificant, resulting in a rejection of a peer effect and the social learning theory.

In the current analysis of ROOM data (table 6), the estimates of a peer effect (0.764, 0.585, and 0.821, respectively) for youth with a medium genetic propensity is about five to eight times as large as those (0.167, 0.083, and −0.288, respectively) for youth with a low or high genetic propensity. Without taking into account genetic propensity, we would reject or find much weaker support for the social learning theory. More generally, many social-contextual influences may be conditional on genetic propensity. In such cases, G × E interaction analysis would be a much more effective way of discovering such effects than estimating an average effect.

Our study design of a natural experiment is crucially important in the test of sociological theories. In the test of social learning theory in the current study, we show that binge drinking could, indeed, be causally influenced by peer drinking. Without establishing causality, the social learning theory cannot be evaluated, and social policies based on social learning theory cannot be implemented.

Our study demonstrates how rGE may be successfully addressed. The G × E interactions based on observational data risk bias from gene-environment correlations (rGE; Jaffee and Price 2007; Wagner et al. 2013). The rGE produces environmental effects that are induced, in part, by genes, leading to ambiguity in interpreting G × E interaction. And rGE is

an extremely thorny issue for G × E interaction studies: rGE is to studies of G × E interaction like friend selection is to studies of peer influence and, more generally, like endogeneity is to causal inferences in social science studies. Just as randomized experiments provide a solution to endogeneity, the randomly assigned roommates in ROOM provide a solution to the thorny issue of rGE. The randomization guarantees that the roommates' precollege behavior is uncorrelated and that the peer influence is exogenous and uncorrelated with self genetic propensity for alcohol use.

In addition to the rGE between genetic propensity and peer influence for binge drinking, rGE between genetic propensity and other environmental factors such as parental unemployment, parental education, family income, family structure, and churchgoing, which were included as controls, was investigated. None of the rGE is statistically significant in ROOM or Add Health.

Randomized experiments protect against rGE, but questions remain about whether the findings from natural experiments can be generalized. In this study, we estimated peer influence by genetic propensity interaction for binge drinking in both experimental data (ROOM) and observational data (Add Health) for assessing the generalizability of the experimental ROOM study. Although peer influence from ROOM and Add Health was based on two different study designs, the empirically estimated peer effects conditional on genetic propensity from ROOM and Add Health are similar.

The research community generally assumes that peer effects estimated from observational studies overstate peer causation because of peer self-selection. That our findings from self-selected school friends are similar to those obtained from randomly assigned roommates came as a surprise and suggested that the biases may be modest, at least in the case of binge drinking. Our conclusions on this issue are, of course, provisional and should be considered as a first attempt on the crucial issue of generalizing experimental findings to observational studies or even vice versa. Future efforts should focus on designing observational studies that are as comparable as possible to experimental studies so that the two sets of findings can be compared with more confidence.

Our G × E interaction analysis is focused on the effects of peers and on the specific question, At what level of self genetic propensity do peers exert the strongest influence on binge drinking? Our G × E interaction analysis is also relevant to research that focuses on the effects of genes. This analysis demonstrates that the effects of genes on binge drinking are larger when one's friends are drinkers rather than nondrinkers. This finding suggests that understanding genetic origins of complex human traits may often require an understanding of environmental circumstances under which the relevant genes are operating.

Appendix

TABLE A1
DETAILS OF GENETIC DATA TESTED IN FDR PROCEDURE IN THE DISCOVERY DATA SET ROOM

Gene and Chromosome	Number of SNPs Tested	SNPs Selected by FDR
ADH1A, 4	2	rs182609 rs4147531 (2)
ADH1B, 4	10	rs1159918 rs1229982 rs7673353 (3)
ALDH2, 12	11	rs10849970 rs2158029 rs2238151 rs671 rs7296651 rs7311852 (6)
ANKK1, 11	1	0
ARVCF, 22	1	rs5993891 (1)
BDNFOS, 11	1	0
CHRM2, 7	5	rs1455858 rs7357341 (2)
CHRNA4, 20	1	rs2236196 (1)
CHRNB2, 1	1	0
CNR1, 6	1	0
COMT, 22	9	rs174696 rs739368 (2)
DBH, 9	7	rs1541332 rs3025410 rs77905 (3)
DDC, 7	3	rs1451371 rs1470750 rs998850 (3)
DEAF1, 11	1	0
DRD2, 11	26	rs1076563 rs1079596 rs11214605 rs1125394 rs12283680 rs12364283 rs2242592 rs2471857 rs2587548 rs2734833 rs4245145 rs4581480 rs7109897 (13)
DRD4, 11	5	rs11604855 rs1800443 rs3758653 rs916457 (4)
FTO, 16	8	rs10521303 rs6499640 (2)
GABRA2, 4	15	rs16859292 rs16859325 rs16859348 rs6857343 rs7678520 (5)
HTR1B, 6	17	rs1213366 rs13212041 (2)
HTR2A, 13	6	rs6304 (1)
LMO3, 12	13	rs11057005 rs16912030 rs16912043 rs7975434 (4)
MAOA, X	11	rs2072744 rs3027405 rs5905859 rs5906729 rs5906883 (5)
MAOB, X	8	rs1040399 rs12394221 rs17462 rs1799836 rs2239441 rs3027459 rs6520902 rs9887047 (8)
SLC18A2, 10	2	rs363333 (1)
SLC6A4, 17	17	rs2054848 rs9903602 (2)
TPH21, 7	10	rs1386483 rs2171363 rs7967586 (3)
TTC12k, 11	1	0
TXNRD2, 22	3	0
Total	186	73

TABLE A2

SNPs Selected by Stepwise Regression from the Entire Panel: Add Health

SNP	Gene	Coefficient	*P*-Value
rs1008098	OPCML	.209	.0095
rs10456876	FYN	.1393	.0326
rs10865408	TACR1	.1747	.0254
rs10894669	OPCML	−.2566	.0036
rs11015015	GAD2	−.1721	.0273
rs11609535	LMO3	.1916	.0263
rs12514354	CAMK2A	.2075	.0043
rs13245899	MUC3B	−.2622	.0029
rs1952586	ESR2	.2424	.0083
rs2000589	OPCML	−.203	.0026
rs2158029	ALDH2	.3572	.008
rs2161382	TRPC7	.2027	.0056
rs238300	CTNNBL1	.1634	.0147
rs324576	CHRM2	.2341	.02
rs376063	APP	.2204	.0179
rs4578395	OPCML	.1663	.0325
rs5911570	GRIA3	−.1519	.0112
rs6869634	CAMK2A	.2869	.0002
rs7135281	LMO3	.2179	.0016
rs7195954	FAM86A	−.1581	.0307
rs759588	TACR1	−.1807	.0119
rs762513	FAM50A	−.3051	.0423
rs7805828	IL6	.1784	.0078
rs7885398	MAOA	.5508	.001
rs806368	CNR1	−.1646	.0329
rs827419	ESR1	−.1981	.0082
rs985933	HTR2A	.136	.0381

Note.—SNPs shown here exclude AIMS. Two of the three bioancestry scores are included in the regression as controls. All 27 SNPs are simultaneously statistically significant at the 5% level in a single regression.

TABLE A3

FULL MODELS OF PEER INFLUENCE BY GENETIC PROPENSITY INTERACTION ON BINGE DRINKING: ROOM, WITH ALTERNATIVE SETS OF SNPS

	FIRST SEMESTER BINGE		PAST SEMESTER BINGE		PAST TWO WEEKS BINGE	
	Medium	Low or High	Medium	Low or High	Medium	Low or High
138 SNPs:						
Roommate drank/self medium	*.694** (*.029*)	.490	*.640** (*.044*)	.434	*.568+* (*.059*)	.356
Roommate nondrank/self low or high	.378	. . .	.112	. . .	.060	. . .
Roommate nondrank/self medium	. . .	−.204	. . .	−.206	. . .	−.212
Roommate drank/self low or high	.204	*.174* (*.695*)	.206	*−.094* (*.810*)	.212	*−.152* (*.713*)
71 SNPs:						
Roommate drank/self medium	*.701** (*.026*)	.460 (.206)	*.591+* (*.064*)	.439 (.200)	*.561+* (*.065*)	.418 (.209)
Roommate nondrank/self low or high	.404 (.344)	. . .	.125 (.741)	. . .	−.013 (.973)	. . .
Roommate nondrank/self medium	. . .	−.241 (.460)	. . .	−.152 (.640)	. . .	−.143 (.651)
Roommate drank/self low or high	.241 (.460)	*.163* (*.730*)	.152 (.640)	*−.028* (*.945*)	.143 (.651)	*−.155* (*.705*)

NOTE.—Genetic propensity is based on two alternative larger sets of SNPs. When setting $P < 1$, one set consists of 138 SNPs that have remained after pruning, which is a statistical procedure that deletes highly correlated SNPs. The other set consists of 71 SNPs obtained when setting $P < .1$. Models include the same controls as in those in table 6. See table 6 note for interpretation of these results. Effects of primary interest italicized. *P*-values in parentheses; $N = 1{,}003$.

+ $P < .10$.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

TABLE A4
FULL MODELS OF PEER INFLUENCE BY GENETIC PROPENSITY INTERACTION ON BINGE DRINKING: ADD HEALTH WAVE 3, WITH ALTERNATIVE SETS OF SNPS

	Binge Drinking Past Two Weeks Monthly Episodes		Binge Drinking Past Year Monthly Days	
	Medium	Low or High	Medium	Low or High
54 SNPs:				
Friend drank/self medium	*.555**	.100	*.598**	.383
	(*.039*)		(*.033*)	
Friend nondrank/self low or high	.455	. . .	.215	. . .
Friend nondrank/self medium	. . .	−.455	. . .	−.215
Friend drank/self low or high	.975***	*.520*	.557*	*.342*
		(*.104*)		(*.233*)
N	1,612	1,612	1,604	1,604
26 SNPs:				
Friend drank/self medium	*.620**	.388	*.553+*	.440
	(*.028*)		(*.057*)	
Friend nondrank/self low or high	*.232*	. . .	.113	. . .
Friend nondrank/self medium	. . .	−*.232*	. . .	−.113
Friend drank/self low or high	.732*	*.501+*	.514+	.401
		(*.081*)		(.142)
N	1,612	1,612	1,604	1,604

NOTE.—Models include the same controls as in those in table 7. See table 7 note for interpretation of these results. Genetic propensity is based on two alternative larger sets of SNPs. When setting $P < 1$, one set consists of 54 SNPs; the other set consists of 26 SNPs when setting $P < .1$. Effects of primary interest italicized. *P*-values in parentheses.

\+ $P < .10$.
* $P < .05$.
** $P < .01$.
*** $P < .001$.

REFERENCES

Akers, Ronald. 1985. *Deviant Behavior: A Social Learning Approach*, 3d ed. Belmont, Calif.: Wadsworth.

———. 2001. "Social Learning Theory." Pp. 192–210 in *Explaining Criminals and Crime: Essays in Contemporary Criminological Theory*, edited by Raymond Paternoster and Ronet Backman. Los Angeles: Roxbury.

Bandura, Albert. 1971. *Social Learning Theory*. New York: General Learning.

Belsky, Jay. 1997. "Theory Testing, Effect-Size Evaluation, and Differential Susceptibility to Rearing Influence: The Case of Mothering and Attachment." *Child Development* 68 (4): 598–600.

Belsky, Jay, and Michael Pluess. 2009. "Beyond Diathesis Stress: Differential Susceptibility to Environmental Influences." *Psychological Bulletin* 135 (6): 885–908.

Belzung, C., and J. R. Anderson. 1986. "Social Rank and Responses to Feeding Competition in Rhesus-Monkeys." *Behavioural Processes* 12 (4): 307–16. q14
Benjamini, Y., D. Drai, G. Elmer, N. Kafkafi, and I. Golani. 2001. "Controlling the False Discovery Rate in Behavior Genetics Research." *Behavioural Brain Research* 125 (1–2): 279–84.
Benjamini, Y., and D. Yekutieli. 2005. "Quantitative Trait Loci Analysis Using the False Discovery Rate." *Genetics* 171 (2): 783–90.
Blanchard, Robert J., Kevin Hori, D. Caroline Blanchard, and James Hull. 1987. "Ethanol Effects on Aggression of Rats Selected for Different Levels of Aggressiveness." *Pharmacology Biochemistry and Behavior* 27 (4): 641–44.
Boardman, J. D., J. Daw, and J. Freese. 2013. "Defining the Environment in Gene-Environment Research: Lessons from Social Epidemiology." *American Journal of Public Health* 103:64–72.
Borsari, B., and K. B. Carey. 2001. "Peer Influences on College Drinking: A Review of the Research." *Journal of Substance Abuse* 13 (4): 391–424.
Boyce, W., and B. Ellis. 2005. "Biological Sensitivity to Context." Pt. 1, "An Evolutionary-Developmental Theory of the Origins and Functions of Stress Reactivity." *Development and Psychopathology* 17:271–301.
Brain, P. F., N. W. Nowell, and A. Wouters. 1971. "Some Relationships between Adrenal Function and the Effectiveness of a Period of Isolation in Inducing Intermale Aggression in Albino Mice." *Physiology and Behavior* 6:27–29.
Bulchand, S., L. Subramanlan, and S. Tole. 2003. "Dynamic Spatiotemporal Expression of LIM Genes and Cofactors in the Embryonic and Postnatal Cerebral Cortex." *Developmental Dynamics* 226 (3): 460–69.
Candland, D. K., and A. I. Leshner. 1974. "A Model of Agonistic Behavior: Endocrine and Autonomic Correlates." Pp. 137–63 in *Limbic and Autonomic Nervous Systems Research*, edited by L. V. DiCara. New York: Plenum.
Caspi, A., J. McClay, T. E. Moffitt, J. Mill, J. Martin, I. W. Craig, A. Taylor, and R. Poulton. 2002. "Role of Genotype in the Cycle of Violence in Maltreated Children." *Science* 297 (5582): 851–54.
Clayton, David G. 1978. "Model for Association in Bivariate Life Tables and Its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence." *Biometrika* 65 (1): 141–51.
———. 2008. "Testing for Association on the X Chromosome." *Biostatistics* 9 (4): 593–600.
Conley, Dalton, and Emily Rauscher. 2013. "Genetic Interactions with Prenatal Social Environment: Effects on Academic and Behavioral Outcomes." *Journal of Health and Social Behavior* 54 (1): 109–27.
Crabbe, J. C., T. J. Phillips, K. J. Buck, C. L. Cunningham, and J. K. Belknap. 1999. "Identifying Genes for Alcohol and Drug Sensitivity: Recent Progress and Future Directions." *Trends in Neurosciences* 22 (4): 173–79.
Crabbe, J. C., T. J. Phillips, R. A. Harris, M. A. Arends, and G. F. Koob. 2006. "Alcohol-Related Genes: Contributions from Studies with Genetically Engineered Mice." *Addiction Biology* 11 (3–4): 195–269.
Dahlberg, M., and E. Johansson. 2002. "On the Vote-Purchasing Behavior of Incumbent Governments." *American Political Science Review* 96 (1): 27–40.
De Bellis, Michael D., Duncan B. Clark, Sue R. Beers, Paul H. Soloff, Amy M. Boring, Julie Hall, Adam Kersh, and Matcheri S. Keshavan. 2000. "Hippocampal Volume in Adolescent-Onset Alcohol Use Disorders." *American Journal of Psychiatry* 157 (5): 737–44.
Dixit, A., and J. Londregan. 1996. "The Determinants of Success of Special Interests in Redistributive Politics." *Journal of Politics* 58 (4): 1132–55.

Duncan, G. J., J. Boisjoly, M. Kremer, D. M. Levy, and J. Eccles. 2005. "Peer Effects in Drug Use and Sex among College Students." *Journal of Abnormal Child Psychology* 33 (3): 375–85.

Eastman, Peggy. 2002. "The Student Perspective on College Drinking." National Institute of Alcohol Abuse and Alcoholism. http://www.collegedrinkingprevention.gov/supportingresearch/student1.aspx.

Ellis, B. J., W. T. Boyce, J. Belsky, M. J. Bakermans-Kranenburg, and M. H. van Ijzendoorn. 2011. "Differential Susceptibility to the Environment: An Evolutionary-Neurodevelopmental Theory." *Development and Psychopathology* 23 (1): 7–28.

Enoch, M., P. Shen, K. Xu, C. Hodgkinson, and D. Goldman. 2006. "Using Ancestry-Informative Markers to Define Populations and Detect Population Stratification." *Journal of Psychopharmacology* 20:19–26.

Fletcher, Jason M., and Dalton Conley. 2013. "The Challenge of Causal Inference in Gene-Environment Interaction Research: Leveraging Research Designs from the Social Sciences." *American Journal of Public Health* 103:S42–S45.

Foster, Gigi. 2006. "It's Not Your Peers, and It's Not Your Friends: Some Progress toward Understanding the Educational Peer Effect Mechanism." *Journal of Public Economics* 90 (8–9): 1455–75.

Goldman, D., and A. Bergen. 1998. "General and Specific Inheritance of Substance Abuse and Alcoholism." *Archives of General Psychiatry* 55 (11): 964–65.

Gottfredson, Mark, and Travis Hirschi. 1990. *A General Theory of Crime*. Stanford, Calif.: Stanford University Press.

Guo, G. 1993. "Use of Sibling Data to Estimate Family Mortality Effects in Guatemala." *Demography* 30 (1): 15–32.

Harris, K. M., F. Florey, J. Tabor, P. S. Bearman, J. Jones, and J. R. Udry. 2003. "The National Longitudinal Study of Adolescent Health: Research Design." Add Health, University of North Carolina at Chapel Hill. http://www.cpc.unc.edu/projects/addhealth/design.

Jaffee, S. R., and T. S. Price. 2007. "Gene-Environment Correlations: A Review of the Evidence and Implications for Prevention of Mental Illness." *Molecular Psychiatry* 12 (5): 432–42.

Johnson, P. B. 1989. "Reactions, Expectancies, and College Student's Drinking." *Psychology Report* 65:1245–46.

Kadrmas, J. L., and M. C. Beckerle. 2004. "The LIM Domain: From the Cytoskeleton to the Nucleus." *Nature Reviews Molecular Cell Biology* 5 (11): 920–31.

Kandel, Denise B. 1978. "Homophily, Selection, and Socialization in Adolescent Friendships." *American Journal of Sociology* 84 (2): 427–36.

Kantor, Glenda Kaufman, and Murray A. Straus. 1990. "The 'Drunken Bum' Theory of Wife Beating." Pp. 203–24 in *Physical Violence in American Families: Risk Factors and Adaptations to Violence in 8,145 Families*, edited by Murray Arnold Straus and Richard J. Gelles. New Brunswick, N.J.: Transaction.

Konig, I. R., C. Loley, J. Erdmann, and A. Ziegler. 2014. "How to Include Chromosome X in Your Genome-Wide Association Study." *Genetic Epidemiology* 38 (2): 97–103.

Kremer, Michael, and Dan Levy. 2008. "Peer Effects and Alcohol Use among College Students." *Journal of Economic Perspectives* 22 (3): 189–206.

Lasek, A. W., F. Giorgetti, K. H. Berger, S. Tayor, and U. Heberlein. 2011. "LMO Genes Regulate Behavioral Responses to Ethanol in Drosophila melanogaster and the Mouse." *Alcoholism-Clinical and Experimental Research* 35 (9): 1600–1606.

Leonard, K. E., R. L. Collins, and B. M. Quigley. 2003. "Alcohol Consumption and the Occurrence and Severity of Aggression: An Event-Based Analysis of Male to Male Barroom Violence." *Aggressive Behavior* 29 (4): 346–65.

Leonard, K. E., B. M. Quigley, and R. L. Collins. 2003. "Drinking, Personality, and Bar Environmental Characteristics as Predictors of Involvement in Barroom Aggression." *Addictive Behaviors* 28 (9): 1681–1700.
Leshner, A. I., and J. A. Moyer. 1975. "Androgens and Agonistic Behavior in Mice: Relevance to Aggression and Irrelevance to Avoidance-of-Attack." *Physiology and Behavior* 15:695–99.
Leshner, A. I., J. A. Moyer, and W. A. Walker. 1975. "Pituitary Adrenocortical Activity and Avoidance-of-Attack in Mice." *Physiology and Behavior* 15:689–93.
Leshner, A. I., W. A. Walker, A. E. Johnson, J. S. Kelling, S. J. Kreisler, and B. B. Svare. 1973. "Pituitary Adrenocortical Activity and Intermale Aggressiveness in Isolated Mice." *Physiology and Behavior* 11 (5): 705–11.
Li, Yun, Cristen Willer, Serena Sanna, and Goncalo Abecasis. 2009. "Genotype Imputation." *Annual Review of Genomics and Human Genetics* 10:387–406.
Liang, K. Y., S. L. Zeger, and B. Qaqish. 1992. "Multivariate Regression Analyses for Categorical Data." *Journal of the Royal Statistical Society* B 54 (1): 3–40.
Lindbeck, A., and J. W. Weibull. 1993. "A Model of Political Equilibrium in a Representative Democracy." *Journal of Public Economics* 51 (2): 195–209.
Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60 (3): 531–42.
Marchini, Jonathan, and Bryan Howie. 2010. "Genotype Imputation for Genome-Wide Association Studies." *Nature Reviews Genetics* 11:499–511.
Maxson, Stephen. 2009. "The Genetics of Offensive Aggression in Mice." Pp. 301–16 in *Handbook of Behavior Genetics*, edited by K. Y. Kim. New York: Springer.
Maxson, Stephen C., and Andrew Canastar. 2003. "Conceptual and Methodological Issues in the Genetics of Mouse Agonistic Behavior." *Hormones and Behavior* 44 (3): 258–62.
Mitchell, C., S. McLanahan, J. Brooks-Gunn, I. Garfinkel, J. Hobcraft, and D. Notterman. 2013. "Genetic Differential Sensitivity to Social Environments: Implications for Research." *American Journal of Public Health* 103:102–10.
Mitchell, C., D. Notterman, J. Brooks-Gunn, J. Hobcraft, I. Garfinkel, K. Jaeger, I. Kotenko, and S. McLanahan. 2011. "Role of Mother's Genes and Environment in Postpartum Depression." *Proceedings of the National Academy of Sciences of the USA* 108 (20): 8189–93.
Moffitt, Robert A., ed. 2001. *Policy Interventions, Low-Level Equilibria and Social Interactions*. Cambridge, Mass.: MIT Press.
Mouw, Ted. 2006. "Estimating the Causal Effect of Social Capital: A Review of Recent Research." *Annual Review of Sociology* 32:79–102.
Ober, C., D. A. Loisel, and Y. Gilad. 2008. "Sex-Specific Genetic Architecture of Human Disease." *Nature Reviews Genetics* 9 (12): 911–22.
Palmer, R. S., T. J. McMahon, B. J. Rounsaville, and S. A. Ball. 2010. "Coercive Sexual Experiences, Protective Behavioral Strategies, Alcohol Expectancies and Consumption among Male and Female College Students." *Journal of Interpersonal Violence* 25 (9): 1563–78.
Panel on Contexts and Consequences. 2002. "High-Risk Drinking in College: What We Know and What We Need to Learn." Final report of the Panel on Contexts and Consequences. Task Force of the National Advisory Council on Alcohol Abuse and Alcoholism, National Institutes of Health and U.S. Department of Health and Human Services.
Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155 (2): 945–59.
Raine, A. 2002. "Biosocial Studies of Antisocial and Violent Behavior in Children and Adults: A Review." *Journal of Abnormal Child Psychology* 30 (4): 311–26.

Remedios, R., L. Subramanian, and S. Tole. 2004. "LIM Genes Parcellate the Embryonic Amygdala and Regulate Its Development." *Journal of Neuroscience* 24 (31): 6986–90.

Ripke, S., B. M. Neale, A. Corvin, J. T. R. Walters, et al. 2014. "Biological Insights from 108 Schizophrenia-Associated Genetic Loci." *Nature* 511 (7510): 421–27.

Roudsari, B. S., M. M. Leahy, and S. T. Walters. 2009. "Correlates of Dating Violence among Male and Female Heavy-Drinking College Students." *Journal of Interpersonal Violence* 24 (11): 1892–1905.

Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics* 116 (2): 681–704.

Schilling, E. A., R. H. Aseltine, J. L. Glanovsky, A. James, and D. Jacobs. 2009. "Adolescent Alcohol Use, Suicidal Ideation, and Suicide Attempts." *Journal of Adolescent Health* 44 (4): 335–41.

Searle, Shayle R., George Casella, and Charles McCulloch. 1992. *Variance Components*. New York: Wiley.

Simons, Ronald L., Man Kit Lei, Steven R. H. Beach, Gene H. Brody, Robert A. Philibert, and Frederick X. Gibbons. 2011. "Social Environment, Genes, and Aggression: Evidence Supporting the Differential Susceptibility Perspective." *American Sociological Review* 76 (6): 883–912.

Singleton, R. A., and A. R. Wolfson. 2009. "Alcohol Consumption, Sleep, and Academic Performance among College Students." *Journal of Studies on Alcohol and Drugs* 70 (3): 355–63.

Sutherland, Edwin H. 1947. *Principles of Criminology*, 4th ed. Philadelphia: Lippincott.

Tapert, Susan F., Lisa Caldwell, and Christina Burke. 2005. "Alcohol and the Adolescent Brain." *Alcohol Research and Health* 28 (4): 205–12.

Thombs, D. L. 1999. *An Introduction to Addictive Behaviors*. New York: Guilford.

Tsai, L. T. Y., R. J. Bainton, J. Blau, and U. Heberlein. 2004. "LMO Mutants Reveal a Novel Role for Circadian Pacemaker Neurons in Cocaine-Induced Behaviors." *Plos Biology* 2 (12): 2122–34.

Vaupel, J. W., K. G. Manton, and E. Stallard. 1979. "Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." *Demography* 16 (3): 439–54.

Wagner, B., J. Li, H. X. Liu, and G. Guo. 2013. "Gene-Environment Correlation: Difficulties and a Natural Experiment-Based Strategy." *American Journal of Public Health* 103:167–73.

Wechsler, Henry, Andrea Davenport, George Dowdall, Barbara Moeykens, and Sonia Castillo. 1994. "Health and Behavioral Consequences of Binge Drinking in College: A National Survey of Students at 140 Campuses." *Journal of the American Medical Association* 272 (21): 1672–77.

Zimmerman, David J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics* 85 (1): 9–23.