Network Effects in Risky Teen Behavior: Evidence from the L.A.FANS

Alvaro Mezza and Emily Wiemers

March 31, 2008

Abstract

This paper uses a dataset of families in Los Angeles called the Los Angeles Family and Neighborhood Survey (L.A.FANS) to examine whether there is evidence of network effects in risky teen behaviors such as smoking, drinking, drug use, skipping school and teen sexual activity. We use a dataset rich in information about attitudes and behaviors of both parents and children and combine it with census data about neighborhoods in Los Angeles and data on behavior from the National Longitudinal Study of Adolescent Health (Add Health), allowing us to disentangle the issues of causality in the measurement of peer effects. We control for factors that are not found in other studies of network and peer effects and that are likely to be correlated with the measures of interest. We examine whether networks are operating via a channel of social norms using information about adults' attitudes toward drugs, alcohol, and cigarettes consumption by adults. Beliefs do not seem to be significant and their inclusion does affect the size or significance of the network measure. Once we control for the endogeneity of the network measure using an IV approach we no longer find evidence of network effects, suggesting that omitted variables may be biasing our network measure.

1 Motivation

It is well documented that many youth behaviors are strongly correlated among peers and within neighborhoods and ethnic groups. In a study of disadvantaged youth in Boston, Case and Katz (1991) show strong correlations among neighbors in drug use, gang membership, alcohol use, church attendance, and crime, even after controlling for individual and family characteristics. The positive correlations among neighbors prompt the question of causality.

Are behaviors among neighbors, peers and ethnic groups correlated because of peer effects where the behavior of an individual is affected by the behavior or attitudes of his peers? Or, are behaviors among neighbors, peers and ethnic groups correlated because there are unobserved individual, group, or environmental factors that cause similar behavior across individuals? If correlations among peers are the result of causal peer influences we should see multiplier effects in behavior because of feedback between individuals and their peer group. Multiplier effects can operate as both a positive and a negative influence. Finding peer effects suggests that combatting risky behaviors among a few individuals in a network can have large effects.

This paper uses a dataset of families in Los Angeles called the Los Angeles Family and Neighborhood Survey (L.A.FANS) to examine whether there is evidence of network effects in risky teen behaviors such as smoking, drinking, drug use, skipping school and teen sexual activity. Our findings suggest that when we use an endogenous measure of peer effects we find strong evidence of peer effects that are robust to the inclusion of many individual and neighborhood characteristics. However, once we control for the endogeneity of peer behavior we do not find evidence of peer effects. We use a dataset rich in information about attitudes and behaviors of both parents and children and combine it with census data and information coming from the sample core of the public-use data of the National Longitudinal Study of Adolescent Health (Add Health) to identify the network effect.

Our work builds on the work of Bertrand, Luttmer and Mullainathan (2000) and Aizer and Currie (2004), who use non-experimental variation to identify peer effects in welfare and in public prenatal care, respectively. They define networks using language groups and ethnic groups. Bertrand, Luttmer and Mullainathan (2000) use language spoken at home as a non-experimental source of exogenous variation to identify peer effects in welfare use. Their approach circumvents the omitted variable biases that are usually present in measures of network effects without experimental variation because using language and geography to proxy for social networks generates variation within local areas and within language groups so that fixed effects for both neighborhood and language groups can be included. Aizer and Currie (2004) extend the analysis to examine the importance of social networks in the use of public prenatal care. They use the same identification strategy but extend the question to explore whether the networks that they find operate through information sharing among network members or through social norms within the network. Our paper also builds on the work of Musick, Seltzer and Schwartz (2007), who use the L.A.FANS to examine the role of neighborhood norms in substance use among teens.

Our study uses the same empirical approach as Bertrand et al (2000) and Aizer and Currie (2004) to examine the differential impact of a stronger network across racial groups and

asks whether having a stronger network increases risky teen behaviors more for adolescents from racial groups with a higher likelihood of risky behavior. Risky teen behaviors such as smoking, drinking and drug use vary by race, with white teens more likely to participate in these activities than their African American and Latino counterparts. Los Angeles is an ideal environment for looking at these questions because it is a city with substantial white, African American, Latino and Asian populations. The population composition at tract level in Los Angeles has changed substantially over time with the influx of Latino immigrants. We use this variation to account for the endogeneity of neighborhood selection. The L.A.FANS has several advantages in addition to its heterogeneity by race and ethnicity. First, it is a survey of households clustered in neighborhoods. It samples 3090 households from 65 tracts of Los Angeles County. The structure of the survey means that we have many children within each neighborhood. Second, the survey includes separate interviews for a randomly selected adult in each household, a randomly selected child in households with children, as well as a randomly selected sibling (if siblings are present) and the primary care giver of the randomly selected child. The survey structure gives us self-reported information on both adults and children within a household. Third, the survey includes information on the neighborhood of residence at the tract and block level. This information provides us with a richer set of controls for characteristics of the block in which the child lives.

2 Background

The underlying question in studying peer effects is whether the behavior of an individual is affected by the behavior and/or attitudes of his or her peers. The empirical question of whether peer or neighborhood effects exist is motivated by the work of William Julius Wilson and others who suggest that the high rates of poverty, joblessness and crime experienced in the inner city is supported by peer influences and the behavior of adult role models (Wilson (1987), Crane (1991)). Early work on peer influences such as Case and Katz (1991) found that outcomes of disadvantaged youth are related to measured family background characteristics and that the spatial patterns of these outcomes are consistent with neighborhood effects operating through peer influences. Though the early work shows some evidence of neighborhood and peer effects, the overall results of the early studies are quite mixed as to whether neighborhood or peer effects exist and the results seem to depend largely on the number and quality of neighborhood and individuals covariates, suggesting omitted variable and selection bias.

Manski (1993) describes the problem with identifying peer effects when only the distribution of behaviors in the population is observed. He shows the difficulty in distinguishing

between three types of effects. The first are contextual effects, in which the characteristics of an individual's group or environment affects his outcomes but there is not feedback between the individual's outcomes and the characteristics of the group or environment on which the outcomes depend. The second are endogenous effects, in which the individual's outcomes feed back into the group and neighborhood characteristics on which the individual's outcomes depend, producing social multiplier effects. The third are correlated effects or spurious effects, in which unobservables or similar shocks in environment are driving the observed correlation among group members. Manski concludes that identifying the effect of interest (the endogenous effect) is difficult in the presence of the other two effects. Brock and Durlauf (2001, 2007) show that these effects can be distinguished using non-linearities in discrete choice models. Bertrand, Luttmer and Mullainathan (2000) and Aizer and Currie (2004) use variation in behaviors across ethnic and racial groups to identify peer effects. They do not separately identify contextual and endogenous effects but do separate correlated effects using a fixed effect strategy.

Several experimental approaches have been used to identify neighborhood or peer effects. Katz, Kling and Liebman (2001) use the Moving to Opportunity experiment in which 4600 families with children living in high poverty public housing projects in New York, Boston, Baltimore, Chicago and Los Angeles were randomized into three groups: a control group in which families continued to be eligible to live in public housing and two treatment groups. In the first treatment group families received geographically unrestricted Section 8 vouchers that could be used to rent apartments in any neighborhood. In the second treatment group families received restricted vouchers that could be used to rent apartments only in low-poverty neighborhoods. Katz, Kling and Liebman use the vouchers as an instrument for neighborhood quality in a regression of youth and adult outcomes on neighborhood and family characteristics. They measure the intent to treat effect and find that there are significantly positive neighborhood effects on outcomes for girls but significantly negative neighborhood effects on outcomes for boys. In a more recent paper Katz, Kling and Liebman (2004) compare their estimates to non-experimental estimates using the L.A.FANS data and show that non-experimental results are not consistent with their experimental results. They also try to identify whether the neighborhood effect is linear and find weak evidence of linearity. Other experiments include Sacerdote's (2001) use of random assignments of college roommates, in which he finds little evidence of peer effects in academic outcomes in the long run but stronger evidence of peer effects in social outcomes.

The evidence on peer effects in teen behavior is inconclusive. While some studies find strong evidence of peer effects in behaviors such as smoking, drinking, drug use, church attendance and dropping out (Gaviria and Raphael (2001), Powell, Tauras and Ross (2003)),

others find more mixed evidence (Clark and Loheac (2007)). The strength of the results seems to depend on how well correlated effects and the endogeneity of peer groups are accounted for in estimation. Gaviria and Raphael (2001) use data from the National Education Longitudinal Study to estimate peer effects in smoking, drinking, drug use, church attendance and dropping out. They argue that contextual peer effects are less important because they are using peers from school. Average peer characteristics, therefore, are valid as instruments for average peer behavior. They find similar results with their OLS and IV strategy. They control for the endogeneity of school choice by estimating peer effects on recent movers and comparing the results with those of people who did not move recently. They find some differences in the estimates although they argue that the differences are not large. They also find significant positive peer effects using an OLS specification after controlling for school level characteristics. This study suggests that peer effects are large, but the instruments they use are only valid in the absence of contextual effects, and the controls for the endogeneity of school choice and for school level correlated effects are not robust. Powell, Tauras and Ross (2005) examine the impact of prices and tobacco control policies on youth smoking while distinguishing between direct effects of the policy and indirect effects through peer effects. They use a school based measure of peer behavior and find that peer effects are the main channel through which tobacco control policies operate but that failing to account for prices and tobacco control policies biases the magnitude of peer effects upwards. Like Gaviria and Raphael (2001), they assume that contextual effects are not important and use average peer characteristics as an instrument for peer behavior. They also use Census measures such as racial diversity, population density and poverty as instruments to account for endogenous sorting of households. As in Gaviria and Raphael (2001) these instruments are arguably endogenous and so the measured peer effects are biased. Clark and Loheac (2007) ask a similar question to Gaviria and Raphael (2001), and estimate the extent of peer effects in smoking, drinking, and drug use in the Add Health survey. Their study improves on Gaviria and Raphael by including school fixed effects to account for correlated effects in schools and for unobserved characteristics that are shared by individuals in the same school. They show that estimated peer effects are reduced substantially by including school fixed effects in their model. After controlling for school fixed effects they find significant peer effects only for boys in alcohol use and smoking. Peer effects for girls and for drug use among boys become insignificant. However, one potential problem remains. The authors include the average behavior of the reference group lagged by one year to eliminate the omitted variable bias that arises if there are unobserved personal characteristics that are correlated within students of the same school. A one year lag may not be distant enough in time for the lag measure to be exogenous considering the relatively high correlations reported by the authors for the

current and lagged behaviors that range from 0.4 to 0.55. Like Clark and Loheac (2007) our paper includes neighborhood fixed effects, controlling for unobserved characteristics shared by individuals in the same neighborhood and for the effect of the neighborhood itself and utilizes an IV approach to avoid omitted variable biases.

3 Data

The analysis in this article is based on the first wave of the L.A.FANS. This survey has been designed for the purpose of studying neighborhood effects on child and family outcomes, making it a suitable dataset for this analysis. The L.A.FANS is a longitudinal study. The baseline wave was completed between the years 2000 and 2001. The fieldwork for the second wave is under process. The sampling units are census tracts from Los Angeles County. From the approximately 1,600 eligible tracts in the county, 65 tracts were selected in which 3,090 households were interviewed. Of these households, 30% are in very poor stratum tracts, 31% in poor stratum tracts, and 39% in non-poor stratum tracts.[1]

At the household level, a randomly sampled adult was interviewed. In those households with children, a randomly selected child and his/her mother or primary caregiver were interviewed. Finally, a sibling was also randomly selected if the randomly selected child had one or more siblings less than 18 years of age who shared the same biological or adoptive mother and the same primary caregiver. Contextual information at block level was also collected by interviewers. This includes information such as evidence of drug use and crime, whether kids are seen hanging out together, homelessness and the presence of graffiti.

Our analysis considers if networks are relevant in determining individual behavior in smoking, drinking, drug use, skipping school and having sex for teenagers between 12 and 17 years old. We define the network of an individual as teenagers from an individual's race who live in the same neighborhood. We ask whether having a strong network increases risky teen behaviors more for individuals from groups with higher participation in risky teen activities. The L.A.FANS gives us some evidence about whether neighborhoods are a relevant definition of peer groups for kids. About 55 percent of 12 to 17 year-olds report that their best friend lives in their neighborhood. Over 90 percent report that they know some or most of the adults and children in their neighborhood. Over fifty percent of teens define their neighborhood as their block with over 90 percent defining their neighborhood as within a fifteen minute walk of their house. Over 80 percent of kids live within five miles of their school. While we may be underestimating the size of the network of an individual,

[1] For a more detailed information about the sampling methods and the dataset in general, check Sastry et al. (2006).

it may include friends of a different race in the same neighborhood or friends from different neighborhoods, the questions on neighborhood size in the L.A.FANS suggest that a Census tract is a good proxy for the neighborhood of the teens in the survey.

We categorize race and ethnicity in the L.A.FANS by first using self-reported race. Children are asked separately whether they are white, black, Latino, Asian, Pacific Islander and Native American. Those children who choose only one race are categorized by this choice. Children who report more than one race are asked which race best describes them. We use the report of the race that describes them best as the race for children who report more than one race. Some children only report their race in the household roster because of a mistake in the sibling questionnaire. When using the household roster we assign race to children who report only one race because they are not asked which race best describes them if they select more than one race. After assigning race in this way, only six children are not assigned a race. Table 1 shows the mean of the risky behaviors by race for smoking, drinking, taking drugs, skipping school and having sex among the 833 teens in our sample. In each case the mean of each race and ethnic group is compared to the group with the highest mean behavior. For each outcome there are significant differences in behavior by race. Latinos, African Americans and Asians are significantly less likely than whites to smoke, drink or take drugs. Whites, Latinos and Asians are less likely than blacks to have sex. Differences in skipping school are only significant for Latinos and Asians, who are less likely to skip school than their black counterparts.

Table 1. Risky Behaviors by Race of Child

	Ever Smoked	Ever Drink	Ever take Drugs	Ever Skip School	Ever have Sex
White	33%	49%	24%	28%	13%**
Latino	25%**	38%***	14%***	23%*	15%**
Black	24%*	31%***	18%	32%	23%
Asian/Pac. Islander	19%**	12%***	13%**	17%**	5%***
* significant at 10% ** significant at 5% ***significant at 1%					

The variation that we use to identify peer effects is differences in behavior by race/ethnicity and Census tract. In order to identify a peer effect we require that at least some individuals live in mixed race tracts. The L.A.FANS categorizes tracts by their ethnic group composition. Table 2 shows the race of the children in our sample by the ethnic group composition of the tract in which they reside. The racial/ethnic group characterization of the tract is measured by a cluster analysis of 2000 Census tracts in five race/ethnic groups (non-Hispanic white, non-Hispanic black, Latino, Asian/Pacific Islander and other). Table 2 shows that

while over 80 percent of Latinos, 78 percent of whites, 37 percent of blacks and 46 percent of Asian/Pacific Islanders live in neighborhoods categorized as within their race/ethnic group, many groups live together. In particular we see that there is substantial overlap between blacks and Latinos and between Asians and both whites and Latinos.

Table 2. Race of Child by Race/Ethnic Group of Neighborhood

	High Asian	White	Latino and Black	Latino	White and Other	Total
Latino	6.41%	1.92%	11.97%	68.59%	11.11%	100%
White	11.37%	45.02%	0.95%	9.48%	33.18%	100%
Black	11.11%	3.33%	37.78%	20.00%	27.78%	100%
Asian/Pac. Islander	46.03%	11.11%	3.17%	14.29%	25.40%	100%
Total	11.18%	13.70%	11.30%	44.23%	19.59%	100%

The L.A.FANS allows us to include a rich set of controls for neighborhood characteristics at block level reported by interviewers. We group some of the relevant observations into seven broad categories: presence of drinking, presence of drugs, presences of teenage groups, presence of gangs, presence of police officers, presence of a church on the block and presence of a bar on the block. The variables on drinking, drugs and gangs are combinations of several variables. We say that drinking is present on a block if the interviewer sees people drinking, sees empty bottles on the street, sees drunk people, or hears from neighbors about drunk loiterers. We say that drugs are present if the interviewer sees drug paraphernalia on the street, sees people selling drugs, or hears from neighbors about drugs being sold. We say that gangs are present if the interviewer sees gangs, or if they hear that gangs are present from neighbors. Table 3 shows the mean of these variables by the strata of the neighborhood. Very poor neighborhoods are more likely to have evidence of drug use, public drinking, bars, police, teen groups and gangs. They are also more likely to have churches. Tracts that are not poor are very unlikely to have gangs or drugs present.

Table 3. Block Characteristics by Tract Strata

Block Level Characteristic	Very Poor	Poor	Not Poor
Church	.60	.37	.21
Drugs	.47	.44	.05
Drinking	.99	.85	.61
Bar	.40	.20	.13
Police	.62	.54	.28
Teen Groups	.71	.41	.37
Gangs	.18	.12	.03

4 Estimation Strategy

We quantify the network effect in risky teen behaviors in the L.A.FANS. Our study extends on the papers of Bertrand, Luttmer and Mullainathan (2000), Aizer and Currie (2004) and Musick, Seltzer and Schwartz (2007). Following Bertrand et al (2000), the true equation to estimate is the following:

$$\Pr(outcome_{ijk}) = Network_{ijk}\alpha + X_i^*\beta + Y_j^*\gamma + Z_k^*\delta + \epsilon_{ijk}$$

where i indexes the individual, j indexes the area and k indexes the ethnic group. This equation says that the probability of an outcome (smoking, alcohol or drug use, having sex or skipping school) for individual i of ethnic group k in area j is a function of his network $Network_{ijk}$, observed and unobserved personal characteristics X_i^*, observed and unobserved area characteristics Y_j^*, observed and unobserved ethnic group characteristics Z_k^* and an error term ϵ_{ijk}. The problem comes in estimation because we cannot measure all of these factors precisely. We do not know exactly who is in the network of an individual and we do not measure all personal, ethnic group and neighborhood characteristics likely associated with the outcome. The standard way to estimate network effects is to regress the probability of an outcome on personal characteristics and on the average probability of this outcome among an individual's neighbors. Estimating this equation introduces measurement error and omitted variable bias. As in Bertrand et al (2000), we circumvent these problems by using geographic and ethnic group variation so we can include both area and ethnic group fixed effects to remove the omitted variable bias generated by the correlation between omitted neighborhood and ethnic group characteristics and the impact of the network. We measure the network of individual i in area j in ethnic group k as a composite of the strength of the network and the "quality" of the network. The strength of the network is given by the proportion of people who belong to the individual's ethnic group that reside in the same neighborhood, where neighbors are measured at the census tract level; while the quality of the network measures the likelihood that an individual in the network will experience the outcome, in our case the mean probability of each outcome among those in an individual's ethnic group that reside in the same neighborhood. Adding the network strength by itself to the estimating equation, although its coefficient is inconsistent, avoids the bias in α generated by the correlation between network strength and omitted personal characteristics. We estimate the following equation:

$$\Pr(outcome_{ijk}) = (Netstrength_{jk} * Netqual_{jk})\alpha + X_i\beta + \gamma_j + \delta_k + Netstrength_{jk}\theta + \epsilon_{ijk} \quad (1)$$

Though it is likely that an individual's network includes people of different races and

ethnicities, there is evidence pointing to the salience of race and ethnicity in friendship formation. For example, Kao and Joyner (2004), using the Add Health survey, find that 92 percent of white adolescents and 85 percent of black adolescents report that their best friend is of their same ethnicity. This still holds when the list of friends is extended to their first five listed friends, with percentages of 90 and 81, respectively. Though the figures for Hispanics and Asians are not as high (51 and 50 percent for best friend and 42 and 41 percent for the first five listed friends), they are still high enough to consider our measure a good proxy for the peer group.[2]

Given our measure of network, we will be asking if being surrounded by one's racial group increases smoking, drug and alcohol use, having sex and skipping school more for individuals from racial groups where these risky behaviors are higher. Although this measure is not as straight-forward to interpret as the average behavior at the neighborhood or school level (as it has usually been measured in the literature), it allows us to control for fixed effects for neighborhood and race. In this way we avoid many of the omitted variable biases associated with omitted personal characteristics, unobserved neighborhood characteristics, and neighborhood selection. However, there are still three potential sources of bias in this equation. The first arises because it is expected that adolescents from the same ethnic group who live in the same area have some similar personal unobservable characteristics that are correlated with the quality measure $Netqual_{jk}$. If this is the case, the error term will be correlated with the network effect, biasing the result of interest. The second source of bias is given by simultaneity. If an adolescent is influenced by his/her peer group, this adolescent will also be influential on the behavior of his/her group, something not corrected here. The last source of bias comes from correlation between $(Netstrength_{jk} * Netqual_{jk})$ and omitted personal characteristics. As Bertrand et al (2000) point out, this type of correlation arises if individuals differentially self-select away from their ethnic group, i.e. if white people, for example, move away from white people for different reasons than black people move away from black people.

To circumvent these problems, we adopt an instrumental variable (IV) approach. The exogenous instrument for our measure of network is given by the strength of the network calculated with data coming from the 1980 Census multiplied by the quality of the network calculated using the sample core of the public-use dataset for the first wave of the Add Health survey. That is, we replace in our original measure of network the proportion of adolescents of a certain race who live in a certain neighborhood by the proportion of the population of a certain race who live in a certain neighborhood in 1980, and the proportion of adolescents

[2]For more literature that supports the relevance of race and ethnicity in network formation, see the review study by McPherson et al (2001)

of a certain ethnic group who live in a certain neighborhood of Los Angeles who experience the outcome by the proportion of adolescents of a certain ethnic group at the national level who experience the outcome.

This instrument is exogenous because of the changes in population between the 1980 and the 2000 Census in many neighborhoods. These changes were driven largely by immigration. Many areas, including areas both in central Los Angeles and in the suburbs, underwent a transition in which Latinos replace both blacks and whites as the majority ethnic group. In South Central Los Angeles, Latinos have replaced blacks as the majority ethnic group (Myers (2002)). These changes can be clearly seen in Figures 1 and 2 (see Figure Appendix), which show that the distribution of the population by majority ethnicity in 1980 and in 2000. The quality measure used in our instrument is exogenous, given the different group of adolescents used to calculate the proportions (based on neighborhoods of Los Angeles County vs. the whole nation) at different points in time (2000-2001 vs 1994-1995). Though the measure is exogenous, since the quality of the network is calculated at the national level, the instrument presents less variability than the endogenous measure of network. This could be a problem because of the small sample size and the high number of neighborhood fixed effects, and could affect the statistical significance of the network.

[Figures 1 and 2 here]

We try to see if the networks are operating via a channel of social norms or via a channel of information sharing. For example, do teens smoke because smoking is cool in their neighborhood or because someone told them where to buy cigarettes? We do this using a set of questions about adult beliefs about smoking, drug use, and drinking. These are the same measures used in Musick, Seltzer and Schwartz (2007). We estimate the following equation:

$$\begin{aligned}\Pr(outcome_{ijk}) = (Netstrength_{jk} * Netqual_k)\alpha + (Netstrength_{jk} * Netbeliefs_k)\varphi \\ + X_i\beta + \gamma_j + \delta_k + Netstrength_{jk}\theta + \epsilon_{ijk}\end{aligned} \quad (2)$$

where $Netbeliefs_k$ measures the average adult beliefs about a particular outcome across ethnic groups. The inclusion of a measurement of beliefs allows us to see if social norms are playing a large role in the measured network effect or whether a combination of social norms and information sharing is driving the network effect.

5 Results

In this section we present three sets of results. First, we present results from our estimation of Equation 1 using endogenous measures of networks but controlling for tract and race fixed

effects. We explore how changes in the quality of the network, as measured by the percent of people in your race and neighborhood who participate in the activity, affects the probability that an individual will smoke, drink, do drugs, have sex or skip school. Second, we show results from estimating Equation 2 in which we see whether there is still evidence of network effects after controlling for neighborhood norms. Finally, we present our IV results in which the network measure is exogenous.

Table 4 (see Table Appendix) shows the summary statistics for the variables that we use in the regressions. Our sample has a large number of Latinos and blacks; about 26 percent of our sample is foreign born. Over 20 percent of the teenagers in our sample have a job. Some of those who are working are under 16 because this measure also includes jobs like babysitting and other jobs done "under the table." About 41 percent of our sample lives in a single parent household. Over 80 percent of the sample goes to public school. The mean income of the sample is $60,000 but the median income is $35,000. Family income includes asset, transfer and labor income of all family members but excludes the income of the child. The average number of kids in the family is 2.27.

[Table 4 here]

We first estimate Equation 1 using an endogenous measure of peer effects. We estimate this equation for each of our five outcomes (smoking, drinking, drug use, having sex, and skipping school). Table 5 (see Table Appendix) shows the results of this estimation. The peer effect has a positive and significant effect for each of the outcomes. Among the controls for the characteristics of the child, age, having a job, and living in a single parent home all have positive and significant effects on the probability of engaging in risky behaviors. Having a primary care giver who smokes not only increases the probability of a child smoking but also increases the probability that they will drink and have sex. Primary care giver drinking and drug use do not have significant effects. Having evidence of drug use on the block, strangely perhaps, seems to decrease the probability of engaging in risky behaviors. Evidence of gang activity increases these probabilities.

[Table 5 here]

The values of the coefficients from a probit regression cannot be interpreted directly so we calculate the average change in the predicted probability when the network increases or decreases in size by 5 and 10 percent of its current value, keeping everything else constant. Since the measure of network is given by the multiplication of the strength of the network (proportion of adolescents of a certain ethnic group who live in a certain neighborhood) and the quality of the network (proportion of adolescents of a certain race who reside in a certain neighborhood that experience the outcome), and the strength of the network is included

separately in the regression, this increase (or a decrease) in the measure can be interpreted as a change in the quality of the network, holding everything else constant. Table 6 shows that smoking is the behavior that is least affected by an increase or a decrease in the quality of the network, while drinking is the most responsive behavior from deviations from its mean value. In the case of smoking, an increase of 10 (5) percent in the quality of the network increases the probability of smoking by 2.8 (1.39) percentage points, and a decrease of 10 (5) percent reduces the probability of smoking in 2.74 (1.38) percentage points. For drinking, the change in the probability ranges from 5.87 (2.95) for an increase of 10 (5) percent to -5.84 (-2.94) for a decrease of 10 (5) percent. Given these values, the effect of networks seem to be relevant in affecting behavior not only in a statistical sense, but also in an economic sense

Table 6. Change in Predicted Probability

	Increase		Decrease	
	10%	5%	10%	5%
smoke	3%	1%	-3%	-1%
alldrugs	3%	2%	-3%	-2%
drink	6%	3%	-6%	-3%
sex	3%	2%	-3%	-2%
skipschool	4.22%	2.10%	-4.05%	-2.06%

In order to examine the effect of neighborhood beliefs on teen behavior, we estimate Equation 2. Beliefs are measured on a scale of 1-3 (do not disapprove, disapprove and strongly disapprove). They are asked of adults, about adult behavior so they may not be directly relevant to teen behavior. Beliefs are asked about smoking, drinking and smoking pot and so we estimate the effect of belief only for teen smoking, drinking and drug use. We take the average beliefs about a particular activity of the adults in the ethnic group of a child in their neighborhood interacted with the size of the ethnic group in the neighborhood as our measure of beliefs. We want to see if including beliefs reduces the size and statistical significance of the network measure. Table 7 shows the estimated coefficients for the network measure and the belief measure. Neither the size nor the significance of the network effect is changed by including information about beliefs. The size and significance of the unreported variables is also unchanged. Beliefs have a positive and significant effect on drinking suggesting the a higher level of disapproval is associated with higher probabilities of drinking among teens. This result could happen for two reasons related to endogeneity. Either adults have negative beliefs about drinking because they know their teens drink, or teens who come from backgrounds with a strong disapproval of drinking rebel by drinking more. Unlike smoking and drug use, most adults do not disapprove of drinking.

Table 7. Beliefs in Probit Regression

	Have you ever smoked	Ever did drugs (pot or other)	Ever drink
Network	6.189	10.538	7.625
	(0.924)**	(1.384)**	(0.800)**
Beliefs	0.325	0.368	1.166
	(0.596)	(0.653)	(0.543)*

The results from the regressions in Table 6 and Table 7 could be biased due to omitted personal characteristics correlated with the network measure, simultaneity, and differential self-selection. We implement an IV approach to correct these sources of bias. Table 8 (see Table Appendix) show that omitted variable bias seems to be driving our results on network effects. None of the estimates associated with the network variable are statistically significant in the IV specification. The statistically significant results found on the previous regressions may be reflecting the existence of correlated effects or omitted variables instead of endogenous effects. However, the loss in variability that occurs when the quality measure calculated at the neighborhood level for each ethnic group is instrumented with a measure calculated at the national level could be a concern in our case, where we have a relatively small sample size and a high number of neighborhood fixed effects. For smoking, drug use and drinking our estimated coefficients are still positive though imprecisely estimated. For having sex and skipping school they are both negative and imprecisely estimated. The coefficient and the significance of the rest of the variables are stable with the biggest difference in the coefficients associated with ethnicity, though none of them, with the exception of the coefficient for Asians and Pacific Islanders for drinking behavior, are significant.

6 Extensions

One limitation of the literature about peer effects is the restrictive nature of the measure used to represent the group of influence. Usually, depending on the dataset used for the analysis, the literature has measured the peer group as individuals from the same neighborhood (e.g, Bertrand et al (2000), Aizer and Currie (2004), Musick et al (2007)) or individuals from the same school (e.g, Gaviria and Raphael (2001), Soetevent and Kooreman (2006), Clark and Loheac (2006), Powell et al (2005)). However, an individual is likely to be affected both by people from the same neighborhood and by people from the same school. Analysis based on peers only from school may also miss the extent to which these groups overlap. One advantage of using the L.A.FANS is that it contains information about the school that the adolescent attends. This allows us to improve the measurement of peer groups by including peers from school and peers from a child's neighborhood. The other issue in measuring peer groups is figuring out which peers are likely to have the strongest influence on a child. All

individuals from the same school or from the same neighborhood may not have the same influence over the individual (see, e.g. Kao and Joyner (2004) and Shrum et al (1988) for the relevance of race and ethnicity in group formation, and McPherson et al (2001) for a review of different characteristics that are important in group formation). To include this in the analysis, we plan on using the first wave of the Add Health survey to construct a weighted average of peer behavior, where the estimated weights reflect how influential each person in the group of influence could be given the characteristics of the individual. Among the characteristics we consider are: race, gender, grade in school, education of the adolescents' mothers, type of family, and IQ score (based on the idea that more intelligent kids hang out with kids of similar intelligence).

Using the second wave of the L.A.FANS, we also plan on improving the identification of a causal peer effect by using lagged and leaded peer behavior in a child's school and neighborhood as an instrument for current peer behavior. The lags and leads are sufficiently large so that none of the same individuals are included in both current and lagged or led behavior. This breaks the simultaneity problem in our regression created by the fact that an individual is affected by his or her peer group but he or she also affects the behavior of the group, and gives us an exogenous measure of peer behavior. This is possible because the L.A.FANS is not only a panel of individuals but also of neighborhoods so the same neighborhoods are present in both waves. This measure is not likely to suffer from the problem of having too little variation to identify effects. We plan on controlling for correlated effects using both fixed effects at tract level and controls for characteristics of the neighborhood at block level as we have done here. To solve the problem that affects the literature of separating the endogenous peer effects from the exogenous peer effects in which characteristics of the group affect behavior of the individual, we plan on using an instrumental variable strategy with lagged and leaded group characteristics as an instrument for current group characteristics. We can also use a similar idea to control for beliefs using lags and leads.

7 References

Aizer, A., Currie, J., 2004. Networks or neighborhoods? Correlations in the use of publicly-funded maternity care in California. Journal of Public Economics, 88(12), 2573-2585

Bertrand, M., Luttmer, E., Mullainathan, S., 2000. Network effects and welfare cultures. Quarterly Journal of Economics, 115, 1019-1056.

Brock, W. and Durlauf, S. 2007. Identification of Binary Choice Models with Social Interaction. Journal of Econometrics, 140, 57-75.

Brock, W. and Durlauf, S. 2001. Discrete Choice with Social Interactions. Review of

Economic Studies, 68, 235-260.

Case, A.C., Katz, L.F., 1991. The company you keep: the effects of family and neighborhood on disadvantaged youth. NBER Working Paper 3705. National Bureau of Economic Research.

Clark, A.E, Loheac, Y., 2007. "It wasn't me, it was them!" Social influence in risky behavior by adolescents. Journal of Health Economics, 26, 763-784.

Crane, J. 1991. The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing. American Journal of Sociology, 96(5), 1226-1259.

Gaviria, A., Raphael, S., 2001. School-based peer effects and juvenile behavior. Review of Economics and Statistics 83 (2), 507-268.

Kao, G., Joyner, K., 2004. Do race and ethnicity matter among friends? Activities among interracial, interethnic, and intraethnic adolescent friends. The Sociological Quarterly 45 (3), 557-573.

Katz, L., Kling, J., and Liebman, J. 2001. Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment. Quarterly Journal of Economics, 116(2), 607-654.

Katz, L., Kling, J., and Liebman, J. 2004. Beyond Treatment Effects: Estimating the Relationship Between Neighborhood Poverty and Individual Outcomes in the MTO Experiment. Working Paper 493. Industrial Relations Section, Princeton University.

Manski, C.F., 1993. Identification of endogenous social effects: the reflection problem. Review of Economic Studies 60 (3), 531-542.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annual Review of Sociology 27, 415-444.

Musick, K., Seltzer, J., Schwartz, C., 2008. Neighborhood norms and substance use among teens. Social Science Research, 37(1), 138-155

Myers, D. 2002. Demographic and Housing Transitions in South Central Los Angeles, 1990 to 2000. Special Report. Population Dynamics Research Group, School of Policy Planning and Development, University of Southern California.

Powell, L.M., Tauras, J.A., Ross, H., 2005. The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior. Journal of Health Economics 24 (5), 950-968.

Sacerdote, B. 2001. Peer Effects With Random Assignment: Results For Dartmouth Roommates. Quarterly Journal of Economics, 116(2), 681-704.

Sastry, N., Ghosh-Dastidar, B., Adams, J., Pebley, A. 2006. The design of a multilevel survey of children, families, and communities: The Los Angeles Family and Neighborhood Survey. Social Science Research 2006, 1000-1024.

Soetevent, A.R., Kooreman, P., 2006. A discrete choice model with social interactions with an application to high school teen behavior. Journal of Applied Econometrics.

Shrum, W., Cheek, N.H., Hunter, S.M., 1988. Friendship in school: gender and racial homophily. Sociology of Education 61 (4), 227-239.

Wilson, W.J. 1987. The Truly Disadvantaged: The Inner City, the Underclass and Public Policy. University of Chicago Press, Chicago.

8 Table Appendix

Table 4. Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Ever Smoke	833	.264	.441	0	1
Ever Drink	833	.374	.484	0	1
Ever Sex	833	.147	.354	0	1
Skip School	833	.168	.374	0	1
Ever Drugs	824	.252	.434	0	1
Network Smoke	832	.175	.140	0	.615
Network Drink	829	.263	.171	0	.636
Network Sex	825	.096	.100	0	.384
Network Skip	824	.172	.143	0	.538
Network Drugs	829	.108	.109	0	.538
Network Size	832	.649	.291	.045	1
Age	832	14.47	1.65	12	17
Male	832	.507	.500	0	1
Female	832	.493	.500	0	1
Latino	832	.558	.496	0	1
White	832	.253	.435	0	1
Black	832	.108	.310	0	1
Asian/Pacific Islander	832	.075	.264	0	1
PCG Drugs	824	.032	.178	0	1
PCG Smoke	825	.127	.333	0	1
PCG Drink	825	.351	.477	0	1
Child Foreign Born	794	.258	.437	0	1
Child Employed	829	.220	.415	0	1
PCG Education	824	11.40	4.55	0	20
Single Parent	832	.411	.492	0	1
Family Income (1000)	823	60.29	89.98	0	981
Public School	781	.834	.371	0	1
Magnet/Charter School	781	.047	.212	0	1
Private School	781	.119	.324	0	1
Church on Block	828	.367	.482	0	1
Evidence of Drugs on Block	828	.285	.451	0	1
Public Drinking on Block	828	.785	.411	0	1
Bar on Block	828	.222	.415	0	1
Police Visible on Block	828	.449	.497	0	1
Teens in Groups on Block	828	.474	.499	0	1
Gangs on Block	828	.103	.305	0	1
Neighborhood is Block	828	.747	.434	0	1
Number of Kids in Family	832	2.27	1.16	1	9
Instrument Network Smoke	832	.276	.172	.002	.565
Instrument Network Drink	829	.277	.167	.001	.546

Table 5. Probit Regression Results

Variable	Have you ever smoked	Ever did drugs (pot or other)	Ever drink	Ever sex	Ever Skip School
Network	6.005	10.028	7.502	13.465	7.681
	(0.901)**	(1.329)**	(0.803)**	(2.577)**	(1.181)**
Network size	-1.615	-2.536	-3.103	-2.698	-1.880
	(0.341)**	(0.397)**	(0.428)**	(0.453)**	(0.391)**
Age	0.261	0.267	0.297	0.319	0.214
	(0.041)**	(0.049)**	(0.043)**	(0.067)**	(0.047)**
Male	-0.053	0.037	-0.016	0.582	0.079
	(0.142)	(0.182)	(0.114)	(0.193)**	(0.130)
Latino	-0.152	-0.104	-0.061	0.187	-0.131
	(0.156)	(0.205)	(0.150)	(0.290)	(0.227)
White (omitted for smoking, drugs, drinking)				0.117	-0.258
				(0.261)	(0.199)
Black (omitted for sex, skipping school)	0.073	-0.351	-0.230		
	(0.198)	(0.240)	(0.198)		
Asian	0.160	0.000	-0.641	0.134	-0.127
	(0.227)	(0.310)	(0.244)**	(0.398)	(0.260)
PCG drugs	-0.155	0.279	0.305	0.493	0.287
	(0.366)	(0.374)	(0.388)	(0.368)	(0.278)
PCG smoke	0.348	0.245	0.344	0.554	0.082
	(0.208)+	(0.234)	(0.193)+	(0.281)*	(0.196)
PCG drink	0.107	0.294	0.124	0.013	0.061
	(0.171)	(0.206)	(0.145)	(0.226)	(0.154)
Not Born in The US	0.058	-0.157	-0.293	-0.318	-0.012
	(0.166)	(0.248)	(0.162)+	(0.219)	(0.159)
For Kids: Have a paid job now/In past month	0.532	0.388	0.305	0.732	0.536
	(0.146)**	(0.200)+	(0.175)+	(0.203)**	(0.173)**
PCG Education	-0.028	0.022	-0.013	0.023	-0.013
	(0.020)	(0.027)	(0.022)	(0.027)	(0.020)
Single Parent Family	0.314	0.319	0.111	0.449	0.027
	(0.148)*	(0.187)+	(0.134)	(0.190)*	(0.147)
Family Income	0.001	0.002	-0.000	0.002	0.000
	(0.001)	(0.001)*	(0.001)	(0.001)*	(0.001)
Public School (omitted category)					
Magnet or Charter School	-0.679	-0.291	0.419	1.017	0.110
	(0.316)*	(0.394)	(0.303)	(0.462)*	(0.281)
Private School	-0.001	-0.358	-0.256	0.048	-0.267
	(0.261)	(0.316)	(0.253)	(0.320)	(0.250)
Church on Block	-0.119	0.083	0.073	-0.184	-0.218
	(0.168)	(0.237)	(0.198)	(0.274)	(0.206)
Drug paraphenalia or selling drugs on block	-0.526	-0.740	-0.223	0.089	-0.181
	(0.303)+	(0.363)*	(0.215)	(0.275)	(0.162)
Evidence of public drinking on block	-0.002	-0.065	-0.369	0.098	0.306
	(0.212)	(0.261)	(0.174)*	(0.301)	(0.197)
Bar on block	-0.243	0.514	0.270	-0.148	0.351
	(0.242)	(0.291)+	(0.179)	(0.245)	(0.189)+
Police seen on block	0.284	0.280	-0.177	-0.026	-0.027
	(0.222)	(0.257)	(0.185)	(0.236)	(0.187)
Teens in groups on block	-0.213	-0.005	0.303	0.264	-0.225
	(0.181)	(0.199)	(0.182)+	(0.237)	(0.160)
Evidence of gangs on block	0.866	0.729	0.179	0.456	0.025
	(0.274)**	(0.324)*	(0.188)	(0.388)	(0.231)
Area is about a block	-0.122	-0.081	-0.209	0.381	0.117
	(0.153)	(0.161)	(0.147)	(0.229)+	(0.167)
Number of Kids in family	0.004	-0.016	-0.083	-0.014	0.023
	(0.062)	(0.076)	(0.054)	(0.071)	(0.057)
Constant	-3.840	-4.267	-4.427	-8.428	-3.663
	(0.903)**	(0.953)**	(0.926)**	(1.322)**	(0.938)**
Observations	678	574	703	578	683

Robust standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 8. Instrumental Variable Probit Regression Results					
	Have you ever smoked	Ever did drugs (pot or others)	Ever Drink	Ever Sex	Ever Skip School
Variable Network	1.861	13.546	0.921	-1.264	-3.112
	(10.789)	(13.645)	(8.260)	(12.999)	(11.713)
Network Size	-0.411	-3.293	-0.305	-0.117	1.673
	(3.151)	(2.961)	(3.524)	(2.373)	(3.784)
Age	0.266	0.254	0.306	0.305	0.214
	(0.042)**	(0.083)**	(0.039)**	(0.073)**	(0.046)**
Male	-0.053	0.037	-0.037	0.507	-0.005
	(0.129)	(0.153)	(0.117)	(0.205)*	(0.150)
Latino	-0.371	0.213	-0.434	-0.178	-0.771
	(0.589)	(1.311)	(0.480)	(0.495)	(0.685)
White (omitted for smoking, drugs, drinking)				0.010	-0.474
				(0.417)	(0.352)
Black (omitted for sex, skipping school)	-0.169	-0.161	-0.646		
Asian/Paci...c Islander	(0.693)	(0.878)	(0.552)		
	-0.004	0.152	-1.190	-0.328	-0.325
	(0.534)	(0.749)	(0.690)+	(0.600)	(0.389)
PCG drugs	-0.182	0.244	0.237	0.349	0.314
	(0.357)	(0.386)	(0.317)	(0.431)	(0.321)
PCG smoke	0.349	0.218	0.408	0.575	0.052
	(0.196)+	(0.263)	(0.186)*	(0.235)*	(0.198)
PCG drinks	0.099	0.286	0.098	-0.018	0.036
	(0.155)	(0.183)	(0.138)	(0.177)	(0.140)
Not Born in The US	0.043	-0.135	-0.287	-0.236	-0.006
	(0.175)	(0.237)	(0.153)+	(0.241)	(0.154)
For Kids: Have a paid job now/In past month	0.515	0.377	0.291	0.640	0.476
	(0.175)**	(0.192)*	(0.148)*	(0.250)*	(0.190)*
PCG Education	-0.028	0.025	-0.009	0.022	0.000
	(0.023)	(0.029)	(0.020)	(0.028)	(0.025)
Single Parent Family	0.357	0.300	0.155	0.469	0.132
	(0.167)*	(0.190)	(0.130)	(0.174)**	(0.162)
Family Income	0.001	0.002	-0.000	0.002	-0.000
	(0.001)	(0.001)+	(0.001)	(0.001)	(0.001)
Public School (omitted category)					
Magnet or Charter School	-0.662	-0.293	0.525	0.909	-0.098
	(0.387)+	(0.376)	(0.316)+	(0.425)*	(0.388)
Private School	-0.116	-0.288	-0.362	-0.197	-0.478
	(0.376)	(0.428)	(0.250)	(0.368)	(0.292)
Church on Block	-0.141	0.069	0.059	-0.097	-0.220
	(0.197)	(0.251)	(0.169)	(0.256)	(0.172)
Drug paraphenalia or selling drugs on block	-0.522	-0.732	-0.217	0.280	-0.118
	(0.229)*	(0.300)*	(0.190)	(0.295)	(0.216)
Evidence of public drinking on block	-0.028	-0.059	-0.382	0.112	0.270
	(0.223)	(0.238)	(0.196)+	(0.274)	(0.214)
Bar on block	-0.278	0.541	0.331	-0.010	0.400
	(0.242)	(0.283)+	(0.202)	(0.310)	(0.197)*
Police seen on block	0.268	0.304	-0.183	0.012	-0.112
	(0.192)	(0.235)	(0.164)	(0.230)	(0.192)
Teens in groups on block	-0.199	-0.016	0.257	0.147	-0.195
	(0.196)	(0.232)	(0.177)	(0.257)	(0.176)
Evidence of gangs on block	0.771	0.778	0.061	0.297	-0.147
	(0.393)*	(0.348)*	(0.277)	(0.364)	(0.296)
Area is about a block	-0.144	-0.063	-0.219	0.276	0.106
	(0.156)	(0.192)	(0.134)	(0.233)	(0.141)
Number of Kids in family	0.009	-0.015	-0.073	-0.021	0.052
	(0.061)	(0.072)	(0.056)	(0.074)	(0.059)
Constant	-3.620	-4.191	-3.340	-5.999	-3.581
	(1.005)**	(1.154)**	(1.734)+	(3.319)+	(0.948)**
Observations	678	574	703	578	683
Robust standard errors in parentheses + significant at 10%; * significant at 5%; ** significant at 1%					

9 Figure Appendix