Education inequality and use of cigarettes, alcohol, and marijuana

Sandro Galea a,b,c,*, Jennifer Ahern a,1, Melissa Tracy a,2, Sasha Rudenstine a,3, David Vlahov b,c,d,4

1 Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, USA
2 Center for Urban Epidemiologic Studies, New York Academy of Medicine, New York, NY, USA
3 Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA
4 Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Received 15 March 2006; received in revised form 5 November 2006; accepted 6 November 2006

Abstract

Education inequality at the neighborhood-level may influence population health and health behavior. We assessed the relations between education inequality and substance use in 59 New York City (NYC) neighborhoods. We used Gini coefficients of education to describe neighborhood education inequality and data from a random-digit-dial phone survey of adult residents of NYC to assess use of substances. Among 1355 respondents (female = 56.2%; white = 35.7%; mean age = 40.4), 23.9% (95% confidence interval [CI] = 20.3–27.5) reported smoking, 39.4% (95% CI = 35.3–43.4) drinking, and 5.4% (95% CI = 3.6–7.3) using marijuana in the previous 30 days. In multilevel models controlling for neighborhood education, neighborhood income inequality, and individual covariates, living in a neighborhood with high education inequality was associated with a greater prevalence of drinking (p = 0.02) and of smoking marijuana (p = 0.004) but among current drinkers it was associated (p = 0.03) with having fewer drinks. The odds of alcohol use (OR = 1.70) and marijuana use (OR = 3.49) were greater in neighborhoods in the 75th percentile of education Gini compared to neighborhoods in the 25th percentile of education Gini. Statistical interactions suggest that there may be a stronger relation between education inequality and marijuana use in neighborhoods with low mean education than in neighborhoods with higher mean levels of education. These findings, taken together, suggest a complex relation between education inequality and substance use; likelihood of the use of alcohol and marijuana was higher in areas with higher education inequality suggesting potential roles for substance use norms and availability, whereas quantity used among drinkers was higher in areas with low education inequality, suggesting potential roles for both disadvantage and norms.

Keywords: Education; Inequality; Neighborhood; Cigarettes; Alcohol; Marijuana; Drug use

1. Introduction

Education is often considered one of the “fundamental” determinants of health and a primary indicator of individual socio-economic status (Link and Phelan, 2000; Adler and Ostrove, 1999; Adler and Newman, 2002). The relation between education and health is well-established; persons who are better educated live longer and suffer less morbidity during their lifetimes (Hemmingway et al., 2000; Bobak et al., 1999; Lynch et al., 1995). However, the relation between education and use and misuse of substances is less consistent. For example, although it repeatedly has been shown that low educational attainment is associated with greater risk of smoking throughout the life course (SAMHSA, 2003; Barbeau et al., 2004; Helmert et al., 2001; Jefferis et al., 2003, 2004; Gilman et al., 2003), persons with higher education are more likely to drink alcohol (SAMHSA, 2003; Moore et al., 2005; Casswell et al., 2003), although they are less likely to binge drink (SAMHSA, 2003;
Casswell et al., 2003; Karlamangla et al., 2006). Persons with higher education are also more likely to use marijuana throughout their lifetime (Stenbacka et al., 1993).

In the past two decades a body of work has considered whether group measures of socioeconomic status are associated with health, independent of the role of individual socioeconomic status. In particular, there is a substantial literature assessing the relationship between the distribution of income (frequently referred to in the public health literature as “income inequality”) and population health (Wilkinson, 1992; Subramanian and Kawachi, 2004; Lynch et al., 2004a,b). Although the evidence in the field remains controversial, recent systematic reviews of the literature suggest that although there is little consistent evidence for a cross-national relation between income distribution and health, there may be a relation between income maldistribution and indicators of poorer health in the United States (US) at the state, city, and neighborhood levels (Lynch et al., 2004a,b).

Although the literature in this regard is sparse, recent work suggests that there also may be a relation between distribution of education and population health (Galea and Ahern, 2005). The presence of people with a wide range of educational attainment within a group may be accompanied by positive externalities (“spill-over” benefits) generated by the presence and actions of persons with high educational attainment (Checchi, 2001; Galea and Ahern, 2005). For example, health education messages developed by health care facilities at the level demanded by their most educated patients would then benefit all of those who use their services. Similarly, more educated persons may have access to persons in power and successfully lobby against cigarette advertising in their neighborhood, as such advertising has been shown to be associated with greater cigarette smoking (Schooler et al., 1996). While such improvements in the determinants of health may be driven by persons who are more educated, they will then be available to all others in a particular area as long as the improved resources are not more costly for individuals to access. Therefore, it is plausible that a small group of persons who are more educated may contribute to the improvement of shared facilities and resources in a given area. These shared facilities and resources, barring significant financial barriers to entry, may contribute to improved well-being among all persons in a particular area. These benefits may be particularly important in the context of health indicators that are likely to be affected by short-term changes in the social environment, such as substance use.

Therefore, distribution of education may be an important determinant of population health. Although in the US everyone has access to primary and secondary school education, there is a wide range in educational attainment (US Census, 2000). There are substantial educational disparities between various racial/ethnic and socioeconomic groups in the US (Christensen and Johnson, 1995; Sandefur and Pahari, 2004; Graetz, 1987) and it has been argued that these disparities may contribute to racial/ethnic inequalities in health (Thomas et al., 2000). However, we are aware of only one previous study that has explicitly studied the role of area-level education distribution as a potential determinant of health. In that paper, the authors showed that education distribution was positively associated with health indicators that may be sensitive to short term changes in the social environment (homicide, infant mortality, low birthweight, late or no prenatal care) when taking into account neighborhood education, income, and income distribution (Galea and Ahern, 2005).

There is relatively little research on the role of social and contextual (or group-level) variables in determining substance use behaviors (Galea et al., 2004, 2005). Although there is an emerging body of literature that documents an association between living in economically deprived areas and higher prevalence of smoking (Diez-Roux et al., 1997; Kleinschmidt et al., 1995; Jones and Duncan, 1985; Reijneveld, 1998; Ecob and MacIntyre, 2000), and drinking (Hill and Angel, 2005), this evidence is inconsistent. We are not aware of previous work that has explicitly studied the role of area-level education distribution as a potential determinant of substance use, or any other health behavior. In this paper we assessed the relation between education distribution and the use of cigarettes, alcohol, and marijuana in New York City (NYC) neighborhoods. We hypothesized that distribution of education at the neighborhood level would be positively associated with substance use, and that in neighborhoods where there is more heterogeneous educational attainment there would be lower use of substances, when accounting for individual income. It was our goal both to assess the potential relation between education distribution and use of substances, and also to further advance empirical inquiry into the role of contextual characteristics in shaping risk of substance use and misuse.

2. Methods

2.1. Individual-level variables

Individual-level data for this study were obtained from a cross-sectional random digit dial (RDD) household telephone survey that included measures of substance use. The survey, carried out between 25 March and 25 June 2002, was designed to assess mental health in the New York City (NYC) metropolitan area in the aftermath of the 11 September terrorist attacks in NYC. The sampling frame for the survey included all adults in the NYC metropolitan area with over-sampling of residents in NYC; this analysis is limited to residents of NYC.

The cooperation rate, based on the sum of the number of completed interviews, quota outs and screen-outs (i.e. 1570 + 518 + 117 + 1362 + 71), was 60%. The final sample of respondents did not differ significantly from the 2000 census estimates of New York City (US Bureau of the Census 2000). The Institutional Review Board of the New York Academy of Medicine reviewed and approved this work. Further details on this survey can be found elsewhere (Galea et al., 2003b; Vlahov et al., 2004).

Respondents were interviewed using a structured questionnaire. The primary outcome variables for this analysis were respondents’ cigarette smoking, alcohol drinking, and marijuana smoking. For each of the three substances we asked the following series of questions. First we asked if the respondent had used the substance in the previous 12 months (e.g., “Have you smoked cigarettes in the last 12 months?”). Respondents who answered “Yes” to this question were asked to report on how many days they had used the substance in the 30 days prior to the survey, and the average number of times the substance was used per day. This information was used to calculate the total number of cigarettes smoked, number of alcoholic drinks consumed, and number of times marijuana was smoked in the past 30 days. Of the sample of 1355 NYC residents, 10 respondents (0.7%) were missing data for the use of cigarettes in the past 30 days, 21 (1.5%) for...
the consumption of alcohol, and 18 (1.3%) for marijuana use. For the analyses presented here we examined use of each of these substances individually. The survey also included assessment of demographic characteristics including age, race/ethnicity, gender, yearly household income, and education.

2.2. Neighborhood definition

NYC is divided into 59 residential community districts (CDs) by the Department of City Planning. CDs are well-defined units, each with an administrative community board, that as such have political and social a priori significance for their residents (Messer and Tardiff, 1986; Marzuk et al., 1997; Suecoff et al., 1999; Galea et al., 2003a). Examples of these CDs include the Upper West Side in Manhattan and Bedford-Stuyvesant in Brooklyn. These CDs will be referred to as neighborhoods hereafter.

2.3. Education and education distribution

We used 2000 US Census data on educational attainment among individuals 25 years or older to estimate mean educational levels and distribution of education in NYC neighborhoods (Bureau of the Census, 2000). Mean educational levels were calculated via the following equation:

$$
\mu = \frac{1}{n} \sum_{i=1}^{n} p_i y_i
$$

where $y_i$ is the proportion of individuals at a given level of schooling in the population of interest and $y_i$ is the midpoint of (or the most likely value for) the schooling category (e.g., $y_i = 5.5$ for completion of fifth and sixth grades, $y_i = 16$ for the completion of a bachelor's degree).

The education Gini coefficient was used to measure the distribution of education and the extent of inequality in each neighborhood (Thomas et al., 2000; Kawachi and Kennedy, 1997; Deaton, 1997). A Gini coefficient of 0 denotes a perfectly equitable education distribution, whereas a coefficient of 1.0 represents maximal maldistribution. The two methods – direct and indirect – used in calculating Gini coefficients have been discussed and explored extensively in the income distribution literature. Briefly, the direct method is “the ratio of the mean of half of the average over all pairs of the absolute deviations between [all possible pairs of people]” (Deaton, 1997 [p. 139]). When the indirect method is used, the Gini coefficient is calculated from the Lorenz curve, which is created by plotting proportions of the population from least to most educated on the x-axis and proportions of educational attainment on the y-axis. The Gini coefficient is the area between the diagonal line indicating no inequality and the concave line representing the education distribution in a particular population. The Gini coefficient was rescaled to range from 0 to 100 for this analysis so that regression parameter estimates could be more easily interpreted.

Given the sample size in this analysis, we used the small-sample Gini estimation (Thomas et al., 2000). This small sample formula is related, through the factor $N(N-1)$, to the large-sample Gini calculation. In practical terms, when a sample is large enough, $N(N-1)$ is approximately equal to 1, and the small-sample approximation is equivalent to the large-sample formula. This definition is mathematically represented as follows:

$$
E = \left(\frac{N}{N-1}\right) \left[\frac{1}{N} \sum_{j=1}^{n} p_j y_j \right]-1
$$

where $E$ is the education Gini coefficient, $N$ the number of individuals in the population of interest, $\mu$ the mean number of years of schooling in the population of interest, $y_j$ and $y_j$ the years of schooling at different educational attainment levels, and $n$ is the number of levels of educational attainment. We used 16 levels of educational attainment in this study (Barro and Lee, 2001).

2.4. Income distribution

In order to adjust for income distribution as a potential confounder of the observed relation, we calculated income Gini coefficients. We used household income data from the 2000 US Census to calculate the Gini coefficient as a measure of income distribution in each NYC neighborhood (Bureau of the Census, 2000).

2.5. Data analysis

Statistical weights were used in all analyses to correct potential bias related to the number of household telephones, persons in the household, and oversampling. Weights represented the inverse probability of selection for interview. Hence, the sample weight included a component that was inverse to number of household telephones, proportional to number of persons in the household, and inversely proportional to the population sampling fraction. We described demographic characteristics of the survey population and compared these characteristics to the demographic distribution suggested by the 2000 US Census (Bureau of the Census, 2000). All survey respondents who could not be geocoded (due to missing or incorrect addresses) were excluded from these analyses; we compared characteristics of the persons included in these analyses to those who were excluded to examine the potential for bias due to the exclusion of these participants. We used logistic regression models to test the bivariate relations between the individual and neighborhood-level covariates of interest and prevalence of substance use for each of the three substances. We used linear regression to test bivariate relations between covariates and the frequency of substance use among those who had used a substance. Since the total number of persons who had used marijuana in the 30 days prior to the survey was small (92) we did not have sufficient power to assess the determinants of frequency of marijuana use among marijuana users. The linearity of the relation between each covariate and outcome was assessed using likelihood ratio tests ($p < 0.1$). The prevalence of the use of each substance was calculated and graphed by thirds of education inequality. Generalized estimating equations (GEE) were used to fit separate multilevel multiple logistic regression models that assessed the relation between neighborhood-level covariates (mean educational level, education inequality, and income inequality) and prevalence of use of each of the three substances (Zeger and Liang, 1986; Merlo, 2003); models were also adjusted for individual age, gender, race/ethnicity, income, and educational attainment. To assess the magnitude of the relation between education distribution and likelihood of substance use we calculated the odds ratio for percentiles of education distribution (for the range of education Gini coefficients in the dataset) setting the 25th percentile as the referent. For the models where there was a relation between education Gini and substance use, we constructed separate models to assess statistical interaction between neighborhood education distribution and neighborhood education through the use of interaction terms between education Gini and mean neighborhood education. In order to interpret the interaction terms in the models which included them we calculated the relative odds of substance use for individuals living in neighborhoods with different levels of education Gini and mean education. Finally, among those who had used cigarettes and among those who had used alcohol in the month prior to the survey, separate multilevel linear GEE models were fit to assess the relations between neighborhood mean educational level and education inequality and the number of cigarettes smoked and the number of alcoholic drinks consumed, respectively, in the 30 days prior to the survey assessment. In order to consider whether the associations documented here were particularly influenced by the circumstances of this study, we also reran all the statistical models also adjusting for a variable that described whether the participants were affected by the 11 September 2001 terrorist attacks. This variable indicated whether participants were in the World Trade Center complex during the attacks, were injured during the attacks, lost possessions or property, had a friend or relative killed, lost a job as a result of the attacks, or were involved in the rescue effort.

3. Results

Overall, 1570 people responded to the telephone survey. Of these individuals we were able to link 1355 to their neighborhood of residence and all analyses presented have been restricted to this latter sample. There were no significant differences in characteristics of the persons included in these analyses and
Table 1
Demographic characteristics of survey population

<table>
<thead>
<tr>
<th>Total</th>
<th>Included</th>
<th>Excluded</th>
<th>Total vs. p-valuea</th>
<th>Included vs. p-valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Total</td>
<td>1570</td>
<td></td>
<td>1355</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–24</td>
<td>157</td>
<td>14.7</td>
<td>144</td>
<td>16.0</td>
</tr>
<tr>
<td>25–34</td>
<td>414</td>
<td>27.0</td>
<td>357</td>
<td>27.8</td>
</tr>
<tr>
<td>35–44</td>
<td>329</td>
<td>19.6</td>
<td>279</td>
<td>18.3</td>
</tr>
<tr>
<td>45–54</td>
<td>286</td>
<td>18.3</td>
<td>248</td>
<td>18.0</td>
</tr>
<tr>
<td>55–64</td>
<td>175</td>
<td>11.2</td>
<td>148</td>
<td>11.0</td>
</tr>
<tr>
<td>65+</td>
<td>190</td>
<td>9.2</td>
<td>134</td>
<td>9.0</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>697</td>
<td>44.1</td>
<td>616</td>
<td>43.8</td>
</tr>
<tr>
<td>Female</td>
<td>873</td>
<td>55.9</td>
<td>739</td>
<td>56.2</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>774</td>
<td>35.8</td>
<td>682</td>
<td>35.7</td>
</tr>
<tr>
<td>Asian</td>
<td>118</td>
<td>6.3</td>
<td>102</td>
<td>6.3</td>
</tr>
<tr>
<td>African-American</td>
<td>264</td>
<td>23.7</td>
<td>220</td>
<td>24.2</td>
</tr>
<tr>
<td>Hispanic</td>
<td>332</td>
<td>28.7</td>
<td>291</td>
<td>29.7</td>
</tr>
<tr>
<td>Other</td>
<td>53</td>
<td>5.5</td>
<td>40</td>
<td>4.2</td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$100,000+</td>
<td>233</td>
<td>10.7</td>
<td>213</td>
<td>11.3</td>
</tr>
<tr>
<td>$75,000–99,999</td>
<td>119</td>
<td>10.3</td>
<td>102</td>
<td>9.8</td>
</tr>
<tr>
<td>$50,000–74,999</td>
<td>221</td>
<td>16.5</td>
<td>196</td>
<td>16.1</td>
</tr>
<tr>
<td>$40,000–49,999</td>
<td>110</td>
<td>7.3</td>
<td>93</td>
<td>6.7</td>
</tr>
<tr>
<td>$30,000–39,999</td>
<td>155</td>
<td>14.7</td>
<td>134</td>
<td>14.9</td>
</tr>
<tr>
<td>$20,000–29,999</td>
<td>161</td>
<td>16.9</td>
<td>142</td>
<td>16.5</td>
</tr>
<tr>
<td>&lt;$20,000</td>
<td>303</td>
<td>23.6</td>
<td>265</td>
<td>24.7</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate degree</td>
<td>282</td>
<td>10.1</td>
<td>255</td>
<td>10.6</td>
</tr>
<tr>
<td>College degree</td>
<td>504</td>
<td>27.4</td>
<td>442</td>
<td>27.9</td>
</tr>
<tr>
<td>Some college</td>
<td>273</td>
<td>22.0</td>
<td>239</td>
<td>21.6</td>
</tr>
<tr>
<td>High school graduate/GED</td>
<td>295</td>
<td>24.7</td>
<td>248</td>
<td>24.0</td>
</tr>
<tr>
<td>&lt;High school graduate</td>
<td>207</td>
<td>15.9</td>
<td>169</td>
<td>16.0</td>
</tr>
<tr>
<td>Directly affected by 11 September</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1074</td>
<td>70.2</td>
<td>909</td>
<td>69.6</td>
</tr>
<tr>
<td>Yes</td>
<td>496</td>
<td>29.8</td>
<td>446</td>
<td>30.4</td>
</tr>
<tr>
<td>Any cigarettes in last 30 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1194</td>
<td>75.2</td>
<td>1037</td>
<td>76.1</td>
</tr>
<tr>
<td>Yes</td>
<td>361</td>
<td>24.8</td>
<td>308</td>
<td>23.9</td>
</tr>
<tr>
<td>Any alcoholic drink in last 30 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>794</td>
<td>60.7</td>
<td>673</td>
<td>60.6</td>
</tr>
<tr>
<td>Yes</td>
<td>747</td>
<td>39.3</td>
<td>661</td>
<td>39.4</td>
</tr>
<tr>
<td>Any marijuana in last 30 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1448</td>
<td>94.8</td>
<td>1245</td>
<td>94.6</td>
</tr>
<tr>
<td>Yes</td>
<td>102</td>
<td>5.2</td>
<td>92</td>
<td>5.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neighborhood characteristics</th>
<th>Mean</th>
<th>S.D.</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean education</td>
<td>12.60</td>
<td>1.47</td>
<td>12.42</td>
<td>10.06–16.10</td>
</tr>
<tr>
<td>Education Gini</td>
<td>0.16</td>
<td>0.04</td>
<td>0.16</td>
<td>0.09–0.26</td>
</tr>
</tbody>
</table>

a Two-tailed $\chi^2$ p-value comparing those included in the analysis and the entire sample.

b Two-tailed $\chi^2$ p-value comparing those included in the analysis and those excluded from the analysis.

the 215 persons excluded from the analyses. Table 1 presents baseline characteristics of the sample used for these analyses. Mean age was 40.4 (standard deviation 12.9), 43.8% were male, 35.7% were white, 6.3% Asian, 24.2% African-American, and 29.7% Hispanic. A plurality of participants had an income of under $20,000 (24.7%), 16.5% had an income between $20,000 and $29,999, and 16.1% had an income between $50,000 and $74,999. Table 1 also compares this group to all the persons
in the sample and all the persons excluded from the analysis (i.e., 215 persons). As shown in Table 1 there were no appreciable differences between the total sample of 1570 persons and the 1355 persons who were included in this analysis, or between those included and those excluded from the analysis. There were no differences between groups either on demographic characteristics or on key measures of substance use considered here, including use of cigarettes, alcohol, or marijuana in the previous 30 days. As shown in Table 1, in the included analytic sample, overall, 23.9% of persons smoked (95% confidence interval [CI] = 20.3–27.5), 40.0% (95% CI = 35.3–43.4) used alcohol, and 5.4% (95% CI = 3.6–7.3) used marijuana in the 30 days prior to the survey. These 30-day prevalences of substance use are comparable to national 30-day prevalence estimates of use of these substances (SAMHSA, 2003). Among those who smoked, a mean of 262 cigarettes were smoked in the past 30 days (standard deviation 258.7). Among those who drank alcohol, a mean of 18 alcoholic drinks were consumed in the 30 days prior to the survey (standard deviation 20.9). Table 1 also shows that median neighborhood education across all 59 New York City neighborhoods was 12.4 years (range 10.1–16.1); and mean education Gini coefficient was 0.16 (range 0.09–0.26). This is comparable to values calculated by other authors for the US as a whole (Thomas et al., 2000).

We calculated the prevalence of past 30-day use of each substance by thirds of neighborhood education inequality. Fig. 1 shows the relations between education inequality and prevalence of cigarette smoking, alcohol drinking, and marijuana smoking, unadjusted for mean neighborhood education. We found no association between education inequality and prevalence of cigarette (p = 0.84) or marijuana smoking (p = 0.65). However, there was a greater prevalence of alcohol use in neighborhoods characterized by low education inequality (p < 0.05).

Table 2 presents three multilevel logistic regression models assessing the relations between neighborhood mean education, education Gini coefficient, and individual substance use (cigarettes, alcohol, and marijuana) in the past 30 days. The models are adjusted for individual-level covariates, and the interaction between neighborhood mean education and education inequality is examined. Neither mean education nor education Gini were associated with cigarette smoking in the model adjusted for neighborhood-level characteristics or in the model additionally adjusted for individual-level covariates. The only significant predictors of cigarette use were age (β = −0.02; p = 0.03), being male (β = 0.65; p = 0.02), and having completed graduate work (β = −1.06; p = 0.04). Mean education (β = 0.62; p < 0.0001) and education Gini (β = 0.10; p = 0.01) were both positively associated with alcohol use, adjusting for neighborhood income inequality. These relations persisted after adjustment for individual-level covariates, with higher mean education (β = 0.43; p = 0.0003) and higher education inequality (β = 0.09; p = 0.02) associated with greater prevalence of alcohol use in the past 30 days. Other significant predictors of alcohol use included individual-level income (β = 0.06; p = 0.01), age (β = −0.03; p < 0.0001), male gender (β = 0.64; p = 0.0002), and race/ethnicity (β = −1.72 for Asians, p = 0.001; β = −0.66 for African-Americans, p = 0.02; and β = −0.90 for Hispanics, p = 0.003, all compared to Whites). In a separate fully adjusted model, the interaction term between education Gini and mean education was not significantly associated with any alcohol use (β = −0.01; p = 0.71). Both mean education (β = 0.69; p = 0.01) and education Gini (β = 0.22; p = 0.004) were positively associated with marijuana use after controlling for neighborhood income inequality and individual-level covariates; male gender (β = −0.10; p < 0.0001) and race/ethnicity (β = −5.95 for Asians, p = 0.01; β = −2.08 for Hispanics, p = 0.005, both compared to Whites) were also associated with marijuana use. In the model including the interaction term between education Gini and mean education, the interaction term was statistically significant (β = −0.05; p = 0.03). Neighborhood income inequality was not associated with use of any of the three substances in adjusted models.

![Fig. 1. Prevalence of substance use in the past 30 days by neighborhood education inequality.](image-url)
Table 2
Multilevel logistic regression models predicting use of substances

<table>
<thead>
<tr>
<th></th>
<th>Any cigarettes in last 30 days, N = 1115&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Any alcoholic drink in last 30 days, N = 1107&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Any marijuana in last 30 days, N = 1110&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta )</td>
<td>p-Value</td>
<td>( \beta )</td>
</tr>
<tr>
<td><strong>Neighborhood-level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>(-0.18)</td>
<td>0.93</td>
<td>(-2.06)</td>
</tr>
<tr>
<td>Mean education</td>
<td>(-0.09)</td>
<td>0.49</td>
<td>0.1</td>
</tr>
<tr>
<td>Education Gini&lt;sup&gt;b&lt;/sup&gt;</td>
<td>(-0.03)</td>
<td>0.54</td>
<td>(-0.002)</td>
</tr>
<tr>
<td>Income&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.02</td>
<td>0.73</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Individual-level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>(-0.06)</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td>Age</td>
<td>(-0.02)</td>
<td>0.03</td>
<td>(-0.03)</td>
</tr>
<tr>
<td>Male&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.05</td>
<td>0.02</td>
<td>0.84</td>
</tr>
<tr>
<td><strong>Race/ethnicity&lt;sup&gt;e&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>(-0.56)</td>
<td>0.36</td>
<td>(-1.72)</td>
</tr>
<tr>
<td>African-American</td>
<td>(-0.11)</td>
<td>0.71</td>
<td>(-0.86)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>(-0.21)</td>
<td>0.62</td>
<td>(-0.9)</td>
</tr>
<tr>
<td>Other</td>
<td>(-0.83)</td>
<td>0.29</td>
<td>(-0.43)</td>
</tr>
<tr>
<td><strong>Education&lt;sup&gt;f&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school graduate/GED</td>
<td>0.46</td>
<td>0.21</td>
<td>0.04</td>
</tr>
<tr>
<td>Some college</td>
<td>0.53</td>
<td>0.19</td>
<td>0.37</td>
</tr>
<tr>
<td>College degree</td>
<td>(-0.2)</td>
<td>0.63</td>
<td>0.69</td>
</tr>
<tr>
<td>Graduate work</td>
<td>(-1.06)</td>
<td>0.04</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Mean education x education Gini

\(-0.01\)  | 0.71          | \(-0.05\)  | 0.03

<sup>a</sup> N reflects sample with non-missing data for all variables included in the model.

<sup>b</sup> Rescaled to range from 0 to 100.

<sup>c</sup> Female as referent.

<sup>d</sup> White as referent.

<sup>e</sup> <High school graduate as referent.
Fig. 2. Relative odds of marijuana use at different values of neighborhood education Gini and neighborhood mean education, predicted from final models including interaction terms. Low education corresponds to the 25th percentile, mid education to the 50th percentile, and high education to the 75th percentile of neighborhood mean education.

To illustrate the magnitude of the associations found in this analysis, we calculated the relative odds of alcohol use by percentiles of education Gini coefficient. Compared to the 25th percentile, in neighborhoods in the fiftieth percentile of education Gini distribution the relative odds of alcohol use were 1.24 (95% confidence interval [CI] = 1.01–1.50) and in neighborhoods in the 75th percentile the relative odds were 1.70 (95% CI = 1.04–2.79). The relative odds of marijuana use by percentiles of education Gini coefficient were as follows. Compared to the 25th percentile, in neighborhoods in the fiftieth percentile of education Gini distribution the relative odds of marijuana use were

Table 3

<table>
<thead>
<tr>
<th>Number of cigarettes in last 30 days, among smokers; N = 259a</th>
<th>Number of alcoholic drinks in last 30 days, among drinkers; N = 571a</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>p-Value</td>
</tr>
<tr>
<td>Neighborhood-level</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>169.13</td>
</tr>
<tr>
<td>Mean education</td>
<td>19.29</td>
</tr>
<tr>
<td>Education Gini</td>
<td>3.62</td>
</tr>
<tr>
<td>Income Gini</td>
<td>−4.73</td>
</tr>
<tr>
<td>Individual-level</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>3.91</td>
</tr>
<tr>
<td>Age</td>
<td>4.64</td>
</tr>
<tr>
<td>Malec</td>
<td>119.67</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>−99.79</td>
</tr>
<tr>
<td>African-American</td>
<td>−159.69</td>
</tr>
<tr>
<td>Hispanic</td>
<td>−117.22</td>
</tr>
<tr>
<td>Other</td>
<td>−255.87</td>
</tr>
<tr>
<td>Educatione</td>
<td></td>
</tr>
<tr>
<td>High school graduate/GED</td>
<td>54.56</td>
</tr>
<tr>
<td>Some college</td>
<td>9.52</td>
</tr>
<tr>
<td>College degree</td>
<td>143.02</td>
</tr>
<tr>
<td>Graduate work</td>
<td>45.21</td>
</tr>
</tbody>
</table>

a N reflects sample with non-missing data for all variables included in the model.
b Rescaled to range from 0 to 100.
c Female as referent.
d White as referent.
e <High school graduate as referent.
percentile of education Gini distribution the relative odds of marijuana use were 1.65 (95% CI = 1.12–2.41) and in neighborhoods in the 75th percentile the relative odds were 3.49 (95% CI = 1.33–9.15).

Fig. 2 shows the relative odds of marijuana use for different levels of education Gini among neighborhoods with low different levels of mean education (ranging from the 25th to the 75th percentile in the data), as predicted from the final models including the statistically significant interaction terms for these two variables. Among neighborhoods with low mean education, relative odds of marijuana use for individuals living in neighborhoods in the 75th percentile of education distribution were 3.39 (compared to the 25th percentile of education Gini). For mid education neighborhoods, relative odds of marijuana use were 2.66 and for high education neighborhoods, relative odds were 2.09, also comparing neighborhoods in the 75th percentile of education Gini to the 25th percentile.

In multilevel linear regression models assessing the relations between education inequality and frequency of cigarette use and alcohol use in the past 30 days among those who used each substance (Table 3), neither neighborhood mean education ($\beta = 0.32$) nor education inequality ($\beta = 0.63$) were significantly associated with the quantity of cigarettes smoked. However, controlling for neighborhood education level, income distribution, and individual-level characteristics, living in a neighborhood with high education inequality was associated with fewer drinks consumed in the past month ($\beta = -2.41; p = 0.03$) among those who reported any alcohol use in the past month. In contrast, living in a neighborhood with high income inequality was associated with more drinks consumed in the past month ($\beta = 1.55; p = 0.004$). Further adjustment to account for exposure to the 11 September 2001 terrorist attacks did not appreciably change any of the parameter estimates in the models of interest in this analysis (data not shown).

4. Discussion

Neighborhood-level distribution of education attainment was significantly associated with alcohol consumption and marijuana use but not cigarette use in this analysis. In multilevel models adjusting for individual and neighborhood education and other covariates, higher education inequality was associated with higher prevalence of drinking. The odds of alcohol use were 1.70 times higher in neighborhoods in the 75th percentile of education distribution compared to neighborhoods in the 25th percentile; comparable relative odds for marijuana use were 3.49. However, among those who drink, living in neighborhoods of higher education inequality was associated with fewer drinks consumed. There was a significant interaction between neighborhood education distribution and neighborhood mean education for marijuana use such that education distribution was more strongly associated with marijuana use in neighborhoods with low mean education, with higher use in neighborhoods with more unequally distributed education levels.

There are several reasons why contextual variables in general, and distribution of education at the neighborhood level in particular, may be associated with substance use. First, neighborhood characteristics may increase levels of psychological distress (Aleshesnel and Sucoff, 1996) and drug use may occur for the relief of states of stress (Rhodes and Jason, 1990; Lindenberg et al., 1994). There is ample research demonstrating that stressful life events occur with greater frequency in neighborhoods with low levels of income and education (Fang et al., 1998) and that substance use may be a way to cope with these events (Boardman et al., 2001). Second, adverse neighborhood conditions may undermine individuals’ psychological coping resources and make use of substances more likely (Wilson, 1996). Third, it is possible that neighborhood disadvantage decreases social resources available to individuals, resulting in more limited assistance in coping with daily stresses, and fewer resources to overcome substance use once initiated.

Fourth, drug-related behaviors may be related to neighborhood social norms through mechanisms unique to different neighborhoods and population groups (Linsky et al., 1986; Kaplan et al., 2001). Fifth, differential neighborhood availability of substances may be directly associated with different levels of drug use independent of individual-level factors. For example, it has been shown that alcohol outlet density is related to higher levels of alcohol consumption (Scribner et al., 2000). Targeted advertising in particular neighborhoods may increase awareness and desirability of substances (Donovan et al., 2002).

The first three postulated mechanisms to explain relations between neighborhood characteristics and substance use suggest that more disadvantaged neighborhoods (e.g., those with lower educational attainment and lower levels of education inequality) would be expected to have higher prevalence and frequency of substance use due to stress, stressful life events, or diminished coping or other resources. The last two postulated mechanisms leave open the question of how neighborhood disadvantage might relate to substance use, as substance use norms and availability of substances may not correlate with disadvantage. For example, in NYC socioeconomically disadvantaged neighborhoods have more liquor stores, but advantaged neighborhoods have more bars and restaurants that serve liquor. Overall, socioeconomically advantaged neighborhoods have more sources of liquor than disadvantaged neighborhoods (New York State Division of Alcoholic Beverage Control/State Liquor Authority, 2002). The complex relations observed in this study between education distribution in neighborhoods and current use of different substances suggest that disadvantage-related mechanisms as well as substance use norm and substance availability mechanisms may be operating on different aspects of substance use. We showed that education distribution is associated with alcohol use in multilevel models such that neighborhoods with higher mean education and higher education inequality had a higher prevalence of alcohol use and of marijuana use, suggesting a role of substance use norms or substance availability mechanisms. However, we also showed that lower education inequality is associated with a higher number of drinks consumed among those who drink, suggesting roles for both the disadvantage and substance use norms-related mechanisms. Therefore, in neighborhoods with narrow distributions of education, those who drink may consume more heavily due to stress levels or lack of coping resources. Conversely, in neighborhoods where at least
some people have higher education, awareness about the conse-
quences of heavy alcohol drinking may moderate the amount of alco-
hol consumed by those who drink.

In general, the broader range of educational attainment within
neighborhoods of high education inequality may result in the avai-
lability of salutary materials and human resources other-
wise absent in neighborhoods with low education inequality.
For example, the presence of a few individuals with higher edu-
cation may benefit all residents of the neighborhood by bringing a
variety of advantageous resources ranging from social and
health services to fresh food markets. Furthermore, as persons
with high educational attainment are likely to have had access
to public goods during their education, they may actively con-
tribute to social welfare and cohesion, which would improve
the general population health. Our results extend previous work
in suggesting that the influence of education distribution and
income distribution differ with respect to health and health be-
haviors (Kawachi and Kennedy, 1999; Kaplan et al., 1996;
Galea and Ahern, 2005). Specifically, while maldistribution of
income was associated with a higher number of drinks consumed
among those who drink, maldistribution of education was asso-
ciated with a lower number of drinks consumed among those
who drink.

We did not find any associations between education distrib-
ution and smoking. This furthers the debate in the literature
between work that has shown that contextual factors are asso-
ciated with cigarette use (Diez-Roux et al., 1997; Kleinschmidt
et al., 1995; Jones and Duncan, 1985; Reijneveld, 1998; Ecob
and MacIntyre, 2000) and work that has not (Tseng et al., 2001).
These differences may be explained both by the methodologic
 Differences (including contextual units of analysis) between
these studies and by the potential presence of multiple mech-
anisms, each relevant in different contexts, that determine
 substance use.

The literature on contextual determinants of illicit drug use
is sparse (Galea et al., 2003c). We are aware of only one other
empirical study that has assessed contextual determinants of
alcohol, cigarette, and marijuana use in the same population
sample (Galea et al., in press). This work also suggests similar-
ities between the determinants of alcohol and marijuana use in
contrast to cigarette use. It is plausible that since cigarettes have
 a greater potential for dependence than either alcohol or mari-
jjuana (Kandel et al., 1997), current cigarette use is determined
by earlier characteristics of the life-course, while contempora-
neous contextual variables may have an effect on alcohol and
marijuana use. This observation, if replicated, has substantial
implications for prevention efforts and merits further research.
We note that we also found statistical interaction suggesting that
education inequality is more strongly associated with marijuana
use in neighborhoods with low mean education, with higher use
in neighborhoods with more unequally distributed education lev-
els. In contrast, there was no statistically significant interaction
in the models predicting alcohol use. This further hints at the
complexity of the role that neighborhood-level social domains
may play in influencing risk of substance use. It is plausible
that the mechanistic relations underlying the education Gini and
marijuana use relation are different than those underlying the
relation between education Gini and alcohol use and that the
key mechanisms that explain the results documented here oper-
ate differently in the determination of marijuana use at different
absolute average level of neighborhood education.

There are several limitations to this study. We used data from
a study of residents of New York City in the aftermath of the 11
September attacks. It is possible that the relations observed here
are particular to a period of heightened concern due to a national
disaster and are not generalizable to other contexts. However,
this data was collected more than 6 months after the attacks
and there is no evidence that the increase in substances used in
New York City after this disaster was differential across geo-
graphic areas suggesting that this concern is unlikely to affect
the observations documented here. To further control for the
possibility that the results documented here were influenced by
study circumstance, we adjusted for exposure to the 11 Septem-
ber 2001 terrorist attacks in all models considered here. The
absence of any, appreciable change in parameters of interest in
these models is reassuring that the findings documented here are
not attributable to the circumstances of this particularly study. By
studying intraurban differences in the largest city in the United
States, the generalizability to other smaller cities or non-urban
environments is potentially limited. Future analyses would have
to consider the role of education inequality, if any, in different
contexts and at different geographic levels. The survey data
used here were collected through telephone interviews, raising
the possibility of under-reporting of substances used. This is
again unlikely given the comparability of substance use docu-
mented here to national estimates and the growing evidence to
suggest that estimates obtained through telephone assessments
are valid when compared to in-person assessments (Midanik and
Greenfield, 2003; Nelson et al., 2003). Although we controlled
for the available relevant individual and neighborhood-level vari-
ables it is possible that residual cross-level confounding or
confounding by covariates not considered here could explain the
observed relations among neighborhood characteristics and sub-
stance use measures. Consistent with previous research (Marzuk
 et al., 1997), we used community districts as proxies for neigh-
borhoods in NYC. Definitions of relevant neighborhood units
is challenging and these units, while large, are probably more
meaningful analytic units than census tracts or zip codes, com-
mon units of analysis in the study of neighborhood-level effects.
However, it is important to note that all findings that assess con-
textual determinants need to be considered carefully with respect
to the contextual levels selected and ultimately depend on the
theoretic rationale as to why a particular neighborhood unit may
matter (Galea and Ahern, 2006). We suggest that the hypo-
thesized mechanisms that may explain an association between
education distribution and substance use involve the sharing of
human and social resources that would primarily manifest at
the small-area level. The observation that education inequality
is associated with substance use at the neighborhood level does
not preclude the possibility that education inequality may be an
important determinant at county, state, or national levels.

In this regard, we note that one of the central challenges
in work like the one carried out here is operationalizing
neighborhoods. Although operationalizing or defining the rel-
relevant group-level unit of analysis is straightforward in some cases such as institutional settings (e.g., schools) or political boundaries (e.g., states), it can be more challenging in other cases—especially with regard to neighborhoods. The bounding or definition of neighborhoods in neighborhood health research has garnered a lot of recent academic attention (O’Campo, 2003; Galea and Ahern, 2006). As with any exposure, bias can result if the neighborhood construct of interest (in this case education inequality) does not map onto the units chosen for operationalizing neighborhoods and their constructs. Defining neighborhoods (or any relevant unit) mainly threatens construct validity, although it threatens internal validity as well (e.g., non-differential exposure misclassification tends to bias the effect estimate towards the null). Therefore incorrectly bounding neighborhoods, or any spatial unit, may result in empirical problems.

The issue of how to define neighborhoods may also be a conceptual one. For example, how an outsider defines neighborhood boundaries may be different from how a resident him/herself does. Is also is likely that there is heterogeneity within a certain neighborhood as to how residents define their neighborhood and that the underlying construct of neighborhood varies for different people. If we are interested in social exposures that affect health, as we are in this paper, then resident perceptions or definitions of neighborhood may be more relevant for that examination than in the study of other factors such as, for example, formal policies which are operationalized according to administrative boundaries (Diez Roux, 2001). Also, some processes occurring in neighborhoods that we may hypothesize to affect health are not necessarily contained within any given spatial boundary (O’Campo, 2003). Although the challenges for operationalizing relevant neighborhoods are important to consider as we weigh choosing a level of analysis, this operationalization should not paralyze empirical investigation (Diez Roux, 2001). Since different phenomena may operate at different scales to affect health, multiple appropriate neighborhood units may be defined to accommodate inclusion of the multiple processes that comprehensively describe how a particular social process shapes health and behavior (O’Campo, 2003). Ultimately a multitude of different groups/units may be relevant for a specific research question (Diez Roux, 2004).

We also note that much as there is concern about generalizability of work such as this across different neighborhood units, there are further considerations about generalizing neighborhood research to different local and national contexts. There is burgeoning evidence that observations about group-level determination of health vary across countries (Lynch et al., 2004a) and future work that aims to replicate the observations we document in this paper may fruitfully consider both different ways of operationalizing neighborhoods and replication (or refutation) of these observations in different countries.

We used Census data from 2000 and it is difficult to know how well this information represents conditions of neighborhoods in NYC in 2002 and if any changes may account for some of the observed associations. Also, inferences about the patterns of marijuana use prevalence are limited by the relatively low prevalence of marijuana use. We also note that although we assess our observations with tests of statistical significance, these tests are predicated on significance levels set for each test alone. Type I error rates (false positives) apply to each test, not to them all taken together, and there remains the possibility that our observations are due to chance alone. Further replication of these observations in other studies would be needed to provide us with confidence that these observations are not accounted for by chance alone. Finally our data pertain strictly to substance use and inferences should not be extended from these observations to either substance abuse or dependence.

This study does not, in and of itself, offer much guidance about policies that can be implemented to reduce substance use and misuse in urban environments. As we note earlier in this paper, this is only the second paper of which we are aware that has explicitly considered how distribution of education at the neighborhood level may be associated with use of substances. The findings here both illustrate the potential contribution of contextual determinants to the use of substances and suggest the complexity of the relations between contextual factors and substance use behavior. Further inquiry is needed to investigate the mechanisms underlying the associations between education inequality and substance use. We suggest that such inquiry is worth pursuing. It has long been established in thinking about population health that change at the level of fundamental “upstream” determinants of health has the potential to influence population health more profoundly than intervention targeted at individual determinants of behavior, health, or disease. Considering neighborhood determinants of individual substance use risk may present important opportunities for intervention once further work has been conducted to clarify and explore the relations that are documented in this analysis.

Acknowledgements

Funded in part by National Institutes of Health grants DA 017642 and MH 066391. The authors would like to thank the anonymous reviewers and the editor for comments that substantially helped improve the manuscript.

References

Adler, N.E., Newman, K., 2002. Socioeconomic disparities in health: pathways and policies. Inequality in education, income, and occupation exacerbates the gaps between the health of “haves” and “have-nots”. Health Aff. (Millwood) 21, 60–76.


S. Galea et al. / Drug and Alcohol Dependence 90S (2007) S4–S15
