

Selective nonresponse bias in population-based survey estimates of drug use behaviors in the United States

Sean Esteban McCabe¹ · Brady T. West^{2,3}

Received: 27 April 2015 / Accepted: 5 September 2015 / Published online: 16 September 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract

Purpose There is a trend of decreasing response rates in population surveys, and selective nonresponse represents a major source of potential bias in population-based survey estimates of drug use behaviors, especially estimates based on longitudinal designs.

Methods This study compared baseline substance use behaviors among initial respondents who did respond ($n = 34,653$) and did not respond ($n = 8440$) to a 3-year follow-up interview in a prospective study of the general U.S. adult population. Differences in nonresponse rates were assessed as a function of past-year drug use behaviors both before and after adjustment for socio-demographic differences potentially associated with these behaviors, and the effects of interactions of the socio-demographic characteristics with the drug use behaviors were assessed in multivariate logistic regression models for response at the 3-year follow-up.

Results Weighted and unweighted nonresponse rates varied between alcohol users and users of other drugs such

as cocaine and marijuana, with rates of nonresponse being higher in the latter drug categories. There were also significant differences in nonresponse rates as a function of frequency of use and demographics. More specifically, being married tends to reduce the probability of non-response, while older age, being male, being Asian or Hispanic, and having lower education all substantially increase the probability of nonresponse at Wave 2, even after controlling for relevant covariates.

Conclusions This study provides the substance abuse field with a methodology that users of longitudinal data can apply to test the sensitivity of their inferences to assumptions about attrition patterns.

Keywords Nonresponse bias · Longitudinal · Population-based · Survey estimates · Drug use

Introduction

There is evidence that the profile of substance use behaviors and substance use disorders (SUDs) among individuals entering substance abuse treatment has changed dramatically in the past two decades according to the Treatment Episode Data Set, which collects data on admissions to U.S. substance abuse treatment facilities [1, 2]. More specifically, there has been a significant shift in the primary substances of abuse observed among those entering substance abuse treatment programs in the U.S. over the past two decades. The percentage of substance abuse treatment admissions reporting alcohol as the primary substance of abuse decreased from 57 % in 1993 to 41 % in 2010, while the percentage of substance abuse treatment admissions for marijuana, opiates, and stimulants as the primary substance increased from 22 % in 1993 to 47 % in 2010 [1, 2]. These

Electronic supplementary material The online version of this article (doi:10.1007/s00127-015-1122-2) contains supplementary material, which is available to authorized users.

✉ Sean Esteban McCabe
plius@umich.edu

¹ University of Michigan Institute for Research on Women and Gender, 204 S. State St., Ann Arbor, MI 48109-1290, USA

² Survey Research Center, University of Michigan Institute for Social Research, P.O. Box 1248, Ann Arbor, MI 48016-1248, USA

³ University of Michigan Center for Statistical Consultation and Research, 915 East Washington Street, Ann Arbor, MI 48109-1070, USA

shifts suggest the importance of obtaining accurate estimates of substance use behaviors and SUDs in the U.S. general population.

There is a general trend toward decreasing response rates in population surveys [3], and selective nonresponse represents a major source of potential bias in survey-based estimates of drug use behaviors. This is especially true for estimates of change in the prevalence of these behaviors and disorders based on longitudinal designs. Therefore, the degree to which different types of substance use behaviors and SUDs among survey respondents are representative of the U.S. general population has important implications for screening, prevention efforts, and substance abuse treatment programs. Previous population surveys have found evidence that alcohol use behaviors were significantly associated with attrition in cross-sectional and longitudinal designs [4–6]. For instance, Dawson and colleagues [4] recently examined whether alcohol consumption differed for respondents and nonrespondents in a longitudinal study. These authors found that the degree to which the prevalence of drinking was underestimated in the total U.S. adult population was 1.6 %, while the extent to which alcohol consumption was overestimated ranges from 1.7 to 2.4 % among drinkers, after correcting for their sociodemographic differences. Dawson and colleagues [4] concluded that the extent of nonresponse bias in estimates of alcohol use behaviors should not affect drinking guidelines and planning for prevention and treatment programs.

At least two other longitudinal investigations have found evidence that alcohol use was not significantly associated with attrition while marijuana and other drug use was an important source of nonresponse bias in longitudinal surveys of young adults in the U.S. military [7] and U.S. college students [8]. However, there have been no nationally representative longitudinal studies of the U.S. adult general population examining the potential nonresponse bias associated with substances other than alcohol. This gap in knowledge needs to be addressed, because non-college and non-military young adults tend to have much higher rates of illicit and nonmedical drug use than their college-attending and military counterparts [9–11]. Thus, investigations are needed to assess potential nonresponse bias associated with estimates of the prevalence of substance use behaviors and SUDs for substances other than alcohol based on prospective nationally representative samples, and especially those that take into account frequency of substance use behaviors. The main objective of this study is to examine whether substance use behaviors measured at the baseline wave of data collection in a nationally representative U.S. population survey are associated with the likelihood of re-interview 3 years later at a follow-up wave, and examine the sensitivity of inferences

regarding substance use behaviors to potential differential attrition.

Methods

This study used the 2001–2002 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC, Wave 1) and 2004–2005 NESARC (Wave 2) as the primary datasets to obtain information regarding substance use behaviors and SUDs among the general population in the United States. The target population for the 2001–2002 NESARC (Wave 1) was the civilian, noninstitutionalized population residing in the United States, 18 years of age and older. Wave 1 and 2 data were collected via face-to-face interviews conducted in respondents' households. The Wave 1 target population included persons living in households, military personnel living off base, and persons residing in certain group quarters: boarding or rooming houses, nontransient hotels, shelters, facilities for housing workers, college quarters, and group homes. The NESARC used a multistage stratified sample design, and had two sampling frames: one for housing units and one for group quarters. The United States Bureau of the Census trained interviewers using the Alcohol Use Disorder and Associated Disabilities Interview Schedule DSM-IV (AUDADIS-IV). This fully structured diagnostic interview was completed via face-to-face personal interviews. The response rate for Wave 1 was 81.0 %; the response rate among those eligible for Wave 2 was 86.7 %, resulting in a cumulative response rate of 70.2 % (the product of the response rates from Waves 1 and 2). More details about the NESARC sample design and data collection methods for both Waves 1 and 2 are available elsewhere [12, 13]. The United States Census Bureau and the United States Office of Budget and Management approved the NESARC research protocol. The University of Michigan Institutional Review Board approved the present study.

Sample

At Wave 1, the sample consisted of 43,093 adults 18 years of age and older, and represented a population that was 52 % women, 71 % White, 12 % Hispanic, 11 % African American, 4 % Asian, and 2 % Native American or another racial category. In addition, an estimated 20 % of the population lived in the Northeast, 35 % in the South, 23 % in the Midwest, and 22 % in the West. Approximately 12 % of the population were 18–24 years of age and 88 % were adults 25 years of age or older. The mean age of the Wave 1 sample (unweighted) was 46.4 (SD = 18.2), and the corresponding weighted estimate of the mean age of the target population was 45.2 (SE = 0.2).

Without adjusting the Wave 1 sampling weights for attrition, the respondents in both waves represented a population that was 53 % women, 73 % White, 10 % Hispanic, 11 % African–American, 4 % Asian, and 2 % Native American or another racial category, with a similar regional distribution. After applying weighting adjustments for attrition from Wave 1, the sample responding in both waves ($n = 34,653$) represented a population with the same gender, region, and race/ethnicity distributions from Wave 1, and an estimated mean age of 44.9 (SE = 0.2).

Measures

Demographic and background characteristics

Demographic and background characteristics collected at the Wave 1 interview included age, sex (male, female), race/ethnicity (white, black, Native American, Asian, Hispanic), educational level (less than high school, high school, some college or higher), employment status (full-time, part-time, not working), income (less than \$19,999, \$20,000–\$34,999, \$35,000–\$69,999, \$70,000 or higher), United States Census geographical region (Northeast, Midwest, South, and West), metropolitan statistical area (MSA) (central city in MSA, not in central city in MSA, not in MSA), and relationship status (married/cohabiting, widowed/separated/divorced, never married).

Past-year substance use behaviors

Alcohol use was assessed at each wave by asking respondents if during the last 12 months they had at least one drink of any kind of alcohol (e.g., beer, champagne, coolers, wine, and liquor), not counting small tastes or sips. Other substance use behaviors were assessed at each wave by asking respondents about use of a variety of medications and other drugs without a doctor's prescription. Respondents were asked about use of medications in the last 12 months that were *not* prescribed to them by a doctor or that were used in a manner not intended by the prescribing clinician (e.g., more often than prescribed, longer than prescribed, or for a reason other than prescribed, such as to get high). Substances included marijuana, cocaine or crack, heroin, hallucinogens, inhalants or solvents, prescription opioids, sedatives, stimulants, and tranquilizers or anti-anxiety drugs.

Past-year substance use disorders

Substance use disorders at each wave were based on DSM-IV criteria from the AUDADIS-IV, which contains symptom questions used to operationalize DSM-IV abuse and dependence separately for all ten substances (alcohol,

marijuana, cocaine, heroin, hallucinogens, inhalants, prescription opioids, sedatives, stimulants, and tranquilizers). A past-year diagnosis of abuse requires the absence of a diagnosis of dependence and at least one of four DSM-IV abuse criteria in the 12 months preceding the interview. A past-year dependence diagnosis is based on presence of at least three of the seven DSM-IV dependence criteria. The diagnoses of abuse and dependence were combined to create binary substance use disorder outcomes. Reliability and validity of the DSM-IV AUDADIS-IV substance use disorder diagnoses have been documented in numerous psychometric studies [14–23], with test–retest reliability ranging from good to excellent (0.70–0.91).

Data analysis

This study compared baseline substance use behaviors among initial respondents who did respond ($n = 34,653$) and did not respond ($n = 8440$) to a 3-year follow-up interview in a prospective study of the general U.S. adult population [12]. Differences in nonresponse rates were assessed as a function of past-year substance use behaviors both before and after adjustment for socio-demographic differences potentially associated with these behaviors, and the effects of interactions of the substance use behaviors with the socio-demographic characteristics were assessed in multivariate logistic regression models for a nonresponse indicator at the 3-year follow-up.

We considered three design-based logistic regression models for the indicator of nonresponse at Wave 2, accounting for the complex sample design features of the NESARC (i.e., using the Wave 1 weights to estimate the model parameters, and the NESARC stratum and cluster codes to estimate sampling variability). This approach enabled us to make representative statements about the expected response behaviors of the larger NESARC population. The initial model only included socio-demographic covariates (age, sex, race/ethnicity, education, and marital status). The second model included categorical variables measuring the self-reported frequency of selected past-year substance use behaviors from Wave 1 (alcohol, marijuana, cocaine, stimulants, opioids, hallucinogens, sedatives, and tranquilizers), in addition to binary indicators of any past-year use of inhalants or heroin. The third model then examined two-way interactions between the substance use measures and the socio-demographic predictor variables; only significant two-way interactions according to a design-adjusted Wald test [24] were retained in the third model. Each model was tested for goodness of fit using design-based methods described by Archer and colleagues [25].

Finally, to examine the possibility that there may be unique subclasses of substance users at Wave 1 based on

these frequency measures and indicators, we performed an exploratory latent class analysis (LCA) using all of these categorical substance use measures at Wave 1, and identified the latent class solution with the optimal fit based on likelihood ratio tests and the BIC criterion. We then grouped Wave 1 respondents into subclasses based on the latent class with the highest posterior predicted probability for each individual, described the resulting subclasses, and compared these subclasses in terms of nonresponse prevalence at Wave 2.

Results

Wave 2 nonresponse rates by Wave 1 substance use behaviors and gender

Table 1 presents both weighted and unweighted nonresponse rates at NESARC Wave 2 for different subgroups of individuals defined by past-year drug use behaviors at Wave 1. Unweighted nonresponse rates varied from 17.45 % (for past-year users of alcohol) to 27.65 % (for past-year cocaine users). While the nonresponse rate for cocaine users is based on a smaller subsample of individuals ($n = 217$), this does suggest that slight negative biases may be present for Wave 2 estimates describing correlates of cocaine use if Wave 1 cocaine use is not used in computing nonresponse adjustments. The weighted estimates, representing expected nonresponse rates if everyone in the target population had been included in the sample, tracked the unweighted response rates fairly closely, ranging from 13.56 % (for past-year heroin users at Wave 1, with $n = 11$ only) to 24.08 % (for past-year cocaine users at Wave 1).

Table 1 also presents design-based comparisons of nonresponse rates for males and females, among each subgroup of users defined by Wave 1 behaviors. Significant differences were found between males and females among Wave 1 users of: (1) any substance ($p < 0.001$, with males having a higher nonresponse rate at Wave 2), (2) alcohol ($p < 0.001$, with males having a higher nonresponse rate at Wave 2), and (3) two or more substances ($p < 0.05$, with males having a higher nonresponse rate at Wave 2).

We replicated the Table 1 analyses for different subgroups defined by substance use disorders (SUDs) in the past year at Wave 1 (results not shown). In general, we found evidence of higher nonresponse rates for subgroups of individuals having various SUDs at Wave 1. This was especially true for those with a cocaine use disorder ($n = 97$; weighted nonresponse rate = 26.38 %, unweighted nonresponse rate = 28.87 %), those with a tranquilizer use disorder ($n = 51$; nonresponse rate = 27.45 % regardless of whether weights were used), and those with SUDs associated with two or more

substances ($n = 462$; weighted nonresponse rate = 23.16 %, unweighted response rate = 23.22 %). In addition, we found evidence of males once again having higher nonresponse rates among individuals with any SUDs at Wave 1, individuals with alcohol use disorders, and individuals with marijuana use disorders, suggesting that longitudinal estimates of change in drug use behaviors for these subgroups may be biased in the direction of the change for females.

Wave 2 nonresponse rates by frequency of Wave 1 substance use behaviors

Table 2 presents weighted and unweighted estimates of nonresponse rates at NESARC Wave 2 for different subgroups of individuals defined by frequency of substance use behaviors in the past year at Wave 1. Substances with notable associations between frequency of use at Wave 1 and nonresponse at Wave 2 included alcohol, where daily drinkers were found to have the highest nonresponse rate; cocaine, where daily and annual users at Wave 1 had relatively higher rates of nonresponse at Wave 2; and stimulants and tranquilizers, where after incorporating the NESARC weights into the estimation, daily and weekly users had a higher nonresponse rates at Wave 2. The general patterns of differences in nonresponse rates were not particularly sensitive to the use of the NESARC weights in estimation, but individual estimates of nonresponse rates did in some cases vary substantially with and without the use of weights in estimation (e.g., cocaine and stimulant use). This suggests that the NESARC weights may be informative about response behavior for selected subgroups, and should thus be used when estimating nonresponse rates for the larger populations represented by these subgroups.

Logistic regression models predicting nonresponse at Wave 2

Table 3 presents results from the multivariate logistic regression analyses, where an indicator of nonresponse at Wave 2 was predicted by various socio-demographics and indicators of different drug use behaviors at Wave 1. Notably, several two-way interactions between socio-demographic covariates and the past-year substance use indicators at Wave 1 emerged as significant predictors, so we focus our discussion on Model 3, which was found to have an improved fit relative to the other models according to the smaller value of the design-based goodness-of-fit test statistic proposed by Archer and colleagues [25].

In general, we found significant two-way interactions involving age, race/ethnicity, education, and alcohol use at Wave 1. Among individuals who did not use alcohol at

Table 1 Nonresponse rates at 3-year follow-up, by baseline substance use behaviors in the U.S.

Baseline past-year substance use behaviors	Overall unweighted Nonresponse rate (weighted rate)	Female (%)	Male (%)	Gender differences: Rao-Scott F-statistic, <i>p</i> value
Any substance				
Use (<i>n</i> = 27221)	17.48 %* (16.59 %)*	14.84	18.17	37.85 (1, 65), <i>p</i> < 0.001
No use (<i>n</i> = 15837)	23.17 % (22.25 %)	21.09	24.07	13.63 (1, 65), <i>p</i> < 0.001
Alcohol				
Use (<i>n</i> = 26946)	17.45 %* (16.58 %)*	14.80	18.18	39.34 (1, 65), <i>p</i> < 0.001
No use (<i>n</i> = 16116)	23.12 % (22.16 %)	21.05	23.88	12.87 (1, 65), <i>p</i> < 0.001
Marijuana				
Use (<i>n</i> = 1603)	20.21 % (19.19 %)	17.22	20.19	1.51 (1, 58), <i>p</i> = 0.225
No use (<i>n</i> = 41479)	19.56 % (18.49 %)	17.35	19.78	29.85 (1, 65), <i>p</i> < 0.001
Cocaine				
Use (<i>n</i> = 217)	27.65 %* (24.08 %)	26.86	22.83	0.28 (1, 32), <i>p</i> = 0.601
No use (<i>n</i> = 42874)	19.54 % (18.48 %)	17.30	19.77	31.46 (1, 65), <i>p</i> < 0.001
Stimulants				
Use (<i>n</i> = 186)	18.82 % (19.60 %)	21.25	18.32	0.15 (1, 32), <i>p</i> = 0.698
No use (<i>n</i> = 42906)	19.59 % (18.51 %)	17.32	19.81	33.03 (1, 65), <i>p</i> < 0.001
Prescription opioids				
Use (<i>n</i> = 686)	20.26 % (18.81 %)	19.37	18.38	0.09 (1, 56), <i>p</i> = 0.771
No use (<i>n</i> = 42403)	19.58 % (18.51 %)	17.31	19.83	32.65 (1, 65), <i>p</i> < 0.001
Hallucinogen				
Use (<i>n</i> = 194)	23.71 % (18.90 %)	26.09	15.86	2.97 (1, 25), <i>p</i> = 0.097
No use (<i>n</i> = 42897)	19.57 % (18.52 %)	17.31	19.83	33.24 (1, 65), <i>p</i> < 0.001
Sedatives				
Use (<i>n</i> = 490)	18.98 % (19.56 %)	17.68	21.13	0.71 (1, 49), <i>p</i> = 0.405
No use (<i>n</i> = 42593)	19.60 % (18.51 %)	17.34	19.79	32.33 (1, 65), <i>p</i> < 0.001
Tranquilizers				
Use (<i>n</i> = 345)	20.87 % (17.04 %)	15.94	17.72	0.15 (1, 40), <i>p</i> = 0.701
No use (<i>n</i> = 42745)	19.57 % (18.53 %)	17.34	19.82	32.55 (1, 65), <i>p</i> < 0.001
Heroin				
Use (<i>n</i> = 11)	18.18 % (13.56 %)	0.00	23.27	1.12 (1, 1), <i>p</i> = 0.482
No use (<i>n</i> = 43082)	19.59 % (18.52 %)	17.34	19.80	32.01 (1, 65), <i>p</i> < 0.001
Inhalants				
Use (<i>n</i> = 39)	23.08 % (19.51 %)	27.90	17.43	0.32 (1, 8), <i>p</i> = 0.587
No use (<i>n</i> = 43054)	19.58 % (18.52 %)	17.33	19.80	32.51 (1, 65), <i>p</i> < 0.001
Illicit drugs other than marijuana				
Use (<i>n</i> = 1408)	19.39 % (17.52 %)	16.23	18.55	0.86 (1, 61), <i>p</i> = 0.357
No use (<i>n</i> = 41672)	19.60 % (18.56 %)	17.37	19.86	31.41 (1, 65), <i>p</i> < 0.001
Two or more drugs				
Use (<i>n</i> = 2241)	19.54 % (18.80 %)	16.04	20.55	4.31 (1, 62), <i>p</i> = 0.042
No use (<i>n</i> = 40852)	19.59 % (18.50 %)	17.40	19.74	27.45 (1, 65), <i>p</i> < 0.001
Total sample (<i>n</i> = 43,093)	19.59 % (18.52 %)	17.34	19.80	32.26 (1, 65), <i>p</i> < 0.001

All weighted estimates and Rao-Scott Chi square tests use design information from Wave 1. Female and Male estimates are weighted

* Indicates significant (*p* < 0.01) difference in nonresponse rate compared to the no-use category, either based on a Pearson Chi square test (ignoring sample design features) for unweighted rates, or a Rao-Scott Chi square test (for weighted rates)

Wave 1, older individuals, Asians, Hispanics/Latinos, and those with less-than-high school education tended to have higher rates of nonresponse (holding other covariates fixed). The age effect was significantly reduced for

different types of users, becoming essentially zero for all four different groups of users defined by frequency of use. A negligible difference in nonresponse rates between black and white persons among non-users became substantially

Table 2 Nonresponse rates at 3-year follow-up, by baseline frequency of past-year substance use behaviors in the U.S.

Baseline frequency of past-year substance use behaviors	Overall unweighted Nonresponse rate (weighted rate)	Chi square (df)/ Rao-Scott F-statistic, <i>p</i> value
Alcohol		
Annually (<i>n</i> = 8652)	16.6 % (15.7 %)	Unweighted:
Monthly (<i>n</i> = 6220)	16.5 % (15.7 %)	$X^2(3) = 16.6, p < 0.01$
Weekly (<i>n</i> = 8794)	18.1 % (17.3 %)	Weighted:
Daily (<i>n</i> = 3075)	19.2 % (18.3 %)	$F(2.8, 181.4) = 3.6, p = 0.02$
Marijuana		
Annually (<i>n</i> = 580)	17.2 % (16.5 %)	Unweighted:
Monthly (<i>n</i> = 340)	21.5 % (21.3 %)	$X^2(3) = 5.3, p = 0.15$
Weekly (<i>n</i> = 323)	21.7 % (20.8 %)	Weighted:
Daily (<i>n</i> = 348)	22.7 % (20.6 %)	$F(2.8, 121.1) = 1.0, p = 0.40$
Cocaine		
Annually (<i>n</i> = 126)	27.8 % (24.2 %)	Unweighted:
Monthly (<i>n</i> = 31)	16.1 % (11.9 %)	$X^2(3) = 3.6, p = 0.30$
Weekly (<i>n</i> = 41)	26.8 % (17.7 %)	Weighted:
Daily (<i>n</i> = 17)	41.2 % (53.0 %)	$F(2.9, 93.2) = 3.3, p = 0.02$
Stimulants		
Annually (<i>n</i> = 82)	17.1 % (14.1 %)	Unweighted:
Monthly (<i>n</i> = 48)	16.7 % (14.1 %)	$X^2(3) = 3.8, p = 0.28$
Weekly (<i>n</i> = 27)	14.8 % (20.3 %)	Weighted:
Daily (<i>n</i> = 28)	32.1 % (40.3 %)	$F(2.9, 93.2) = 2.8, p < 0.05$
Prescription opioids		
Annually (<i>n</i> = 328)	21.0 % (18.5 %)	Unweighted:
Monthly (<i>n</i> = 152)	20.4 % (19.2 %)	$X^2(3) = 0.3, p = 0.96$
Weekly (<i>n</i> = 85)	18.8 % (16.6 %)	Weighted:
Daily (<i>n</i> = 109)	19.3 % (21.9 %)	$F(2.9, 72.5) = 0.2, p = 0.88$
Hallucinogen		
Annually (<i>n</i> = 160)	21.3 % (18.4 %)	Unweighted:
Monthly (<i>n</i> = 22)	40.9 % (23.3 %)	$X^2(3) = 4.5, p = 0.21$
Weekly (<i>n</i> = 11)	27.3 % (14.0 %)	Weighted:
Daily (<i>n</i> = 1)	0.0 % (0.0 %)	$F(2.8, 69.2) = 0.3, p = 0.83$
Sedatives		
Annually (<i>n</i> = 231)	17.3 % (18.5 %)	Unweighted:
Monthly (<i>n</i> = 91)	20.9 % (19.0 %)	$X^2(3) = 2.4, p = 0.50$
Weekly (<i>n</i> = 86)	17.4 % (19.7 %)	Weighted:
Daily (<i>n</i> = 77)	24.7 % (25.5 %)	$F(2.9, 143.7) = 0.4, p = 0.76$
Tranquilizers		
Annually (<i>n</i> = 197)	17.8 % (13.4 %)	Unweighted:
Monthly (<i>n</i> = 57)	24.6 % (16.4 %)	$X^2(3) = 3.0, p = 0.40$
Weekly (<i>n</i> = 44)	25.0 % (28.0 %)	Weighted:
Daily (<i>n</i> = 45)	26.7 % (29.6 %)	$F(2.8, 112.8) = 2.6, p = 0.06$

The sample sizes for heroin and inhalants were too small to examine frequency

larger for annual users, where blacks had a substantially higher nonresponse rate. In addition, the already positive effects of being Asian and Hispanic/Latino on nonresponse rates became even stronger among monthly users, suggesting that more frequent alcohol users in these demographic groups have an increased risk of being lost over

time in longitudinal data collections. Finally, a non-significant negative effect of having more than High School education on nonresponse among non-users becomes even stronger and significant among annual, monthly, and weekly users of alcohol, suggesting that past-year users of alcohol with more than High School education tend to have

Table 3 Estimated coefficients in design-based logistic regression models predicting nonresponse at NESARC Wave 2 as a function of demographics and Wave 1 substance use behaviors ($n = 43,030$ with non-missing data on all variables, for all models)

Predictor variables	Model 1: demographic characteristics Estimate (95 % CI)	Model 2: demographic characteristics + Wave 1 substance use behaviors Estimate (95 % CI)	Model 3: demographic characteristics + Wave 1 substance use behaviors + interactions Estimate (95 % CI)
Intercept	-1.398 (-1.479, -1.317)***	-1.258 (-1.350, -1.165)***	-1.352 (-1.473, -1.230)***
Age (centered at 43)	0.009 (0.006, 0.011)***	0.007 (0.005, 0.010)***	0.012 (0.009, 0.016)***
Sex			
Male	0.213 (0.157, 0.269)***	0.229 (0.170, 0.288)***	0.223 (0.163, 0.283)***
Female	Ref	Ref	Ref
Race/ethnicity			
White NH	Ref	Ref	Ref
Black NH	0.197 (0.100, 0.294)***	0.158 (0.058, 0.259)**	0.033 (-0.097, 0.164)
Amer. Indian NH	-0.070 (-0.318, 0.178)	-0.089 (-0.339, 0.161)	-0.210 (-0.631, 0.211)
Asian NH	0.715 (0.550, 0.880)***	0.668 (0.500, 0.836)***	0.573 (0.328, 0.819)***
Hispanic/Latino	0.553 (0.446, 0.660)***	0.532 (0.424, 0.640)***	0.423 (0.265, 0.580)***
Education			
Less than high school	0.238 (0.136, 0.340)***	0.215 (0.113, 0.317)***	0.249 (0.112, 0.387)**
High school	Ref	Ref	Ref
More than high school	-0.223 (-0.302, -0.144)***	-0.204 (-0.284, -0.125)***	-0.273 (-0.138, 0.084)
Marital status			
Married	-0.479 (-0.538, -0.421)***	-0.475 (-0.533, -0.417)***	-0.451 (-0.509, -0.394)***
Not married	Ref	Ref	Ref
Wave 1 past-year substance use behaviors			
Alcohol			
No use			
Annual use			
Monthly use			
Weekly use			
Daily use			
Marijuana			
No use			
Annual use			
Monthly use			
Weekly use			
Daily use			
Cocaine			
No use			
Ref			

Table 3 continued

Predictor variables	Model 1: demographic characteristics Estimate (95 % CI)	Model 2: demographic characteristics + Wave 1 substance use behaviors Estimate (95 % CI)	Model 3: demographic characteristics + Wave 1 substance use behaviors + interactions Estimate (95 % CI)
Annual use	0.288 (-0.242, 0.819)	0.272 (-0.266, 0.809)	
Monthly use	-0.884 (-2.063, 0.296)	-0.846 (-2.050, 0.359)	
Weekly use	-0.184 (-1.185, 0.697)	-0.241 (-1.171, 0.689)	
Daily use	1.395 (0.122, 2.668)*	1.446 (0.206, 2.685)*	
Stimulants			
No use	Ref	Ref	Ref
Annual use	-0.459 (-1.195, 0.276)	-0.477 (-1.209, 0.254)	
Monthly use	-0.214 (-1.166, 0.739)	-0.235 (-1.189, 0.719)	
Weekly use	-0.048 (-1.080, 0.984)	-0.136 (-1.125, 0.853)	
Daily use	1.097 (-0.005, 2.198)	1.122 (0.004, 2.240)*	
Opioids			
No use	Ref	Ref	Ref
Annual use	0.101 (-0.235, 0.437)	0.079 (-0.254, 0.411)	
Monthly use	-0.047 (-0.603, 0.509)	-0.059 (-0.625, 0.507)	
Weekly use	-0.191 (-0.895, 0.512)	-0.187 (-0.878, 0.503)	
Daily use	-0.070 (-0.691, 0.551)	-0.057 (-0.668, 0.554)	
Hallucinogens			
No use	Ref	Ref	Ref
Annual use	-0.875 (-0.587, 0.412)	-0.162 (-0.675, 0.351)	
Monthly use	0.620 (-0.442, 1.683)	0.564 (-0.499, 1.628)	
Weekly use	-0.186 (-1.860, 1.488)	-0.304 (-2.001, 1.393)	
Daily use	-	-	-
Sedatives			
No use	Ref	Ref	Ref
Annual use	0.017 (-0.406, 0.440)	0.018 (-0.404, 0.440)	
Monthly use	0.061 (-0.490, 0.613)	0.050 (-0.501, 0.601)	
Weekly use	-0.057 (-0.693, 0.579)	-0.048 (-0.683, 0.587)	
Daily use	0.206 (-0.404, 0.817)	0.217 (-0.376, 0.810)	
Tranquilizers			
No use	Ref	Ref	Ref
Annual use	-0.408 (-0.926, 0.110)	-0.421 (-0.948, 0.107)	
Monthly use	-0.210 (-0.987, 0.567)	-0.183 (-0.964, 0.599)	
Weekly use	0.536 (-0.312, 1.383)	0.574 (-0.262, 1.409)	
Daily use	0.207 (-0.462, 0.875)	0.212 (-0.443, 0.867)	

Table 3 continued

Predictor variables	Model 1: demographic characteristics Estimate (95 % CI)	Model 2: demographic characteristics + Wave 1 substance use behaviors Estimate (95 % CI)	Model 3: demographic characteristics + Wave 1 substance use behaviors + interactions Estimate (95 % CI)
Heroin		-0.910 (-3.497, 1.677)	-0.939 (-3.520, 1.642)
Inhalants		-0.0001 (-0.982, 0.982)	-0.0002 (-0.979, 0.980)
<i>Interaction terms</i>			
Age × alcohol			Ref
Age × no use			-0.009 (-0.014, -0.004)***
Age × annual			-0.010 (-0.017, -0.004)**
Age × monthly			-0.013 (-0.019, -0.008)***
Age × weekly			-0.009 (-0.017, -0.001)*
Age × daily			
Race × alcohol			
White NH × no use			Ref
Black × annual			0.324 (0.111, 0.538)**
Asian × monthly			0.512 (0.131, 0.892)**
Hispanic × annual			0.262 (0.029, 0.495)*
Hispanic × monthly			0.385 (0.124, 0.646)**
Education × alcohol			
More than HS × no use			Ref
More than HS × annual			-0.282 (-0.476, -0.088)**
More than HS × monthly			-0.293 (-0.514, -0.073)*
More than HS × weekly			-0.298 (-0.511, -0.085)**
A–L goodness of fit test	F(9,57) = 1.50, p = 0.17	F(9,57) = 1.00, p = 0.45	F(9,57) = 0.97, p = 0.47

Non-significant interaction coefficients were not included in the table for ease of interpretation purposes

Ref reference group, – empty cell

*** p < 0.001; ** p < 0.01; * p < 0.05

an *increased* probability of responding relative to other more frequent users with less education. These effects are illustrated via plots of marginal predicted probabilities of nonresponse in Fig. 1 (see supplementary files online).

Predicted probabilities of nonresponse at 3-year follow-up, two-way interactions (Fig. 1, Note: Supplemental files for online publication only).

In addition to the interaction effects found in Model 3, we also find very strong main effects of sex and marital status, where being married tends to reduce the probability of nonresponse at Wave 2, while being male substantially *increases* the probability of nonresponse at Wave 2. We also find that daily users of cocaine and prescription stimulants have substantially increased odds of not responding at Wave 2, holding all other covariates in Model 3 fixed. In general, these results supported initial design-based tests of bivariate associations between the different socio-demographic predictors and the indicator of nonresponse at Wave 2 (results not shown) and the tests of association presented in Table 2.

Latent class analysis results

The optimal latent class solution (Entropy = 0.935) identified four classes (see Table 4; Fig. 2, Note: Supplemental files for online publication only), which could be described as follows:

- Class 1 (estimated percentage in population: 4.25 %) includes less severe “polydrug” users, who experiment with multiple substances but not at high rates.
- Class 2 (estimated percentage: 0.86 %) includes users of multiple nonmedical prescription drugs.
- Class 3 (estimated percentage: 94.36 %) includes users who drink occasionally but do not tend to use other substances.
- Class 4 (estimated percentage: 0.54 %) is the more severe group of “polydrug users” that includes heavy users of all drugs, and is at the high end of drinking frequency.

We re-fit the three logistic regression models including this four-category predicted latent class variable (instead of the aforementioned categorical measures of substance use at Wave 1), and found that these four classes did not differ significantly in terms of the probability of nonresponse at Wave 2. This was likely due to the fact that such a high proportion of the population was predicted to be in Class 3, limiting the power of our comparisons. Despite the potentially interesting presence of these four different classes, we cannot say that individuals belong to particular classes with certainty (despite the strong fit of the LCA model), and the individual measures of the different substance use behaviors at Wave 1 (especially past-year

alcohol use) appear to be more sensitive predictors of nonresponse at Wave 2.

Discussion

The findings of the present study suggest that nonresponse rates at follow-up waves in national longitudinal surveys of substance use vary considerably between alcohol users and users of other individual drug classes, such as cocaine and stimulants. As a result, these differences should be accounted for when computing nonresponse adjustments to reduce nonresponse bias, given that baseline measures of these drug use behaviors have significant associations with both Wave 2 response propensity and the same measures of interest at Wave 2 [26]. While previous national studies have found evidence that alcohol use behaviors were significantly associated with attrition in cross-sectional and longitudinal designs [4–6, 27], the findings of the present study extend prior longitudinal work on selected samples (e.g., college students, military personnel) that indicated other drug use behaviors were important sources of nonresponse bias in longitudinal surveys [7, 8].

The findings of the present study indicated notable associations between frequency of drug use behaviors at Wave 1 and nonresponse at Wave 2. Most notably, for alcohol, non-drinkers were found to have the highest nonresponse rates. Previous work has shown that alcohol abstainers have the highest attrition in population surveys relative to occasional drinkers and this could be associated with individual characteristics of abstainers that may influence ongoing participation in surveys [5, 6]. However, this same trend was not found for abstainers of other drug use behaviors, suggesting a unique effect for alcohol use. We found evidence of higher nonresponse rates for subgroups of individuals having various substance use behaviors (e.g., daily cocaine and stimulant use) and SUDs at Wave 1 (e.g. cocaine use disorders, tranquilizer use disorders, SUDs involving multiple drug classes). While the sample sizes of these subgroups defined by frequency of use and SUD were fairly small, accounting for frequency of use and SUDs in nonresponse adjustments may be especially important if one also desires subsequent analyses of frequency of use and SUD prevalence at later waves or time points.

The results also suggest that nonresponse rates in follow-up waves for specific subgroups of substance users may vary depending on gender. This implies that nonresponse adjustments only accounting for Wave 1 substance use behaviors and SUDs that do not incorporate gender and interactions between substance use behaviors and gender may be less effective than adjustments that incorporate these factors. Thus, our findings provide strong evidence of

the need to consider interactions between gender and indicators of substance use behaviors and SUDs when performing nonresponse adjustments.

The findings indicated that being married tends to reduce the probability of non-response, while higher age, being male, being Asian or Hispanic, and having lower education all substantially increase the probability of nonresponse at Wave 2. Furthermore, multivariate analyses provided evidence of interactions suggesting that some of these effects are exacerbated by frequency of alcohol use (e.g., being Asian or Hispanic). These findings are similar to previous cross-sectional and longitudinal studies and have important nonresponse bias implications for any drug use behaviors or disorders that are strongly related to these specific socio-demographic features, as the composition of the respondents is clearly changing from Wave 1 to Wave 2 [4, 27]. For example, Wave 2 respondents will tend to be younger, female, non-Asian, non-Hispanic, higher educated, and married. If individuals with these features tend to have unique drug use behaviors or changes in drug use behavior over time, weighted analyses that do not adjust for these demographic features may produce biased estimates of these behaviors and trends.

The conclusions of the present study apply primarily to a longitudinal setting and should not be used to make inferences about differences between respondents and non-respondents in Wave 1 of the NESARC. Past survey nonresponse studies have developed a theoretical model for nonresponse in later waves of a longitudinal survey, and found empirical support for a number of predictors of nonresponse in follow-up waves that have also been found to predict nonresponse in cross-sectional studies [28, 29]. Despite these previous studies, caution should be taken not to infer that the same predictors of nonresponse would be operating in the initial wave. While our study primarily focuses on the effects of various drug use behaviors on response in the second wave of a longitudinal study of substance abuse, our findings related to socio-demographic predictors are largely consistent with these previous studies.

The findings of the present study do suggest that the NESARC sampling weights did carry some information about nonresponse tendencies in particular subgroups, and this has important implications for interpretation of non-response rates (depending on whether weights were used) and nonresponse adjustments. Increasingly, the survey statistics literature suggests that base sampling weights should be used when computing nonresponse adjustments [30, 31]. Based on some of the slight differences found in the present study, future research should examine the sensitivity of adjusted estimates (and their standard errors) to the use of weights in computing the adjustments. Some of the differences that were observed in this study were for subgroups with fairly small sample sizes (e.g., cocaine and

stimulant users), especially when stratifying the results by frequency of use.

The findings from the present study extend our knowledge regarding the nonresponse bias associated with estimates of substance use behaviors and SUDs. This investigation is one of the few national and prospective studies to examine substance use behaviors and SUDs based on DSM-IV criteria, and future research is needed to examine potential nonresponse bias associated with SUDs based on DSM-5 criteria. The relatively large sample size of the NESARC allowed for the calculation of prevalence estimates for several drug classes. Finally, the nationally representative nature of the sample allows for generalization to the civilian non-institutionalized population, 18 years of age and older residing in the United States.

Despite these strengths, there are some limitations that must be weighed carefully when considering the implications of the findings. First, the findings from the NESARC should not be generalized to populations outside of the United States. Second, the sub-samples of some drug classes were relatively small (e.g., heroin and inhalants). Third, this study represented a secondary analysis and not all of the survey items were ideal for our purposes. For example, some of the most commonly misused prescription drugs were not specifically listed as examples in the Wave 1 NESARC survey, including hydrocodone (e.g., Vicodin[®]) and methylphenidate (e.g., Ritalin[®]) [11, 32]. Finally, although the NESARC includes persons formerly institutionalized who are now living in households and high risk group quarters such as nontransient public lodging, including hotels, shelters, and group homes, the exclusion of currently institutionalized individuals omits an important subgroup at increased risk for chronic DUD and relapse [33].

We also emphasize that this was a study of nonresponse bias in a longitudinal setting, where the baseline responses provide a convenient set of auxiliary variables for examining nonresponse bias, including measures on key survey variables. Assessments of nonresponse bias in cross-sectional national studies of substance use will tend to be much more difficult. In these cases, potential candidates for auxiliary variables that could be used to assess nonresponse bias include interviewer observations of features known to be correlated with substance use (e.g., drug papers, beer bottles, etc.) recorded for all sampled households [34], or survey-based estimates of drug use behaviors for large geographic regions (where, for example, if nonresponse rates are higher in regions known from previous studies to have higher rates of substance use, estimates may be biased). Identification of other effective auxiliary information for assessing potential nonresponse bias in cross-sectional substance use surveys is certainly an area where more work is needed.

In conclusion, the findings of the present study should provide the substance abuse field with a methodology that users of publicly available longitudinal data can apply to test the sensitivity of their inferences to assumptions about attrition patterns. In addition, the findings from these analyses provide researchers with empirical evidence in support of differential tracking strategies for individuals having different baseline substance use behaviors, frequencies and disorders in longitudinal studies of substance use, possibly enabling reductions in the risk of nonresponse bias in estimates based on longitudinal data through evidence-based study designs.

Acknowledgments The development of this manuscript was supported by research grants R01DA036541 and R01DA031160 from the National Institute on Drug Abuse, National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors would like to thank the respondents for their participation in the study and the anonymous reviewers and editorial staff for their helpful comments to a previous version of this article.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- Substance Abuse and Mental Health Services Administration (2006) Trends in substance abuse treatment admissions: 1993 and 2003. The DASIS report. Office of applied studies. <http://oas.samhsa.gov/2k6/TXtrends/TXtrends.cfm>. Accessed 26 January 2015
- Substance Abuse and Mental Health Services Administration (2012) Results from the 2011 national survey on drug use and health: summary of national findings, NSDUH series H-44, HHS publication no. (SMA) 12-4713. Substance Abuse and Mental Health Services Administration, Rockville
- Brick JM, Williams D (2013) Explaining rising nonresponse rates in cross-sectional surveys. *Ann Am Acad Pol Soc Sci* 645:36–59
- Dawson DA, Goldstein RB, Pickering RP, Grant BF (2014) Nonresponse bias in survey estimates of alcohol consumption and its association with harm. *J Stud Alcohol Drugs* 75:695–703
- Skogen JC, Harvey SB, Henderson M, Stordal E, Mykletun A (2009) Anxiety and depression among abstainers and low-level alcohol consumers. The Nord-Trøndelag Health Study. *Addiction* 104:1519–1529. doi:10.1111/j.1360-0443.2009.02659.x
- Torvik FA, Rognmo K, Tambs K (2012) Alcohol use and mental distress as predictors of non-response in a general population health survey: the HUNT study. *Soc Psychiatry Psychiatr Epidemiol* 47:805–816. doi:10.1007/s00127-011-0387-3
- Cunradi CB, Moore R, Killoran M, Ames G (2005) Survey nonresponse bias among young adults: the role of alcohol, tobacco, and drugs. *Subst Use Misuse* 40:171–185
- McCabe SE, Schulenberg JE, Johnston LD, O'Malley PM, Bachman JG, Kloska DD (2005) Selection and socialization effects of fraternities and sororities on U.S. college student substance use: a multi-cohort national longitudinal study. *Addiction* 100:512–524
- Bachman JG, O'Malley PM, Schulenberg JE, Johnston LD, Bryant AL, Merline AC (2002) The decline of substance use in young adulthood: changes in social activities, roles, and beliefs. Lawrence Erlbaum Associates, Mahwah
- Johnston LD, O'Malley PM, Bachman JG, Schulenberg JE, Miech RA (2014) Monitoring the future national survey results on drug use, 1975–2013. Volume II: college students and adults ages 19–55. University of Michigan Institute for Social Research, Ann Arbor
- Substance Abuse and Mental Health Services Administration (2014) Results from the 2013 national survey on drug use and health: summary of national findings, NSDUH series H-48, HHS publication no. (SMA) 14–4863. Substance Abuse and Mental Health Services Administration, Rockville
- Grant BF, Kaplan K, Shepard K, Moore T (2003) Source and accuracy statement for wave 1 of the national epidemiologic survey on alcohol and related conditions (NESARC). National Institute on Alcohol Abuse and Alcoholism, Bethesda
- Grant BF, Kaplan KD (2005) Source and accuracy statement for the wave 2 national epidemiologic survey on alcohol and related conditions (NESARC). National Institute on Alcohol Abuse and Alcoholism, Rockville
- Grant BF (1996) DSM-IV, DSM-III-R and ICD-10 alcohol and drug abuse/harmful use and dependence, United States, 1992: a nosological comparison. *Alcohol Clin Exp Res* 20:1481–1488
- Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R (2003) The alcohol use disorder and associated disabilities interview schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug Alcohol Depend* 71:7–16
- Grant BF, Harford TC, Dawson DA, Chou PS, Pickering R (1995) The alcohol use disorder and associated disabilities schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. *Drug Alcohol Depend* 39:37–44
- Hasin D, Carpenter KM, McCloud S, Smith M, Grant BF (1997) The alcohol use disorder and associated disabilities interview schedule (AUDADIS): reliability of alcohol and drug modules in a clinical sample. *Drug Alcohol Depend* 44:133–141
- Hasin D, Grant BF, Cottler L et al (1997) Nosological comparisons of alcohol and drug diagnoses: a multisite, multi-instrument international study. *Drug Alcohol Depend* 47:217–226
- Hasin D, Li Q, McCloud S, Endicott J (1996) Agreement between DSM-III-R, DSM-IV and ICD-10 alcohol diagnoses in a U.S. community-sample of heavy drinkers. *Addiction* 91:1517–1527
- Hasin DS, Van Rossem R, McCloud S, Endicott J (1997) Alcohol dependence and abuse diagnoses: validity in a community sample of heavy drinkers. *Alcohol Clin Exp Res* 21:213–219
- Muthen BO, Grant BF, Hasin DS (1993) The dimensionality of alcohol abuse and dependence: factor analysis of DSM-III-R and proposed DSM-IV criteria in the 1988 National Health Interview Survey. *Addiction* 88:1079–1090
- Nelson CB, Rehm J, Usten B, Grant BF, Chatterji S (1999) Factor structure for DSM-IV substance disorder criteria endorsed by alcohol, cannabis, cocaine and opiate users: results from the World Health Organization reliability and validity study. *Addiction* 94:843–855
- Pull CB, Saunders JB, Mavreas V et al (1997) Concordance between ICD-10 alcohol and drug use disorder criteria and diagnoses as measured by the AUDADIS-ADR, CIDI, and SCAN: results of a cross-national study. *Drug Alcohol Depend* 47:207–216
- Heeringa SG, West BT, Berglund PA (2010) Applied survey data analysis. Chapman and Hall, London
- Archer KJ, Lemeshow S, Hosmer DW (2007) Goodness-of-fit tests for logistic regression models when data are collected using

- a complex sampling design. *Comput Stat Data Anal* 51:4450–4464
26. Little RJA, Vartivarian S (2005) Does weighting for nonresponse increase the variance of survey means? *Surv Methodol* 31:161–168
 27. Zhao J, Stockwell T, Macdonald S (2009) Non-response bias in alcohol and drug population surveys. *Drug Alcohol Rev* 28:648–657
 28. Groves RM, Couper MP (1998) *Nonresponse in household interview surveys*. Wiley, New York
 29. Lepkowski JM, Couper MP (2002) Nonresponse in the second wave of longitudinal household surveys. In: Groves et al. (ed) *From survey nonresponse*, pp 259–271. Wiley, New York
 30. Kott PS (2012) Why one should incorporate the design weights when adjusting for unit nonresponse using response homogeneity groups. *Surv Methodol* 38:95–99
 31. West BT (2009) A simulation study of alternative weighting class adjustments for nonresponse when estimating a population mean from complex sample survey data. In: *Proceedings of survey research methods section, 2009 joint statistical meetings*, pp 4920–4933
 32. Johnston LD, O'Malley PM, Bachman JG, Schulenberg JE, Miech RA (2014) *Monitoring the future national survey results on drug use, 1975–2013. Volume I: secondary school students*. University of Michigan Institute for Social Research, Ann Arbor
 33. Compton WM, Dawson D, Duffy SQ, Grant BF (2010) The effect of inmate populations on estimates of DSM-IV alcohol and drug use disorders in the United States. *Am J Psychiatry* 167:473–475. doi:[10.1176/appi.ajp.2009.09081087](https://doi.org/10.1176/appi.ajp.2009.09081087)
 34. West BT (2013) An examination of the quality and utility of interviewer observations in the National Survey of Family Growth. *J R Stat Soc Ser A* 176:211–225