Alcohol-Induced Deaths and the Minimum Legal Drinking Age

Jeffrey DeSimone*

December 31, 2017

Abstract

Using regression discontinuity based on annual 2007–2015 U.S. averages and age in years, I estimate how reaching the minimum legal drinking age affects alcohol-induced mortality, for which a death certificate ICD-10 code reflects alcohol as a direct cause. At age 21, the alcohol-induced death rate jumps by 30–50%, simultaneously with large alcohol use increases. Estimates from 1999–2006 are similar, closely replicating previous findings using age in days. Comparable effects are absent at other ages, on population size, and for other leading causes of death. Among deaths with alcohol mentions, effects are significant for motor vehicle accidents (MVA) and alcohol-induced underlying causes, but not suicides or drug-induced causes. Unlike in earlier years, effects are insignificant for non-alcohol MVAs and suicides. Estimates are robust to various local and OLS regression specifications, evident among both genders, and proportionately large among non-Hispanic whites and non-MSA residents, for alcohol poisoning, and during the Great Recession.

* Department of Marketing, Industrial Distribution and Economics, University of Alabama at Birmingham; BEC 219, 1150 10th Ave. S., Birmingham, AL 35205; desimone@uab.edu. I thank Kitt Carpenter along with participants in sessions at the 2017 WEA and SEA meetings, particularly Dhaval Dave and Alison Cuellar, for helpful comments.

1. Introduction

Using a regression discontinuity (RD) design based on annual U.S. averages over 2007–2015 and year of age, I estimate the impact of reaching the minimum legal drinking age (MLDA) on alcohol-induced mortality. Results imply that lowering the MLDA would substantially raise deaths caused by alcohol use among those for whom drinking would be newly legal.

In 2008, two decades after the last U.S. state raised it to 21, the MLDA became the focus of renewed attention when as signatories to the Amethyst Initiative, a group of college and university presidents argued that alcohol consumption among 18–20 year olds would be less dangerous if drinking was legal among this cohort. However, using data on age in days among 19–22 year olds, Carpenter & Dobkin (2009 & 2011, henceforth CD) convincingly showed that mortality rates, particularly from external causes such as suicide and motor vehicle accidents (MVAs), spike upward just after the 21st birthday.

As documented by the Centers for Disease Control and Prevention (CDC, e.g. http://nccd.cdc.gov/DPH_ARDI/Info/ICDCodes.aspx), alcohol contributes to many different types of mortality, directly leading to acute causes (e.g. motor vehicle crashes, injuries and suicides) and elevating the risk of chronic causes (e.g. various cancers and cardiovascular diseases). I focus on deaths for which the corresponding death certificate lists an ICD-10 code reflecting an underlying or additional multiple cause of death that the CDC categorizes as 100% alcohol attributable, i.e. occur only as the direct result of excessive drinking. This categorization is identical to what CD called "deaths with a mention of alcohol," but substitutes the CDC's "alcohol-induced" terminology. The most relevant for this age group are alcohol poisoning, alcoholic psychosis, alcohol abuse and dependence, and high alcohol blood levels, but also

included are alcohol-related liver disease, pancreatitis, cardiomyopathy, nervous system degeneration and gastritis.

Effects of alcohol control policies such as drunk driving laws, taxes and MLDAs on mortality from many different causes that are partially or indirectly attributable to alcohol have been studied extensively.[1] In contrast, little research has isolated the policy responsiveness of deaths that are directly induced by alcohol. Sloan et al. (1994), using 1982–1988 panel data on U.S. states, found that higher alcohol prices lowered some types of alcohol-related mortality, but not deaths for which alcohol is the immediate cause. Wagenaar et al. (2009) estimated that in Alaska, alcohol-induced mortality fell by 20% and 9%, respectively, following alcohol tax increases in 1983 and 2002, although the latter decline was not statistically significant.

My study builds on CD, particularly their finding that alcohol-induced deaths increase abruptly at age 21 by 35–45%. My sample period begins upon the discontinuation of ICD-10 code F10.0 (mental and behavioral disorders due to alcohol use, acute intoxication) in 2007, and thus is distinct from that covered by CD. Moreover, I study all alcohol-induced deaths, whereas CD examined only those for which suicide and MVAs, which they studied separately, were not the underlying cause of death.

A distinguishing feature of my study is the use of death rates specific to only year of age, which is commonly available in publicly accessible data. To implement RD in this setting, my sample includes ages 14–27, or a seven-year range on each side of age 21. Like CD, I compare age 21 discontinuities estimated from local linear regressions (LLR) with those from OLS models that include an age polynomial (of order five, the maximum identified). Further, I address potential rounding bias from using coarse age data by estimating local quadratic

[1] Recent examples include Cook & Durrance (2013), Cook et al. (2005), and Johansson et al. (2014).

regressions (LQR), which Calonico et al. (2014) show correct for smoothing bias. Likewise, I estimate OLS models with separate 3rd-order age polynomials on each side of age 21, and adjust the RD estimate based on the coefficients of the linear and quadratic age term interactions with the age 21 indicator, as outlined by Dong (2015).

My analysis begins by establishing, in 2007–2015 NSDUH data, that drinking increases discontinuously at age 21. This corroborates abundant existing evidence, including from CD, but covers my more recent sample period, and verifies that this increase is detectable in data that are aggregated to the age-by-year level and identify only year of age. Consistently across local and OLS regressions, past month alcohol use and binge drinking participation and days each rise at age 21 by about 20% relative to the levels predicted along the corresponding age profiles. Estimates are large and significant across genders, racial/ethnic groups, and MSA status.

The main result of the analysis is that across regression specifications, the alcohol-induced death rate likewise increases sharply and significantly at age 21. This is true for LLRs and LQRs using various bandwidths and kernels, and for both local and OLS regressions regardless of bias correction or logging the death rate. Correcting for rounding bias increases the estimates by allowing the death rate age profile to flatten approaching age 21 from each side. Mirroring CD's estimates, effect size is 30–50% of the predicted age 21 death rate, with preferred models yielding estimates near the center of that range.

I next estimate models on data extending back to 1999, when death certificate coding switched from ICD-9 to ICD-10. Estimates from 1999–2006 data are similar to those from 2007–2015, suggesting that discontinuing code F10.0 had little effect. As CD's sample period was 1997–2004, I show that excluding 2005 and 2006 also has minimal impact. When I mimic

CD by restricting the sample to deaths not attributable to suicide or MVAs (or homicide), I obtain bias-corrected estimates of between 30–45%, closely replicating CD's findings.

Results of various falsification exercises suggest that the significant RD effects are not attributable merely to using less-specific age information. Effects are not consistently large or significant at any age from 17–25 other than 21. The age-specific population increase at the MLDA is always highly insignificant and typically below 0.5%. Excluding suicides and MVA deaths, which CD established respond to the MLDA, the age 21 death rate increase is significant (only marginally) for just one of the next 12 leading causes of death among 14–27 year olds, representing all categories with more than 2,000 sample period mentions among this cohort.

Decomposition by underlying cause reveals that roughly 40% of the age 21 increase in death rates is attributable to MVAs (21% of sample alcohol mentions), with another 20% each to alcohol-induced (19% of mentions) and causes other than alcohol, drugs, MVAs or suicide (13% of mentions). Age 21 increases are significant for each of these individual categories, but insignificant for drug-induced causes (39% of mentions) and suicide (8% of mentions). Effects are much smaller and insignificant for deaths with no alcohol mention, both combined and for each of these separate categories.

Finally, subgroup analysis shows that while MLDA effects are generally large and significant across subsamples defined in a variety of ways, they are proportionately largest among non-Hispanic whites, outside MSAs, for poisoning, and during the Great Recession.

In conclusion, consistent with CD but contradicting the claims of the Amethyst Initiative, the results imply that a MLDA of 21 rather than 18 saved an average of at least 79 lives annually over the sample period. Adjusting the EPA's statistical value of life to 2016 terms (and fully allocating it to ages 18 and over) implies a corresponding minimum savings of nearly $800

million per year. More broadly, the close replication of CD's estimates suggests that MLDA effects on many additional outcomes of alcohol use can be productively studied in publically available data. While my standard errors imply a limited ability to identify small effects, this appears to be a function of the specific outcomes rather than the research design: in 1999–2004 data, standard errors for age 21 effects on MVA, suicide and homicide deaths are no larger than those from CD.

2. Data

Mortality data come from the Multiple Cause of Death database, produced by the CDC's Division of Vital Statistics and available for 1999 onward at https://wonder.cdc.gov/mcd.html. The database encompasses all death certificates of U.S. residents filed in the U.S. Each death certificate lists a single underlying cause of death (UCD), up to twenty additional multiple causes (MCD), and demographic data. Death rates are based on Census Bureau age-specific population estimates that are also included in the database, consisting of April 1 modified Census counts in 2000 and 2010 and July 1 intercensal or postcensal estimates in other years.

A limitation of the publicly available demographic information is that only year of age is reported. Balancing the goals of maintaining sufficient age variation to identify MLDA effects while including only ages reasonably close to the 20-to-21 dividing line and old enough so that alcohol deaths are consistently nonzero, my sample encompasses ages 14 through 27, i.e. a seven-year radius around the MLDA threshold.

I consider deaths to be alcohol-induced if any ICD-10 code labeled as such by the CDC, corresponding to those occurring only as the direct result of excessive drinking, is listed among the multiple causes on the death certificate. Relevant codes include F10 (mental and behavioral

disorders due to alcohol use), G31.2 (degeneration of nervous system due to alcohol), I42.6 (alcoholic cardiomyopathy), K29.2 (alcoholic gastritis), K70 (alcoholic liver disease), K85.2 and K86.0 (alcohol-induced acute and chronic pancreatitis, respectively), R78.0 (alcohol in blood), and X45, X65 and Y15 (poisoning accidentally, intentionally, and with undetermined intent, respectively).

An important consideration for sample construction is that ICD-10 code F10.0 (acute intoxication), which comprised 43% of all alcohol mentions (and was listed for 46% of alcohol-induced deaths) over 1999 to 2006, was discontinued in 2007. This necessarily resulted in a reallocation of deaths to other alcohol categories, largely the three poisoning codes along with F10.9 (unspecified mental and behavioral disorder). Subsequent alcohol-induced deaths are also more likely to have multiple alcohol codes. In case this also affected whether deaths were coded as alcohol-induced at all, I focus on the post-discontinuation 2007–2015 period, while separately estimating results in the earlier period for comparison. Since the unit of observation in the analysis is an age-by-year cell, the full sample contains 126 observations, representing 14 age-years in each of nine years.

Of the 392,774 deaths among 14–27 year olds during 2007–2015, alcohol is listed as a MCD for 16,083, or 4.1%. Not surprisingly, this share increases with age, from 1.9% to 3.7% to 5.1% among ages 14–19, 20–21, and 22–27, respectively. By far the most-often listed specific causes are accidental poisoning and unspecified mental and behavioral disorders from alcohol use, which are mentioned in 61.5% and 27.0%, respectively, of alcohol-induced deaths (bearing in mind that the number of mentions exceed the number of deaths by over 20%, since some deaths have multiple alcohol MCDs).

The MCD database also enables splitting the sample by various other characteristics, including gender, race/ethnicity, metropolitan area (MSA) status, day of week, and place of death. I also compare results from during the Great Recession to the subsequent period.

To establish that alcohol consumption increased at age 21 in the sample period, and that such an increase can be detected in aggregated data on age in years, I also examine data on past month drinking from the 2007 to 2015 waves of the National Survey on Drug Use and Health (NSDUH, U.S. DHHS 2016). A further limitation of the NSDUH data is that beyond age 21, only age groups 22–23, 24–25 and 26–29 are observed. To balance the age range on the lower end, I include respondents as young as 12 years old, the minimum survey age. I separately estimate age 21 increases in any alcohol use, any binge drinking, and binge drinking days, with bring drinking defined as having at least five drinks on an occasion (although this was changed to four drinks for females in 2015). In order to make the analysis of alcohol use comparable to that of mortality, I aggregate the drinking outcomes into year-by-age averages using the NSDUH person-level analysis weights before running regressions.

3. Empirical Specification

I estimate effects of reaching age 21 within a RD framework. The corresponding regression equation, where a and t represent sample ages and years, respectively, is

$$R_{at} = \beta \times D_{at} + D_{at} \times f(a_t) + (1 - D_{at}) \times g(a_t) + u_{at}.$$

Response R depends on D, an indicator for ages 21 and above, age profiles among those at least (f) and younger than (g) age 21, and an unobserved error u. The RD estimator of the discontinuous change in R at the MLDA is β.

Conceptually, RD takes the limit of f(a) and g(a) as each approaches the MLDA, projecting outcomes when just reaching age 21 by slightly extending each profile f(a) and g(a) and estimating β as the difference between these predicted rates. Computationally, I estimate the MLDA discontinuities and smoothed age profiles nonparametrically using local linear regression (LLR), and compare the results to those from OLS models that parameterize the age profile using an age polynomial and estimates the discontinuity as the shift in this profile at the MLDA.

At age a^0, LLR estimates a weighted least squares regression of the outcome on (age – a^0) for ages within a local region of a^0 defined by the specified bandwidth, and predicts a smoothed value of the outcome as the intercept of that regression. The RD estimate is simply the difference between the LLR predictions at age a^0 =21 from each profile, and thus does not depend on the accuracy of the model at any other age. To provide confirmation that the underlying LLR specifications fit the data well, I show diagrams that plot the corresponding smoothed age profiles for the entire range of sample ages.

I operationalize the use of information only on age in years in the LLR context by specifying a threshold age of 21, and coding observed ages as (age + 0.5). This locates the discontinuity exactly halfway between the age profiles on either side, reflecting the premise that exact ages have an approximately uniform distribution. Dong (2015) presents evidence that any departures from uniformity in the distribution of births within a year, which others have found to be statistically significant because of seasonality, are quite small from a practical perspective. Specifically, differences between the first four empirical moments of the birth date distribution in three separate data sets and the corresponding moments of a true uniform distribution are never more than a few thousandths of a point.

Because the LLR RD estimate depends on extending the age profiles to the MLDA from the closest observations along the age distribution, linearization can create bias. This issue of rounding bias is inherent to LLR, because it stems from the unavoidable discretization of the running variable (age in this case), but clearly is more likely to be of practical importance with the relatively coarse age-in-years data that I use. However, Calonico et al. (2014), who also wrote the Stata routine I use to estimate LLRs, showed that rounding bias can be corrected simply by using a higher-order polynomial in age. Given that they recommend a quadratic as a starting point and I have only seven age values on each side of the MLDA, I therefore also show results from local quadratic regressions (LQR), for which the methodology is identical to that explained above other than using a quadratic in age.

The key parameter for local regression is the bandwidth, or the age-year radius within which observations are considered "local." For the models of primary importance, I show results using all bandwidths chosen by at least one of the three most commonly used selection algorithms, those proposed by Ludwig & Miller (2007), Imbens & Kalyanaraman (2012), and Calonico et al. (2014), respectively (henceforth LM, IK and CCT, which tend to produce values from larger to smaller in the order listed), rounded to the nearest integer, along with any integer-valued bandwidths within the extreme values of these.[2] With many of the robustness checks, for brevity I show results for the median value of the bandwidths chosen by the three methods above. In all cases, bandwidths are determined separately for LLR and LQR.

[2] At age 21 along the lower age profile, for example, any bandwidth of at least 2.5 years but no greater than 3.5 years will include ages 18, 19 and 20. However, as the bandwidth increases from 2.5 to 3.5, the absolute weight on each of the three observations increases, thus raising the relative (actual) weight on ages that are further from age 21. To eliminate this variation, I use only integer-valued bandwidths, corresponding to the midpoint of the range of bandwidths that each use observations from the same age values (e.g. 3 years in this example).

Weights in the weighted least squares local regressions are determined by the kernel function. I follow the preponderance of the empirical LLR literature in using a triangular kernel, which assigns linearly decreasing weight moving from the center to the endpoints of the bandwidth. For the main estimates, I also show robustness to using two common alternatives, a rectangular kernel which assigns the same weight to each observation in the bandwidth, and the Epanechnikov kernel which is a parabolic and slightly more efficient version of the triangular.

To ensure that my RD estimates are not somehow an artifact of using local regression methods on the unique format of my data, I also estimate MLDA discontinuities using OLS, with the a terms in the equation above equal to actual age in years without 0.5 added. First, I simplify this equation by setting $f(a) = g(a)$, thus specifying a single polynomial in age that yields a continuous profile throughout the age distribution which potentially shifts at age 21. This is efficient, and eliminates the possibility that a discontinuity emerges only because of slope differences in the age profiles on either side of the MLDA. I use an age polynomial of order five, the maximum identified without inducing perfect collinearity (i.e. including the 6^{th} power of age forces the linear age term to be dropped from the model).

Of course, if the slopes of the age profiles on either side of the MLDA are indeed different, constraining them to be the same is inappropriate. I therefore also show OLS estimates that again allow $f(a)$ and $g(a)$ to differ by specifying separate 3^{rd}-order age polynomials for those younger and older than age 21. An advantage of this specification is that it allows for the use of a correction for rounding bias of the type discussed above, developed by Dong (2015). Under the assumption of a uniform exact age distribution, this correction merely adjusts the estimated β, i.e. coefficient of the age 21 indicator, based on the estimated effects of the linear and quadratic

age terms that are interacted with the age 21 indicator, i.e. from the upper age profile (along with the corresponding moments of the uniform distribution).

Standard errors for both local and OLS regressions are constructed from the conventional Huber–Eicker–White estimator, so are robust to arbitrary heteroskedasticity. I do not cluster on age, which almost universally produces smaller standard errors in my regressions, both to be conservative and because the number of clusters is small.

Since I focus primarily on unlogged death rates, the tables also show predicted values of outcomes at age 21 in the absence of the MLDA, for use as a base in assessing proportionate effects. These are the right endpoints of the lower age profiles from the LLRs and LQRs, sometimes averaged across bandwidths and/or linear and quadratic models as indicated.

4. Results

a. Changes in Binge Drinking at Age 21

I begin by using the 2007–2015 NSDUH data to investigate whether past month drinking increases discontinuously at age 21, which is necessary for any observed upward spike in alcohol-induced deaths to be plausible. Along with CD, Yörük & Yörük (2011) estimated large and significant increases in alcohol use upon turning 21 years old in the RD framework using age in days, as did Crost & Guererro (2012) using age in months. However, no corresponding evidence exists for my more recent sample period, or using aggregated age-by-year data that identify only year of age.

Recalling that for respondents beyond age 21, the NSDUH reports only age ranges 22–23, 24–25, and 26–29, I use the midpoint values of 23, 25 and 28, respectively, for these categories, constructed to be consistent with my strategy of adding 0.5 to single years of age. I

correspondingly estimate LLRs using a bandwidth of five years, the minimum integer value that includes three distinctly measured age levels beyond the MLDA. Estimates are larger using bandwidths of 3 or 4 years, which the previously described selection procedures uniformly indicate to be optimal, or the bias correction procedures (with the same 5-year bandwidth for LQR), but only slightly given the linearity of the age profiles near the MLDA as shown below.[3]

The top row of Table 1 shows results for the full sample. Regardless of regression method and drinking measure, alcohol use increases at age 21 by close to 20%, relative to the levels predicted along the corresponding lower LLR age profile.[4] Estimates are significant, and of similar proportionate size, across genders, racial/ethnic groups, and MSA status.[5]

Figure 1 graphically illustrates these discontinuities for the full sample. Alcohol consumption rises steadily with age through the MLDA, then begins to decline immediately thereafter. The significant RD estimates indicate, as is clear from the diagrams, that alcohol consumption at age 21 is more prevalent than expected given the slope of the age gradient until that point, even more so considering that this gradient becomes negative beyond age 21.

The observed responsiveness of drinking to attaining legal drinking status might be surprising given that alcohol use is widespread among those younger than the MLDA, as shown by the predicted drinking prevalence rates and days in Table 1 (and Figure 1), but this finding is consistent with those of CD and the other studies mentioned above.[6] This confirms that a

[3] LQR effects using a bandwidth of 8 years, the minimum to include all four of the observed older age levels, are slightly smaller but remain significant at 5%.
[4] Although largely beside the point for the purposes of this analysis, the estimated age 21 increases in binge drinking likelihood and days stem from being more likely to drink at all (both are insignificant among drinkers).
[5] Because non-Hispanic whites comprise about 60% of the U.S. population, compared to about 20% for the next largest group identified in the data (Hispanics of all races), and alcohol-induced mortality is relatively infrequent, I do not further subdivide the nonwhite and Hispanic group. Moreover, the only nonwhite or Hispanic subgroup for which the alcohol-induced death rate is not substantially lower than for non-Hispanic whites is non-Hispanic Native Americans, who make up well under 1% of the U.S. population.
[6] While my estimated effects are not directly comparable to those of CD, who primarily use lifetime and past year drinking measures, they appear to reflect similar or larger proportionate increases in drinking at the MLDA.

drinking increase upon reaching the MLDA can be identified in aggregate data on age in years, and that it persists in my more recent sample, making plausible a positive effect on alcohol-induced deaths.

b. Changes in Alcohol-Induced Deaths at Age 21

Table 2 addresses the main question of interest, showing RD estimates of the increase in the rate of alcohol-induced deaths upon reaching the MLDA from a variety of specifications. In panels A and B, columns represent bandwidths. For all specifications shown other than log deaths using LLR, the CCT, IK, and LM bandwidths were the narrowest, 2nd-widest, and widest listed, respectively. I also show the 2nd narrowest listed (3 years for LLR, 4 years for LQR) for completeness. In panel C, columns represent each of the four OLS specifications possible from varying whether the age polynomial is interacted with the age 21 indicator (i.e. estimated separately for those not legal and legal to drink) and whether the death rate is logged. Estimates for the separate-polynomial models reflect the previously described Dong (2015) bias correction.

All of the Table 2 estimates support the main conclusion of the analysis, which is that the alcohol-induced death rate increases sharply and significantly at age 21. Two patterns that emerge are that the RD estimates decline in magnitude as the bandwidth for the local regressions increases, and are larger for bias-corrected models (i.e. LQR and OLS with separate age profiles). Figure 2 shows that both patterns are explained by the age profiles on both sides of the MLDA, but particularly that to the left, flattening upon approaching age 21, a phenomenon more accurately modeled using narrower bandwidths and more flexible specifications.

Panel A shows that alternative kernels make little difference. As such, remaining local regressions utilize triangular kernels, which place greater weight on observations from ages

closer to the MLDA in a straightforward linear way. Age profiles for log deaths flatten substantially approaching age 21, resulting in all three selection procedures choosing bandwidths of 2 years (though the estimate using 3 years is shown for comparison), and LQR estimates in panel B being larger and more stable across bandwidths than the corresponding LLR estimates. OLS estimates, particularly for unlogged death rates, are quite similar to those from local regressions using triangular kernels and the median optimal bandwidths of 4 years (LLR) and 5 years (LQR) which are used to construct Figure 2 (both from the IK procedure).

Relative to the death rate predicted at age 21 along the lower age profiles (or based directly on the log model coefficients), the Table 2 estimates range in size from 30% to 55%. The preferred models – those using the unlogged death rate, a triangular kernel, and the median optimal bandwidths (i.e. in the 3rd column of the top row of panels A and B, and the first two columns of panel C) – narrow that range to between 37% and 47% (matching the LQR age 21 prediction with the bias-corrected OLS estimate). These are large effects, but correspond quite closely to the 35–45% range of effects estimated by CD.[7]

c. Replication of CD

While the similarity of my estimates to those of CD are reassuring, they reflect a completely non-overlapping time period. To more directly address whether CD's results can be replicated using my data structure, I next estimate models on data from 1999–2006, which represents the period during which death certificates were coded using ICD-10 (after switching

[7] The age 21 discontinuity is also large and highly significant when the outcome is specified as the proportion of all deaths in which alcohol is a MCD, even though overall mortality is estimated to increase by as much as 6–7% at age 21 (albeit insignificantly). The effect size varies between 0.009 and 0.012 for the unlogged LLR (4-year bandwidth), LQR (5-year bandwidth) and OLS models, representing 29–38% of the predicted fraction at age 21, with t-statistics between 3.2 and 3.6.

from ICD-9) prior to my sample period. I examine results for three slices of these data: the full sample, for comparison with the subsequent period on which I focus; excluding the last two years, to more closely match CD's 1997–2004 sample period; and over 1999–2004 but excluding UCDs from homicide, MVA and suicide, to match CD's sample construction.

As is clear from Figure 3, the age profiles over this earlier period show much greater nonlinearity than those for the subsequent period, which is further exacerbated when the last two years are excluded. Because LLRs and the single-polynomial OLS model do not adequately capture these nonlinearities, I show only LQR and bias-corrected OLS results. Both the LM and CCT selection procedures indicate optimal LQR bandwidths of 3 years for all three of the samples described above (using unlogged death rates). As this is the minimum bandwidth required to identify a quadratic profile, I also display estimates using a 4-year bandwidth.

Table 3 contains these results. Estimates from 1999–2006 are similar in magnitude and significance with those from 2007–2015, as expected given the close correspondence between the latter and those from CD, suggesting that discontinuing code F10.0 had little impact. Proportionately, the discontinuities grow slightly when 2005–2006 are excluded, with the largest approaching 60% of the predicted death rate at age 21. As the table and Figure 3 imply, restricting the sample to deaths not attributable to homicide, suicide or MVAs eliminates roughly half of all alcohol-induced deaths. Estimates in this sample are therefore noisier, but generally remain significant at 10%. They vary in size from 30–45% of the age 21 prediction, which closely replicates CD's findings. The two unlogged LQR estimates represent increases at the MLDA of 35–45%, the exact range encompassed by CD's estimates. This is striking, particularly considering that the added nonlinearity that appears in their sample period compared

with mine is most apparent for ages 19 and 20, which represents the entire below-MLDA group that CD studied but would presumably be especially hard to capture using only age in years.[8]

d. Falsification

Next I undertake various falsification exercises. Results suggest that the significant RD MLDA effects are not attributable merely to using less-specific age information, or any other idiosyncratic aspect of the data or setting.

Table 4 estimates effects for each age between 17 and 25 other than 21. I show LLR and LQR results for bandwidths ranging from the minimum necessary to identify the corresponding polynomial (2 and 3 years, respectively), to the maximum allowing a symmetric number of age-years to be included on either side, along with OLS coefficients unadjusted for rounding bias. While power to identify small effects is particularly lacking at extreme ages, the estimated RD effects are never consistently significant or nearly as large in magnitude as those for the MLDA.[9]

Table 5 explores increases at the MLDA in population size (logged, although unlogged results are identical), which should not vary abruptly at any specific age.[10] Effects are uniformly quite small and highly insignificant. Magnitudes are typically below 0.5% and rarely above 1.0%, while t-statistics never approach 1 and are usually well below 0.5.

[8] Despite excluding the first two years of their sample period, I am also able to closely replicate CD's estimates for homicide, suicide, MVAs and drug-induced MCDs, the other specific categories they study, in 1999–2004 data.
[9] An exception is the age 20 effect using LQR, which is consistently negative and significant. However, this is confounded by the large death rate increase at age 21. Under the alternative and presumably counterfactual assumption that the age profile is continuous from age 20 onward, it must sharply increase through age 21, meaning that the small component to the left of age 20 that is pivotal for identifying the RD effect is particularly steep. This likely explains why the age 20 effect is never significant at 5% for logged deaths using LQR (not shown) and that the displayed LLR effects become insignificant and then positive as the bandwidth increases.
[10] The bandwidths displayed for local regressions are the median optimal bandwidths among the three selection procedures for the full sample log population models.

Table 6 examines the age 21 increase in other MCDs. To be consistent with the alcohol analysis, I analyze all mentions rather than just UCDs. However, I select causes based on their frequency as UCDs in the sample age group, specifically the 12 leading UCDs excluding suicides and accidental deaths, which are either mechanically or plausibly related to alcohol consumption.[11] These represent all UCDs with at least 2,000 sample period mentions, or about 16 annually for each year of age, among this cohort. I list these in order of their frequency as UCDs, which is similar though not identical to how often they are mentioned.[12] An alcohol-induced cause is included as a co-occurring MCD in fewer than 3% of deaths for each of these categories. For local regressions, I show results for the median among the optimal bandwidths selected by the three relevant procedures for the main death rate model in Table 2, a theme continued in Tables 7 and 8.

Among these major MCDs, other than in one of four specifications for anemias (at 10%), the age 21 death rate increase is significant only for septicemia, and beyond 5% only in the bias-corrected OLS model.[13] Figure 4 provides graphical evidence that for the top 6 causes, the estimated discontinuities are small, even for septicemia relative to that for alcohol mentions.

e. Decomposition by UCD

Having established that the abrupt upward shift in alcohol-induced deaths at the MLDA is unlikely to be spurious, I next decompose this effect by UCD. The most common UCD

[11] CD established that suicides and MVA deaths, both of which I examine in the next subsection, significantly increase at the MLDA, even though only 3–4% are accompanied by an alcohol mention. By comparison, over 14% of non-MV accidental deaths, which include accidental poisoning from alcohol or drugs, also list alcohol as a MCD.
[12] For example, septicemia is the UCD only about 20% of the time in which it is listed as a MCD. In addition, the pregnancy death rate is among females rather than the entire population.
[13] It is unclear whether an effect on deaths from septicemia is realistic: alcohol is concurrently mentioned with septicemia in only 1.3% of sample deaths, but the timing of its role would presumably precede that of mortality. Regardless, the RD estimate is significant, and only at 10%, for only one other LQR bandwidth.

category for which alcohol appears as a MCD is drug-induced causes, which comprise 39% of alcohol mentions. Another 19% list an alcohol-induced cause as the UCD. Two UCDs which CD show to be highly responsive to turning age 21, MVAs and suicide, comprise an additional 21% and 8%, respectively. I group the remaining 13% into a combined category encompassing all other UCDs.[14]

Panel A of Table 7 presents the results. Roughly 40% of the age 21 increase in death rates is attributable to MVAs, with another 20% each corresponding with causes that are alcohol-induced and ones other than alcohol, drugs, MVAs or suicide. Age 21 increases are significant for each of these individual categories, but not for drug-induced causes or suicide.[15] Figures 5 a.–e. verify graphically that the discontinuities are more substantial for alcohol-induced UCDs, MVAs, and residual-cause UCDs than for the remaining two categories.

Panel B similarly decomposes MLDA effects among deaths with no alcohol mention. The increase at age 21 is less than 6%, two-thirds of which is among MVAs. Both of these effects are statistically insignificant, as are those for the other specific categories. Figures 5 f.–h. show that the discontinuities are non-trivial in size, despite their insignificance, for overall and MVA deaths without alcohol mentions, but small for remaining non-alcohol MCD deaths.[16]

[14] Among these residual causes, 63% of alcohol mentions accompany accidents other than MVAs or poisoning from alcohol or drugs, 20% occur along with one of the causes studied in the falsification exercise of Table 6, and 7% correspond with events of undetermined intent (leaving 10% that accompany all other underlying causes).

[15] The insignificance and small magnitude of the age 21 suicide effect, both with and without accompanying alcohol MCD (as described next), contradicts CD's findings. Since I replicate CD's large suicide effect in the earlier period, and this persists when the sample is extended through 2009, the difference appears to reflect a sudden decline in suicide responsiveness beginning after the Great Recession. However, gun suicides, which have driven the effect all along, continue to increase sizably and significantly upon reaching the MLDA in the recent post-recession period.

[16] The implied effect of 15% for non-alcohol MVA deaths is actually in line with that estimated by CD for all MVA deaths (which is 17% here though not shown in the table), but my standard error is considerably larger. Further analysis indicates that the MVA effect variability increased substantially during the Great Recession: over 2010–2015, the MVA effect is essentially the same proportionately but has a standard error that is three times smaller. This results in a significant age 21 effect on both all deaths with no alcohol mention and overall mortality in this later period, the latter consistent with CD (who do not examine the former), although the estimated discontinuity is unchanged for non-MVA deaths with no alcohol mention.

f. Subgroup Analysis

Finally, Table 9 shows results for subgroups formed by dividing the sample along the same gender, racial/ethnic, and MSA status lines used for the alcohol use analysis in Table 1, along with by time, specific death category, day of the week and place of death. Other than for Hispanics and nonwhites, among whom the age 21 increase in deaths is above 15% yet small relative to the overall discontinuity and insignificant, MLDA effects are persistently large and significant, though proportionately largest among non-Hispanic whites, outside MSAs, for poisoning, and during the Great Recession. Figures 6 a.–b. graphically depict the effects for both genders, while c.–h. do so for the category within each grouping with the largest discontinuity.

In accordance with alcohol-induced deaths being four times as likely among males, their MLDA increases are larger in magnitude, but not proportionately in bias-corrected models that adequately account for the greater nonlinearity in female age profiles. Comparison with Table 1, however, implies that any specific drinking occasion is much more likely to become deadly for males than females, given that males are only about 50% more likely to binge drink at all and do so less than twice as frequently.

The same does not appear to be true when comparing non-Hispanic whites with others, as their death and alcohol use rates are larger by roughly the same proportion. Instead, reaching the MLDA seems to induce problematic drinking to a greater extent for non-Hispanic whites.

Similarly, alcohol-induced mortality varies considerably more across MSA status than does alcohol consumption. Even beyond that, reaching the MLDA raises deaths from alcohol by a substantially larger proportion for those living outside than within an MSA.[17] While the

[17] The MSA status classification schemes differ between the mortality and NSDUH data for two reasons. First, MSA categorizations were done retrospectively using definitions as of 2013 in the mortality data, but concurrently at

fraction of deaths occurring at home decreases moving to smaller and then out of MSAs, so does that in outpatient facilities and ERs, while the likelihood of DOA deaths increases. As such, the greater incidence of alcohol deaths and MLDA effect size outside of MSAs is more likely explained by the sparser density of emergency facilities than being more likely to drink heavily at home.

Alcohol deaths were more frequent, and the MLDA effect was disproportionately larger, during the recession years of 2007–2009. Compared with the subsequent period, the recession age 21 effect was twice as large for deaths occurring at home, in an ER or outpatient facility, or as DOA, but only about 40% larger for other deaths (not shown in Table 8). This suggests that during the recession, heavy drinking shifted away from public venues into homes, where emergency assistance was less available.[18]

Age 21 effects are much stronger for poisoning, regardless of intent, than other alcohol-induced causes. This is not surprising under the assumption that drunk driving is less responsive to attaining legal drinking age than is heavy drinking that does not involve driving. In addition, as the quantitatively substantive (and statistically significant, outside of 2007–2009) effects on non-alcohol MVA deaths implies, many fatal MVAs resulting from drunk driving are likely not coded as alcohol-induced, depending on whether the driver was killed and other circumstances regarding the law enforcement response.

While nearly as many drinking deaths occur on the weekend as during the week, as one might expect, any related difference in drinking intensity does not translate to proportionately

---

the time of the survey in the NSDUH. Second, the NSDUH separately groups together all counties in large metro areas and those in smaller metro areas; the closest approximation of this possible in the mortality data is dividing between large MSA counties with principal cities and all other MSA counties.

[18] The Table 6 results differ little between the two periods, except that the septicemia effect is significant only during the recession. An earlier footnote outlines how reaching the MLDA significantly raised suicides until after the recession, when the effect on MVAs regained high significance.

larger weekend effects.[19] Lastly, while only about 40% of deaths occur at home, in an ER or outpatient facility, or as DOA, MLDA effects are slightly larger for those locations, perhaps due to differences in the availability of emergency treatment as hypothesized for the recession.

5. Conclusion

This study has found evidence that reaching age 21 significantly and substantially increases alcohol-induced mortality. Estimates suggest that upon reaching the MLDA, deaths with a MCD resulting directly from alcohol use increase by at least 30%. Several implications follow from this finding.

Behaviorally, this result echoes CD in providing tangible evidence that alcohol consumption has a causal effect on mortality. Clearly, alcohol-induced mortality cannot occur in the absence of alcohol use. However, as has been well-established, heavy drinkers are different from moderate drinkers and non-drinkers in ways that likely imply excess mortality among heavy drinkers even in the absence of drinking. The fact that an exogenous shock, reaching age 21, raises the alcohol-induced death rate is a direct indication that drinking raises mortality – though given the additional results from CD, likely serves as a lower bound for the overall extent of this phenomenon.

From a policy perspective, the age 21 effect contradicts the claims of the Amethyst Initiative, implying that a MLDA of 21 rather than 18 saves rather than costs lives. The preferred effect in Table 2, 0.112, represents 45% of the expected death rate at age 21 in the absence of a change in legal drinking status. Multiplying this percentage by the average number

[19] Consistent with literature that infers alcohol-related MVA deaths using early-morning timing, most Friday and Saturday night alcohol-induced deaths appear to occur after midnight, since counts are nearly identical on Saturdays and Sundays, and Friday counts are much closer to those on Mondays through Thursdays than to those on weekend days.

of sample deaths among 18–20 year olds yields an estimate of 116 lives saved annually. Adjusting the EPA's statistical value of life to 2016 terms, and fully allocating it to ages 18 and over, yields a corresponding minimum savings of well over $1.1 billion per year. Unless the consumption value lost exceeds other costs saved (starting with deaths avoided from other external causes, as per CD), this calculation understates the value of having a MLDA of 21 rather than 18.

Methodologically, the close replication of CD's estimates suggests that MLDA effects on many additional outcomes of alcohol use can be productively studied in publically available data. While my standard errors imply a limited ability to identify small effects, this appears to be a function of the frequency and variability of the outcome studied, rather than being inherent to the research design. In 1999–2004 data, for example, standard errors for age 21 effects on MVAs, suicide and homicide deaths are no larger than those from CD. Standard errors are also sufficiently small using the main analysis sample to identify significant effects of 3% for alcohol use (Table 1), and 6% for deaths from heart disease (Table 6) as well as from drug use and suicide with no alcohol mention (Table 7).

References

Calonico, Sebastian, Matias D. Cattaneo & Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," Econometrica, 82(6), November 2014, 2295–2326.

Carpenter, Christopher & Carlos Dobkin, "The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the Minimum Drinking Age," American Economic Journal: Applied Economics, 1(1), January 2009, 164–182.

Carpenter, Christopher & Carlos Dobkin, "The Minimum Legal Drinking Age and Public Health," Journal of Economic Perspectives, 25(2), Spring 2011, 133–156.

Cook, Philip J. & Christine Piette Durrance, "The Virtuous Tax: Lifesaving and Crime-Prevention Effects of the 1991 Federal Alcohol-Tax Increase," Journal of Health Economics, 32(1), January 2013, 261–267.

Cook, Philip J., Jan Ostermann & Frank A. Sloan, "The Net Effect of an Alcohol Tax Increase on Death Rates in Middle Age," American Economic Review, 95(2), May 2005, 278–281.

Crost, Benjamin & Santiago Guerrero, "The Effect of Alcohol Availability on Marijuana Use: Evidence from the Minimum Legal Drinking Age," Journal of Health Economics, 31(1), January 2012, 112–121.

DeSimone, Jeffrey & Angela Dills, "Alcohol Consumption and the Minimum Legal Drinking: Regression Discontinuity Evidence using only Year of Age," Working Paper, 2016.

Dong, Yingying, "Regression Discontinuity Applications with Rounding Errors in the Running Variable," Journal of Applied Econometrics, 30(3), April/May 2015, 422–446.

Johansson, Per, Tuomas Pekkarinen & Jouko Verho, "Cross-Border Health and Productivity Effects of Alcohol Policies," Journal of Health Economics, 36, July 2014, 125–136.

Ludwig, Jens and Douglas L. Miller, "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," Quarterly Journal of Economics, 122(1), February 2007, 159–208.

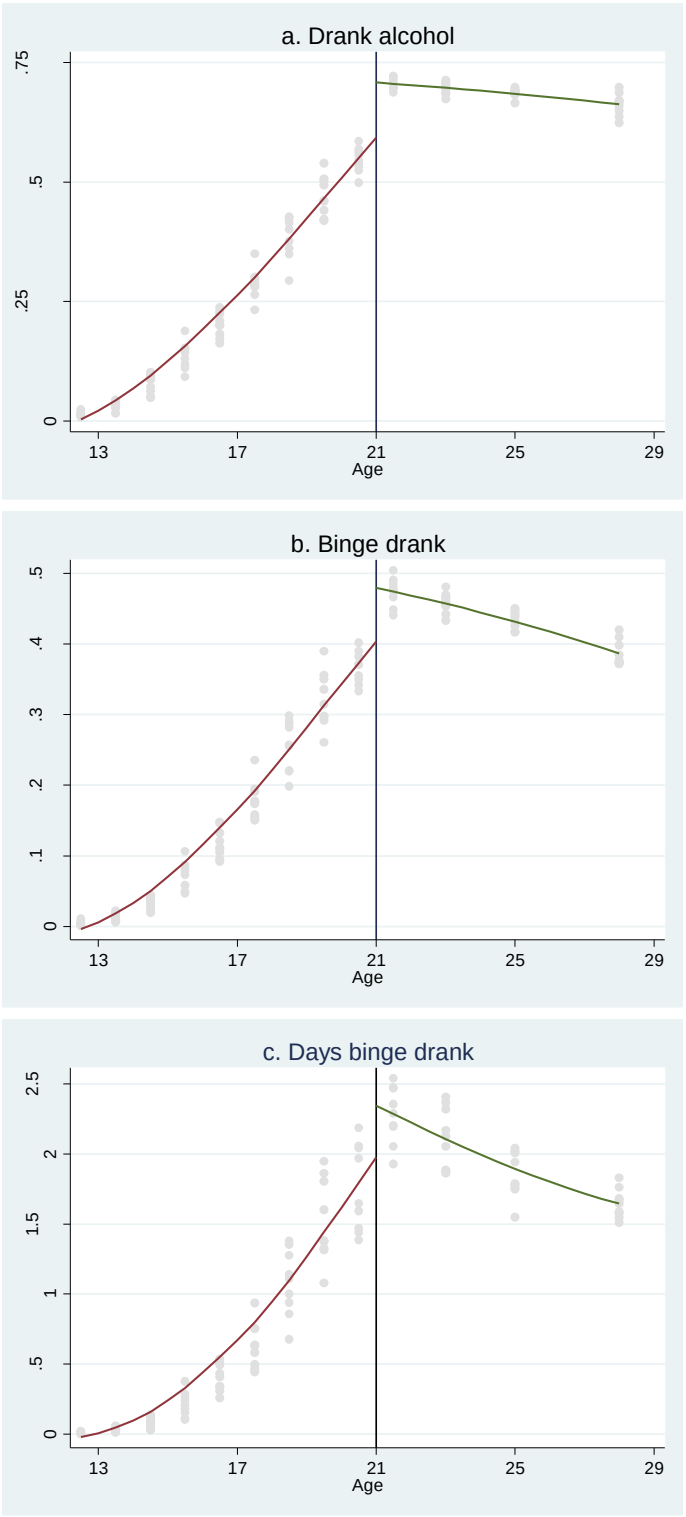
Sloan, Frank A., Bridget A. Reilly & Christoph Schenzler, "Effects of Prices, Civil and Criminal Sanctions, and Law Enforcement on Alcohol-Related Mortality," Journal of Studies on Alcohol, 55(4), July 1994, 454–465 (1994).

United States Department of Health and Human Services (U.S. DHHS), Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2014 (and earlier years), ICPSR 36361–v1, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2016.

Wagenaar, Alexander C., Mildred M. Maldonado-Molina & Bradley H. Wagenaar, "Effects of Alcohol Tax Increases on Alcohol-Related Disease Mortality in Alaska: Time-Series Analyses From 1976 to 2004," American Journal of Public Health, 99(8), August 2009, 1464–1470.

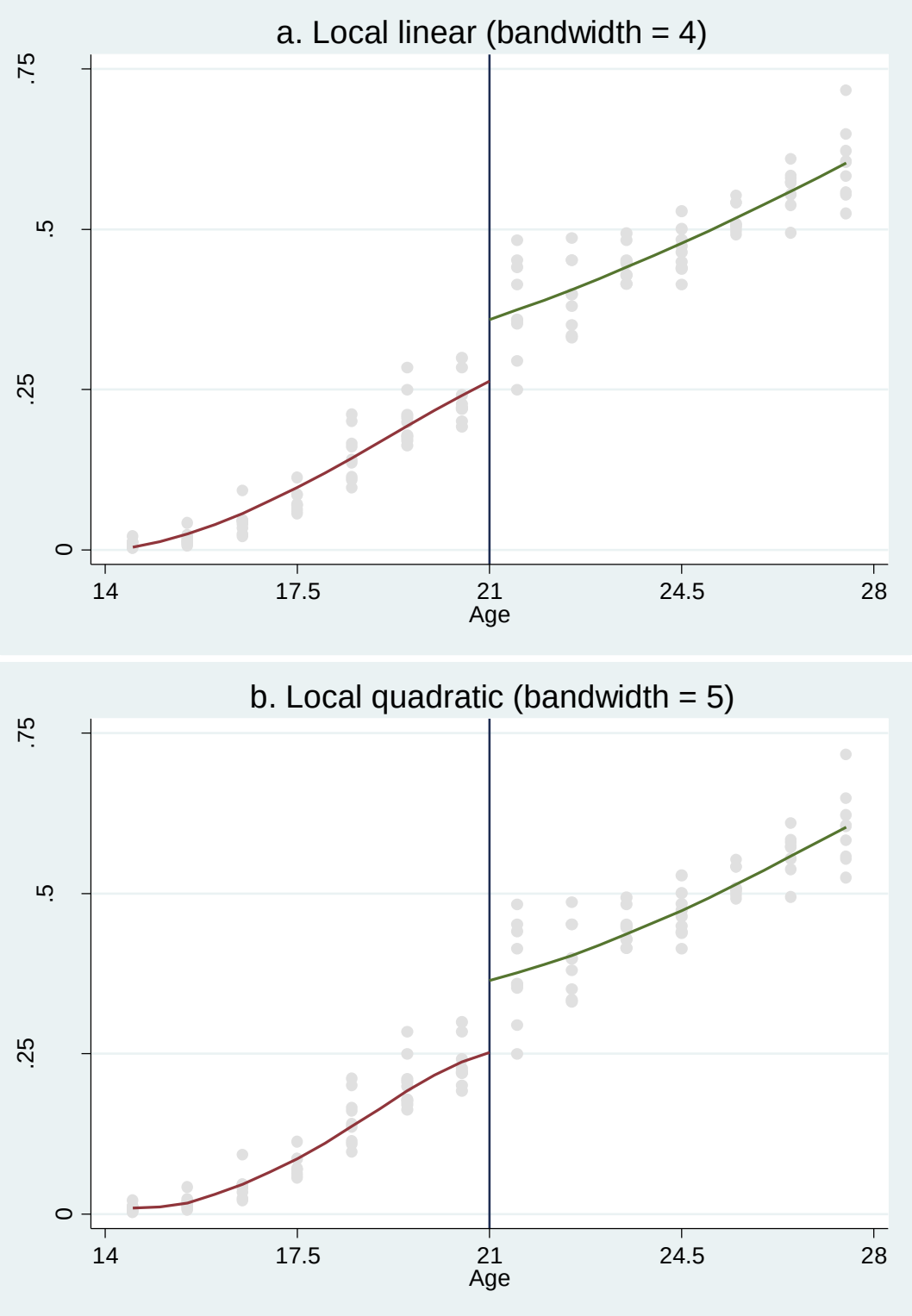
Yörük, Baris K. & Ceren Ertan Yörük, "The Impact of Minimum Legal Drinking Age Laws on Alcohol Consumption, Smoking, and Marijuana Use: Evidence from a Regression Discontinuity Design Using Exact Date of Birth," Journal of Health Economics, 30(4), July 2011, 740–752.

Figure 1 – Age Profiles for Past Month Drinking



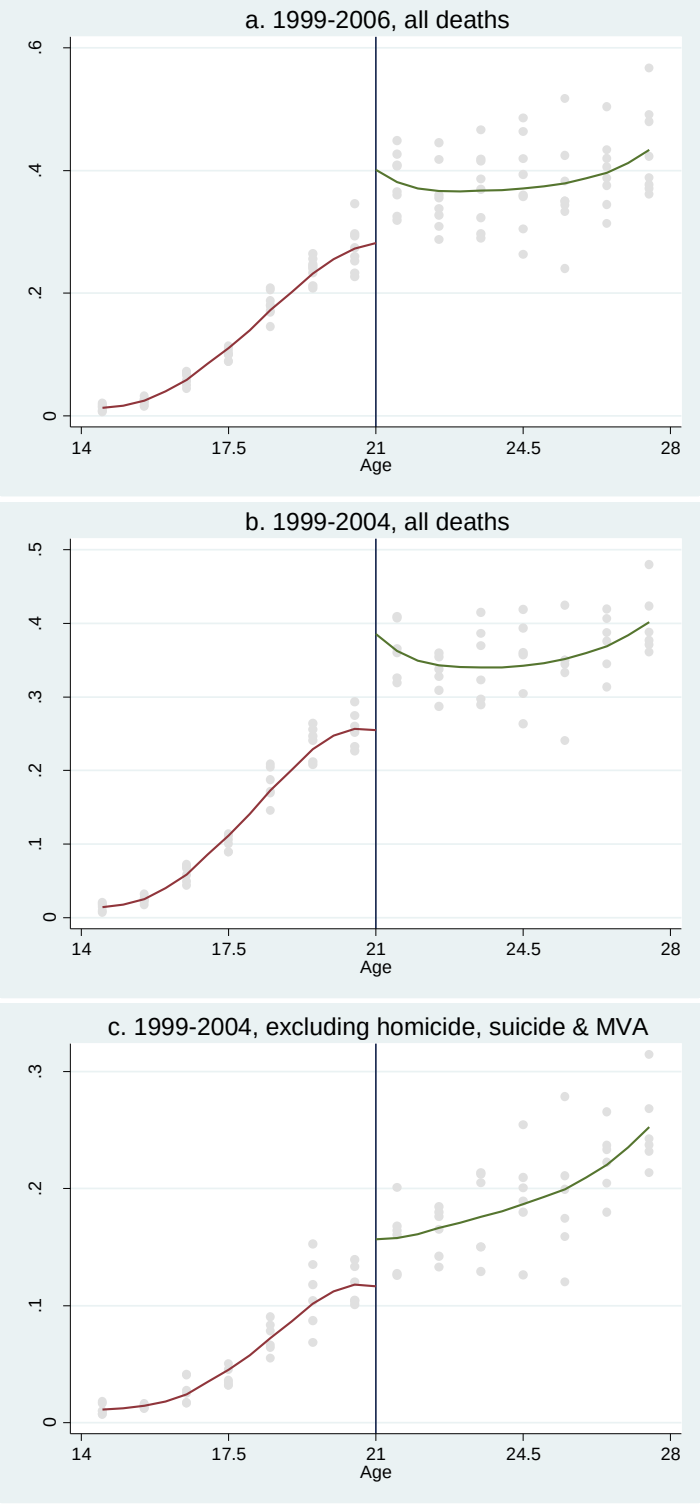
Data are annual national sample-weighted average rates from 2007–2015 for each year of age 12–21 and age groups 22–23, 24–25 and 26–29, coded as (age + 0.5) or the midpoint of the age range. Lines represent local linear regressions on each side of age 21 using a triangular kernel and a bandwidth of 5 age-years, plotted at age intervals of 0.5 years.

Figure 2 – Age Profiles for Alcohol-Induced Death Rate



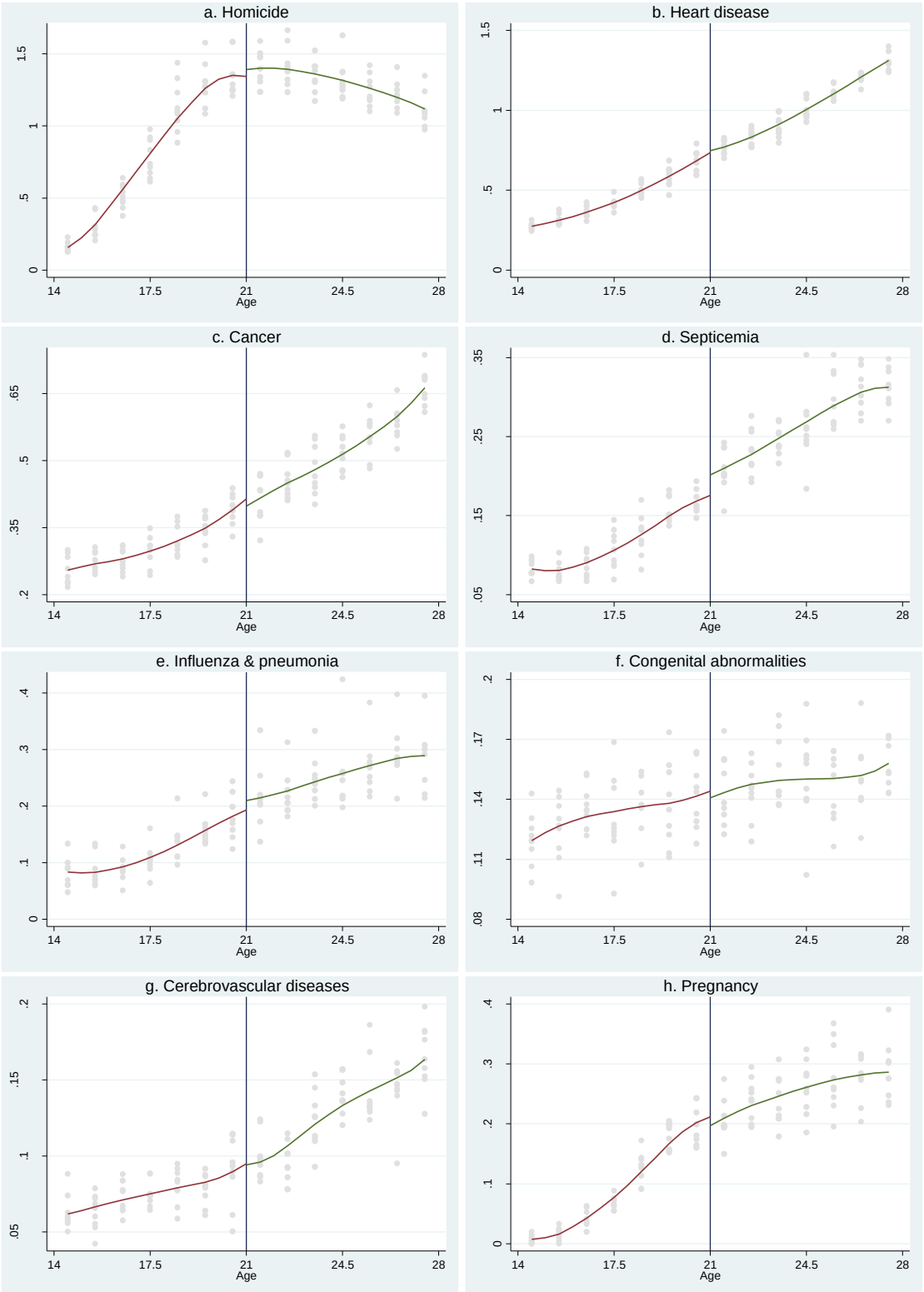
Data are annual national average rates from 2007–2015 for each year of age 14–27, coded as (age + 0.5). Lines represent local regressions on each side of age 21 using a triangular kernel and the indicated bandwidth, plotted at age intervals of 0.5 years.

Figure 3 – Age Profiles for Alcohol-Induced Death Rate, 1999–2006



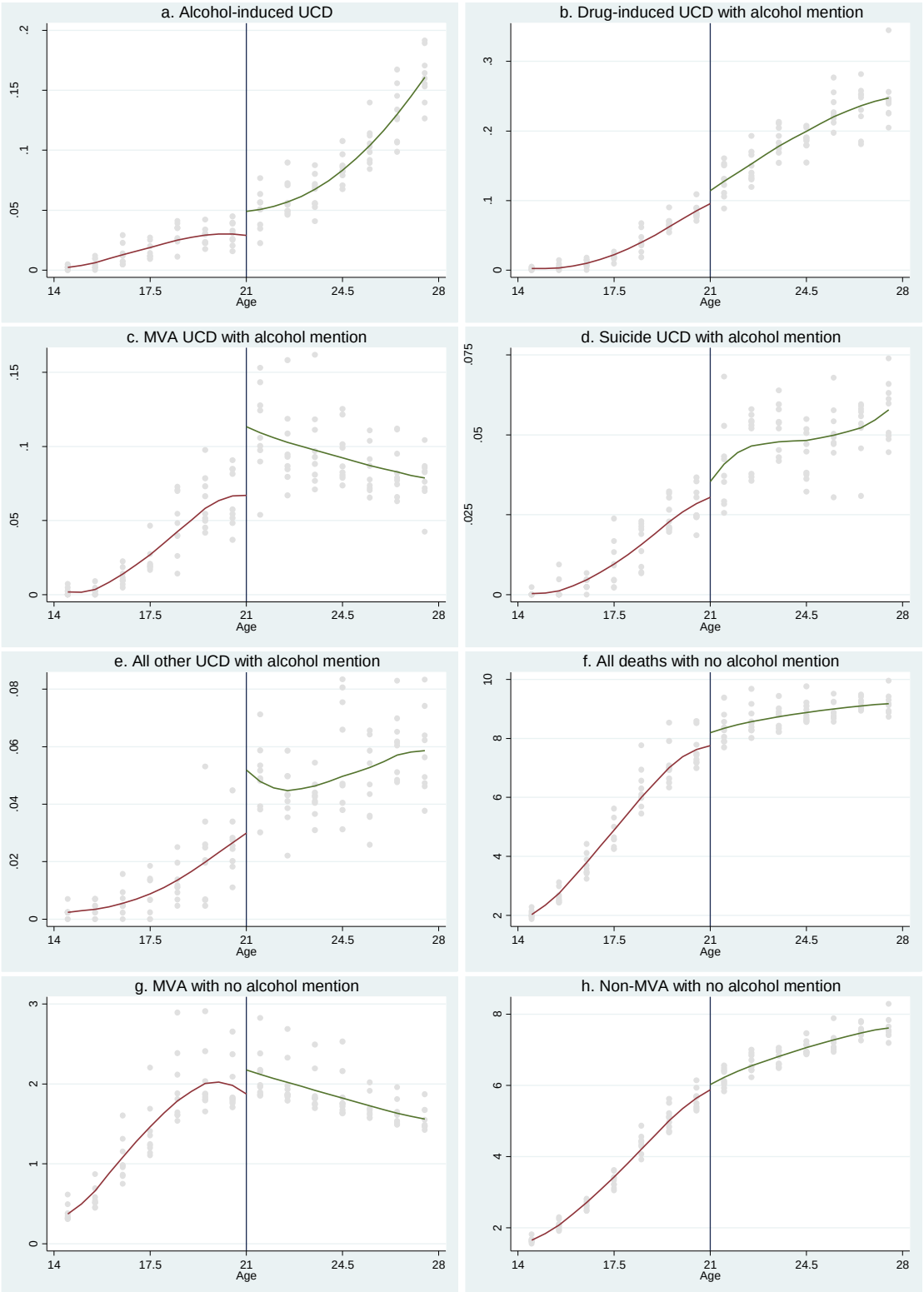
Data are annual national average rates from 1999–2004 (or 2006 in a.) for each year of age 14–27, coded as (age + 0.5). Lines represent local quadratic regressions on each side of age 21 using a triangular kernel and a bandwidth of 4 age-years, plotted at age intervals of 0.5 years.

Figure 4 – Age Profiles for Rates of Death from Other Causes



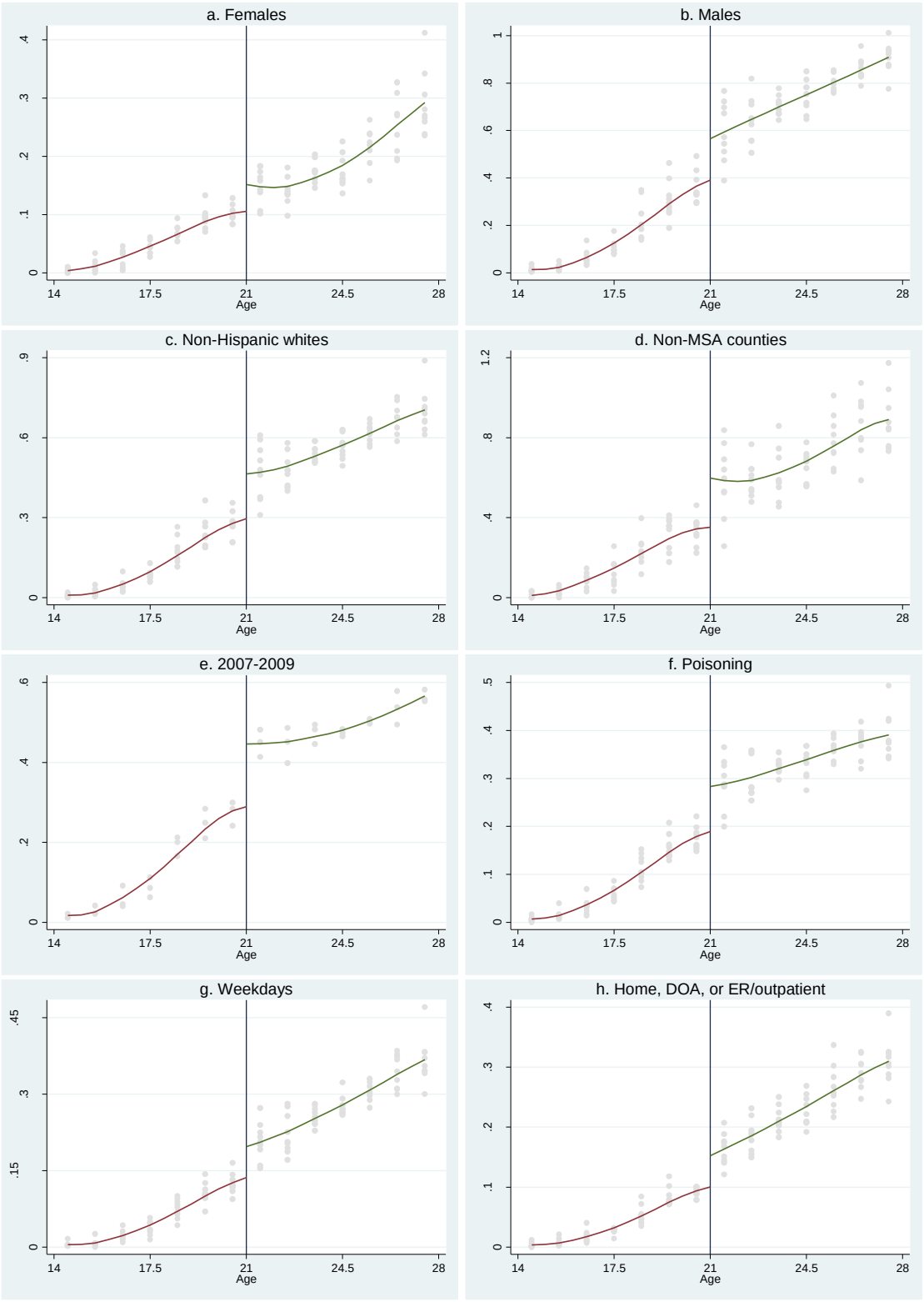
Data are annual national average rates from 2007–2015 for each year of age 14–27, coded as (age + 0.5). Lines represent local quadratic regressions on each side of age 21, using a triangular kernel and a bandwidth of 5 age-years, plotted at age intervals of 0.5 years.

Figure 5 – Age Profiles for Death Rates by Alcohol Mention & Underlying Cause



Data are annual national average rates from 2007–2015 for each year of age 14–27, coded as (age + 0.5). Lines represent local quadratic regressions on each side of age 21, using a triangular kernel and a bandwidth of 5 age-years, plotted at age intervals of 0.5 years.

Figure 6 – Age Profiles for Alcohol-Induced Death Rates by Various Characteristics



Data are annual national average rates from 2007–2015 for each year of age 14–27, coded as (age + 0.5). Lines represent local quadratic regressions on each side of age 21, using a triangular kernel and a bandwidth of 5 age-years, plotted at age intervals of 0.5 years.

Table 1 – Age 21 Increases in Past Month Drinking

	Drank alcohol		Binge drank		Days binge drank	
	LLR	OLS	LLR	OLS	LLR	OLS
Cohort	(1)	(2)	(3)	(4)	(5)	(6)
Full sample	0.115	0.121	0.076	0.085	0.367	0.363
	(0.010)	(0.015)	(0.011)	(0.015)	(0.126)	(0.165)
	0.593		0.403		1.976	
Females	0.119	0.129	0.076	0.089	0.340	0.413
	(0.011)	(0.017)	(0.009)	(0.012)	(0.079)	(0.107)
	0.565		0.338		1.395	
Males	0.113	0.115	0.079	0.082	0.419	0.336
	(0.017)	(0.024)	(0.018)	(0.024)	(0.174)	(0.244)
	0.620		0.466		2.534	
Non-Hispanic whites	0.111	0.121	0.080	0.095	0.430	0.434
	(0.012)	(0.018)	(0.013)	(0.017)	(0.155)	(0.217)
	0.666		0.475		2.505	
Hispanics & nonwhites	0.119	0.120	0.069	0.069	0.264	0.265
	(0.011)	(0.017)	(0.011)	(0.016)	(0.081)	(0.112)
	0.491		0.302		1.219	
MSAs with pop. ≥ 1 mil.	0.106	0.105	0.075	0.082	0.342	0.345
	(0.011)	(0.016)	(0.012)	(0.016)	(0.131)	(0.181)
	0.586		0.383		1.831	
MSAs with pop. < 1 mil.	0.122	0.141	0.074	0.086	0.401	0.420
	(0.014)	(0.020)	(0.012)	(0.018)	(0.145)	(0.208)
	0.616		0.433		2.156	
Outside MSAs	0.131	0.134	0.088	0.094	0.403	0.314
	(0.018)	(0.028)	(0.018)	(0.025)	(0.129)	(0.183)
	0.570		0.409		2.077	

Coefficients represent regression discontinuity (RD) estimates of the increase at age 21 in the drinking outcome indicated in the column heading, for the cohort listed in the row heading, from local linear regressions (LLRs) with a bandwidth of 5 age-years in odd-numbered columns and OLS 6^{th}-order polynomial regressions in even numbered columns. Parentheses contain heteroskedasticity-robust standard errors. Entries beneath LLR regression estimates represent predicted values for drinking outcomes upon turning 21 years old, i.e. from the left along the younger age profile. Data are annual national sample-weighted proportions from the 2007–2015 waves of the National Survey on Drug Use and Health, for each year of age 12–21 and age groups 22–23, 24–25 and 26–30. The sample size is 117, corresponding to 9 years of observations on each of 13 age levels, although LLRs use only observations encompassed by the bandwidth of 5 age-years around age 21, i.e. ages 16–20 on the left and ages 21, 22–23 and 24–25 on the right.

Table 2 – Age 21 Increases in Alcohol-Induced Deaths

Specification	Bandwidth (LLR & LQR) or order/interacted (OLS)			
A. Local linear	2	3	4	5
Triangular kernel	0.119 (0.041)	0.106 (0.033)	0.096 (0.029)	0.091 (0.025)
Epanechnikov kernel	0.119 (0.041)	0.103 (0.032)	0.093 (0.028)	0.089 (0.024)
Rectangular kernel	0.119 (0.041)	0.097 (0.030)	0.088 (0.025)	0.086 (0.022)
Log deaths	0.367 (0.128)	0.305 (0.105)		
B. Local quadratic	3	4	5	6
Triangular kernel	0.138 (0.033)	0.124 (0.045)	0.112 (0.039)	0.102 (0.035)
Log deaths	0.461 (0.181)	0.464 (0.144)	0.422 (0.127)	0.400 (0.113)
C. OLS	5 (No)	3 (Yes)	5 (No, log)	3 (Yes, log)
	0.098 (0.033)	0.117 (0.042)	0.325 (0.120)	0.451 (0.164)

Coefficients represent RD estimates of the increase at age 21 in the rate per 10,000 residents of alcohol-induced deaths using the specification listed in the row heading, and bandwidth (LLR/LQR) or age polynomial order and interaction with the age 21 indicator (OLS) listed in the column heading. LLRs and LQRs for log deaths use a triangular kernel. Parentheses contain heteroskedasticity-robust standard errors. Data are annual national average rates from 2007–2015 for each year of age 14–27, resulting in a sample size of 126 for OLS regressions and 18 × the bandwidth for local regressions (i.e. 9 years × the number of age-years included on each side of age 21). In the unlogged local regressions with triangular kernels, the average predicted death rate upon turning 21 years old (from the left along the younger age profile) is 0.258 in the linear model and 0.249 in the quadratic model.

Table 3 – Age 21 Increases in Alcohol-Induced Deaths, 1999–2006

Period	Log?	Excludes homicide, suicide & MVA?	Local quadratic (bandwidth=3) (1)	Local quadratic (bandwidth=4) (2)	OLS (3)	Predicted at Age 21 (4)
1999–2006	No	No	0.131 (0.046)	0.120 (0.036)	0.143 (0.053)	0.282
1999–2006	Yes	No	0.413 (0.138)	0.418 (0.110)	0.431 (0.120)	
1999–2004	No	No	0.150 (0.037)	0.131 (0.030)	0.155 (0.044)	0.253
1999–2004	Yes	No	0.492 (0.121)	0.472 (0.097)	0.495 (0.114)	
1999–2004	No	Yes	0.050 (0.029)	0.040 (0.022)	0.052 (0.034)	0.112
1999–2004	Yes	Yes	0.361 (0.223)	0.342 (0.160)	0.295 (0.161)	

Entries represent RD estimates of the increase at age 21 in the rate per 10,000 residents of alcohol-induced deaths for the period and specification listed in the row headings and regression model listed in the column headings in columns 1–3, and the average predicted death rate upon turning 21 years old (from the left along the younger age profile) from the displayed LQRs in column 4. LQRs use a triangular kernel while OLS regressions include a 4^{th}-order polynomial in age (or 3^{rd}-order for log models) fully interacted with the age 21 indicator. Parentheses contain heteroskedasticity-robust standard errors. Data are annual national average rates from 1999–2004 (or 2006 in the first 2 rows) for each year of age 14–27, resulting in a sample size of 84 (or 112) for OLS regressions and 12 (or 16) × the bandwidth for LQRs (i.e. 6 or 8 years × the number of age-years included on each side of age 21).

Table 4 – Changes in Alcohol-Induced Deaths at Ages Other Than 21

Bandwidth:		2	3	4	5	6	OLS	At age
Age	Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)
17	LLR	–0.019	–0.010				–0.010	0.055
		(0.014)	(0.011)				(0.017)	
	LQR		–0.031					0.063
			(0.023)					
18	LLR	0.030	0.038	0.030			0.016	0.088
		(0.022)	(0.017)	(0.016)			(0.019)	
	LQR		0.018	0.049				0.094
			(0.031)	(0.025)				
19	LLR	0.004	–0.014	–0.009	0.005		–0.020	0.170
		(0.027)	(0.024)	(0.020)	(0.018)		(0.021)	
	LQR		0.031	–0.016	–0.031			0.193
			(0.038)	(0.032)	(0.029)			
20	LLR	–0.070	–0.044	–0.020	0.002	0.022	–0.024	0.230
		(0.029)	(0.022)	(0.021)	(0.021)	(0.021)	(0.026)	
	LQR		–0.110	–0.088	–0.074	–0.059		0.239
			(0.047)	(0.032)	(0.027)	(0.024)		
22	LLR	–0.075	–0.040	–0.020	–0.012	–0.005	–0.027	0.410
		(0.045)	(0.038)	(0.032)	(0.029)	(0.026)	(0.036)	
	LQR		–0.128	–0.082	–0.052	–0.039		0.452
			(0.059)	(0.050)	(0.044)	(0.040)		
23	LLR	0.027	–0.011	–0.024	–0.025		–0.027	0.435
		(0.032)	(0.025)	(0.023)	(0.021)		(0.025)	
	LQR		0.085	0.021	–0.008			0.406
			(0.050)	(0.036)	(0.030)			
24	LLR	–0.029	–0.021	–0.032			–0.033	0.469
		(0.023)	(0.019)	(0.017)			(0.022)	
	LQR		–0.041	–0.006				0.462
			(0.037)	(0.028)				
25	LLR	0.018	0.012				0.034	0.479
		(0.021)	(0.018)				(0.024)	
	LQR		0.026					0.465
			(0.031)					

Coefficients represent RD estimates of the increase in the rate per 10,000 residents of alcohol-induced deaths at the age listed, using LLR or LQR with a triangular kernel and the indicated bandwidth in columns 1–5, and OLS with a 5^{th}-order age polynomial in column 6. Parentheses contain heteroskedasticity-robust standard errors. Column 7 represents the average LLR or LQR predicted death rate upon turning the corresponding age (from the left along the younger age profile) among the listed bandwidths. Data are annual national average rates from 2007–2015 for each year of age 14–27, resulting in a sample size of 126 for OLS regressions and 18 × the bandwidth for local regressions (i.e. 9 years × the number of age-years included on each side of age 21).

Table 5 – Age 21 Increases in Log Population

Cohort	LLR Bandwidth = 5 (1)	LQR Bandwidth = 6 (2)	OLS 5^{th} order (3)	OLS 3^{rd} order interacted (4)
Total	–0.002	0.002	–0.000	0.004
	(0.014)	(0.019)	(0.018)	(0.024)
Females	–0.003	0.002	–0.001	0.004
	(0.014)	(0.019)	(0.018)	(0.024)
Males	–0.001	0.003	0.000	0.004
	(0.014)	(0.019)	(0.018)	(0.024)
Non-Hispanic whites	–0.000	0.003	0.001	0.004
	(0.012)	(0.016)	(0.016)	(0.021)
Hispanics & nonwhites	–0.004	0.002	–0.002	0.004
	(0.027)	(0.038)	(0.036)	(0.047)
Counties with central cities in MSAs ≥ 1 million	–0.011	–0.014	–0.016	–0.006
	(0.015)	(0.021)	(0.021)	(0.026)
Other MSA counties	0.005	0.010	0.007	0.007
	(0.016)	(0.023)	(0.022)	(0.028)
Outside MSAs	–0.001	0.013	0.009	0.010
	(0.015)	(0.020)	(0.020)	(0.026)

Coefficients represent RD estimates of the increase at age 21 in the population of the group listed in the row heading, using the model indicated in the column heading. Local regressions use a triangular kernel and bandwidths representing the median integer value among the Calonico et al. (2014), Imbens & Kalyanaraman (2012) and Ludwig & Miller (2007) procedures for the total population regressions, while OLS regressions include an age polynomial of order 5 (column 3) or order 3 fully interacted with the age 21 indicator (column 4). Parentheses contain heteroskedasticity-robust standard errors. Data are annual national average rates from 2007–2015 for each year of age 14–27, resulting in a sample size of 126 for OLS regressions and 18 × the bandwidth for local regressions (i.e. 9 years × the number of age-years included on each side of age 21).

Table 6 – Age 21 Increases in Deaths with Other Causes Mentioned

Cause	LLR Bandwidth = 4 (1)	LQR Bandwidth = 5 (2)	OLS 5th order (3)	OLS 3rd order interacted (4)	Predicted at age 21 (5)
Homicide	–0.031	0.048	0.006	0.101	1.397
	(0.068)	(0.091)	(0.079)	(0.102)	
Heart disease	0.003	0.011	0.003	0.013	0.731
	(0.028)	(0.039)	(0.034)	(0.043)	
Cancer	–0.005	–0.016	–0.015	–0.030	0.409
	(0.021)	(0.030)	(0.025)	(0.033)	
Septicemia	0.022	0.026	0.026	0.037	0.177
	(0.011)	(0.015)	(0.014)	(0.017)	
Influenza & pneumonia	0.013	0.017	0.014	0.023	0.193
	(0.023)	(0.031)	(0.027)	(0.035)	
Congenital abnormalities	–0.001	–0.003	–0.002	–0.007	0.143
	(0.008)	(0.011)	(0.010)	(0.013)	
Cerebrovascular diseases	–0.003	–0.001	–0.007	–0.006	0.093
	(0.009)	(0.013)	(0.011)	(0.014)	
Pregnancy	–0.022	–0.015	–0.024	–0.011	0.218
	(0.016)	(0.023)	(0.020)	(0.026)	
Diabetes	–0.007	–0.005	–0.003	0.002	0.082
	(0.005)	(0.007)	(0.006)	(0.008)	
Chronic respiratory diseases	0.005	0.003	0.006	0.005	0.069
	(0.007)	(0.010)	(0.009)	(0.011)	
Anemias	0.004	0.008	0.007	0.015	0.055
	(0.005)	(0.007)	(0.006)	(0.008)	
HIV	0.006	0.004	0.006	0.005	0.026
	(0.006)	(0.008)	(0.008)	(0.010)	

Coefficients represent RD estimates of the increase at age 21 in the rate per 10,000 residents of deaths with mentions of the cause listed in the row heading, using the model indicated in the column heading. Local regressions use a triangular kernel and bandwidths representing the median integer value among the Calonico et al. (2014), Imbens & Kalyanaraman (2012) and Ludwig & Miller (2007) procedures for the baseline regressions in Table 2, while OLS regressions include an age polynomial of order 5 (column 3) or order 3 fully interacted with the age 21 indicator (column 4). Column 5 represents the average of the LLR and LQR predicted death rates upon turning 21 years old (from the left along the younger age profile). Parentheses contain heteroskedasticity-robust standard errors. Data are annual national average rates from 2007–2015 for each year of age 14–27, resulting in a sample size of 126 for OLS regressions and 18 × the bandwidth for local regressions (i.e. 9 years × the number of age-years included on each side of age 21).

Table 7 – Age 21 Increases in Deaths by Alcohol Mention & Underlying Cause

Category	LLR Bandwidth = 4 (1)	LQR Bandwidth = 5 (2)	OLS 5th order (3)	OLS 3rd order interacted (4)	Predicted at age 21 (5)
A. Alcohol Mentions	0.096	0.112	0.098	0.117	0.258
	(0.029)	(0.039)	(0.033)	(0.042)	
Alcohol-induced	0.013	0.020	0.017	0.021	0.031
	(0.006)	(0.009)	(0.008)	(0.010)	
Drug-induced	0.020	0.019	0.016	0.020	0.096
	(0.009)	(0.013)	(0.011)	(0.015)	
Motor vehicle accident	0.038	0.046	0.039	0.048	0.070
	(0.012)	(0.017)	(0.014)	(0.019)	
Suicide	0.006	0.003	0.004	0.002	0.026
	(0.004)	(0.006)	(0.005)	(0.006)	
All others	0.020	0.023	0.022	0.026	0.035
	(0.006)	(0.008)	(0.007)	(0.009)	
B. No Alcohol Mention	0.156	0.444	0.277	0.619	7.929
	(0.278)	(0.377)	(0.321)	(0.413)	
Drug-induced	–0.010	–0.041	–0.036	–0.034	1.018
	(0.035)	(0.050)	(0.051)	(0.068)	
Motor vehicle accident	0.086	0.300	0.184	0.382	1.978
	(0.161)	(0.221)	(0.190)	(0.244)	
Suicide	0.053	0.049	0.061	0.061	1.247
	(0.049)	(0.065)	(0.058)	(0.075)	
All others	0.027	0.136	0.068	0.210	3.687
	(0.159)	(0.218)	(0.184)	(0.239)	

Coefficients represent RD estimates of the increase at age 21 in the rate per 10,000 residents of deaths from the underlying cause listed in the row heading, with and without an alcohol-induced cause also mentioned in panel A and B, respectively, using the model indicated in the column heading. Local regressions use a triangular kernel and bandwidths representing the median integer value among the Calonico et al. (2014), Imbens & Kalyanaraman (2012) and Ludwig & Miller (2007) procedures for the baseline regressions in table 2, while OLS regressions include an age polynomial of order 5 (column 3) or order 3 fully interacted with the age 21 indicator (column 4). Column 5 represents the average of the LLR and LQR predicted death rates upon turning 21 years old (from the left along the younger age profile). Parentheses contain heteroskedasticity-robust standard errors. Data are annual national average rates from 2007–2015 for each year of age 14–27, resulting in a sample size of 126 for OLS regressions and 18 × the bandwidth for local regressions (i.e. 9 years × the number of age-years included on each side of age 21).

Table 8 – Age 21 Increases in Alcohol-Induced Deaths for Subsamples

Panel	Group	LLR Bandwidth = 4 (1)	LQR Bandwidth = 5 (2)	OLS 5th order (3)	OLS 3rd order interacted (4)	Predicted at age 21 (5)
A.	Females	0.028	0.046	0.037	0.053	0.109
		(0.012)	(0.016)	(0.015)	(0.019)	
	Males	0.161	0.175	0.156	0.178	0.399
		(0.049)	(0.067)	(0.056)	(0.072)	
B.	Non-Hispanic whites	0.141	0.167	0.145	0.177	0.303
		(0.039)	(0.053)	(0.045)	(0.058)	
	Hispanics & nonwhites	0.032	0.032	0.031	0.032	0.195
		(0.021)	(0.029)	(0.024)	(0.031)	
C.	With central cities in MSAs ≥ 1 million	0.071	0.076	0.074	0.082	0.185
		(0.022)	(0.030)	(0.026)	(0.033)	
	Other MSA counties	0.083	0.096	0.077	0.093	0.268
		(0.033)	(0.046)	(0.038)	(0.049)	
	Outside MSAs	0.190	0.245	0.220	0.281	0.366
		(0.066)	(0.092)	(0.077)	(0.101)	
D.	2007–2009	0.132	0.157	0.134	0.161	0.299
		(0.025)	(0.033)	(0.031)	(0.039)	
	2010–2015	0.079	0.089	0.080	0.095	0.237
		(0.027)	(0.037)	(0.031)	(0.040)	
E.	Poisoning (X45, X65, Y15)	0.081	0.094	0.080	0.097	0.194
		(0.020)	(0.027)	(0.023)	(0.029)	
	All other causes	0.032	0.032	0.035	0.036	0.114
		(0.018)	(0.024)	(0.020)	(0.026)	
F.	Monday through Friday	0.053	0.061	0.055	0.065	0.138
		(0.015)	(0.020)	(0.017)	(0.022)	
	Saturday & Sunday	0.043	0.051	0.043	0.052	0 .119
		(0.017)	(0.023)	(0.019)	(0.025)	
G.	Home, DOA, or ER/outpatient	0.047	0.052	0.048	0.056	0.103
		(0.010)	(0.013)	(0.012)	(0.015)	
	All other places	0.049	0.060	0.050	0.062	0.155
		(0.020)	(0.028)	(0.023)	(0.030)	

The table is constructed analogously to Tables 6 and 7. County groups reflect the 2013 6-category NCHS classification scheme. In panel E, other causes include ICD–10 codes F10 (mental and behavioral disorders), G31.2 (nervous system degeneration), I42.6 (cardiomyopathy), K29.2 (gastritis), K70 (liver-related), and K85.2 and K86.0 (pancreatitis), all due to alcohol, plus R78.0 (finding of alcohol in blood), and deaths with a cause from each row are included in both. Panel F excludes one death on an unknown day. Other places in panel G include inpatient, hospice, nursing home/long-term care and unknown facilities, and "other" places. The entire population is used to construct rates in panels E–G.