Liquid Assets? The Short-Run Liabilities of Binge Drinking[*]

Marco Francesconi
University of Essex

Jonathan James
University of Bath

June 15, 2018

Abstract

We estimate the effect of binge drinking on road accidents, accident and emergency (A&E) attendances, and arrests using a variety of unique English data and a two-sample instrumental variables estimation procedure. Drinking 10 or more units of alcohol in a single session increases road accidents by 18.6%, injury-related A&E attendances by 6.6%, and arrests by 72%. The marginal increase from 8 to 10 or more alcoholic units implies nearly 6,100 extra road accidents every year, 63,000 additional A&E attendances, and 100,000 additional arrests. The externality per mile driven by a binge drinker is about 5 pence and the punishment that internalizes this externality is equivalent to a fine of £22,800 per drunk driving arrest.

I. Introduction

Background and Motivation — More than two and a half billion people worldwide are alcohol users. With a global death toll of 5.9% (approximately 3.3 million deaths every year) and a burden of disease of 5.1% (about the same as that caused by tobacco), harmful use of alcohol has been identified as one of the leading preventable causes of death and a key risk factor for chronic diseases, injuries, and cancer around the world, especially in advanced economies (World Health Organization 2010, 2014; Stewart and Wild 2014).

Heavy episodic drinking (HED), commonly known as "binge drinking", is an alcohol abuse pattern characterized by episodes of heavy social drinking followed by periods of abstinence; it often results in acute impairment and is believed to contribute to a substantial proportion of alcohol related deaths and injuries. Some countries face high binge drinking rates. For instance, between 10% and 20% of all adult men binge drink (i.e., drink the alcohol equivalent of at least one bootle of wine or four pints of beer in one session) once a week in Canada, Australia, Germany, and the United States, while the rates for English men are staggering, at around 30%, as are those for women at about 20%. Almost everywhere, bingeing is far more prevalent among young people than among older individuals.

Past, mainly medical, research has attempted to document the effects of HED on various life domains (Wechsler et al. 1994; Single et al. 1999; Miller et al. 2007). While these studies are valuable, they suffer from a number of important limitations. First, they refer to specific subgroups of the youth population, such as college students, and do not attempt to deal with sample selection issues. The extent to which their conclusions are generalizable is thus unknown. Second, by covering only few, very particular outcome domains — such as alcohol poisoning, ischaemic heart disease, alcoholic liver cirrhosis, and alcohol-related suicides — they cannot provide a broad picture of the consequences of bingeing. Third, they are correlational and lack a credible research design, in the sense they fail to isolate any source of variation in bingeing behavior that is likely to reveal causal effects on outcomes.

In this paper, we adopt a radically different strategy for estimating the causal effects of HED. Specifically, we rely on a variant of the two-sample instrumental variables procedure developed by Angrist and Krueger (1992) and Arellano and Meghir (1992) using a variety of unique data from England. We use this procedure because measures of alcohol involvement are generally not available from the same data sources that collect information on outcomes. This method combines the first stage and reduced form results to produce estimates of the effect of bingeing on a range of outcomes obtained from multiple sources.

An important contribution of our study is to develop a strategy for estimating the impact of HED on several outcomes, more than in previous studies, without relying on policy changes that typically affect alcohol availability only for specific age groups. Our analysis combines two sources of variation. The first is the *difference* in age among alcohol consumers, and the second is the *difference* in the timing of alcohol consumption. Our instrument is the interaction between such two differences. This is similar to the idea used by Lindo, Siminski, and Swensen (2018) who consider the effect of American football college games, which are expected to intensify partying and drinking among college students,

on the incidence of rape at universities with top football teams. In our case, the population of interest is all young individuals, not just college students, while our time variation compares weekends to weekdays, and not game days to non-game days.

On the age variation, extensive medical research eloquently motivates the importance of the age gradient in alcohol consumption patterns: young individuals (e.g., those aged less than 30) are substantially more likely to experience HED than older people (e.g., those aged 50 or more). This pattern, which has been documented for several countries, emerges very strongly in England (Naimi et al. 2003; Williamson et al. 2003; Grant et al. 2004). On the time variation, many studies have emphasized the social aspects of HED (Cook and Moore 2000) as well as the need for group time coordination that favors weekends over weekdays for alcohol consumption across all developed countries, including Britain (Simpson, Murphy, and Peck 2001; Parker and Williams 2003; Taylor et al. 2010).[1] Indeed, drinkers in England are twice as likely to engage in HED on Fridays and Saturdays as opposed to the other days of the week. Notice that, although the double difference in consumers' age and in the time of alcohol consumption is our identifying variable, age and day of week are expected to influence our outcomes of interest directly. Age and day of the week therefore are determinants of both HED and outcomes, while the interaction of the differences in age and time affects outcomes only through its effect on bingeing.

Another way to motivate our instrument is in terms of the cue-theory of consumption. The combination of early adulthood and weekends generates the ideal conditions in which salient environmental cues (such as the smell and sight of alcohol) and social stimuli (such as the sight of familiar "wet" milieux and feed-forward mechanisms associated with near-term gratification and pleasure, or with the display of physical prowess and agonistic behavior) induce alcohol cravings and lead to bingeing (Domjan, Cusato, and Villareal 2000; Laibson 2001; Bernheim and Rangel 2004; Gul and Pesendorfer 2007; Barker and Taylor 2014). In Section II we shall return to the instrument and the identification strategy in greater detail. In particular, we provide clean evidence to deal with issues related to other risky behaviors and self-selection into several activities other than drinking that young adults can be involved in during the weekend.

Our study is the first to provide a broad picture of the consequences of bingeing estimating the impact of HED on road accidents, accident and emergency (A&E) attendances, and arrests. We find that a binge of just over 4 glasses of wine (or 4 pints of beer) increases the daily mean number of road accidents by 18.6%, the average number of daily injury-related A&E attendances by 6.6%, and the average number of arrests for all alcohol-related incidences by another 72%. Such estimates translate into approximately 6,100 extra road accidents (including 300 additional fatalities), 63,000 additional A&E attendances, and 100,000 extra arrests every year.

Literature — Not only are these effects sizeable, they are also new. Economic research on binge

---

[1]Hamermesh, Myers, and Pocock (2008) document that complementarities in consuming goods and services generate time coordination of individual economic activities. Another likely form of group coordination is over places (e.g., bars, pubs, restaurants, dorms, and private homes). Although this does not play a role in our main analysis, we will go back to it as a robustness exercise.

drinking, in fact, is not very systematic. More attention has been given to problem (excessive) drinking — which does not necessarily coincide with bingeing, since the latter has a strong social dimension and is concentrated in short time episodes — and to specific groups, such as teenagers and college students. The most authoritative survey to date by Cawley and Ruhm (2012) discusses binge drinking in the context of other risky behaviors, and provides a brief review of the slender set of studies on HED. Much of this research identifies the effect of drinking through policy changes that affect alcohol availability, such as state variation in alcohol (or "beer") taxes, changes in the minimum legal drinking age (MLDA), and "dry laws", which ban alcohol sales at certain times of the day, or on certain days of the week, or in specific premises.[2] As mentioned, one of our contributions is to estimate the effects of bingeing without relying on such policies.

A large body of empirical research looks at the effect of alcohol consumption (but not necessarily HED) on a wide range of outcomes. The existing evidence suggests that as alcohol consumption increases so do motor vehicle accidents and alcohol related deaths (Chaloupka, Saffer, and Grossman 1993; Ruhm 1996; Dee 1999; Carpenter and Dobkin 2009, 2011), fatal and nonfatal injuries and A&E admissions (Levitt and Porter 2001; Adams, Blackburn, and Cotti 2012; Marcus and Siedler 2015; Carpenter and Dobkin 2017; Hansen 2015), arrests and crime (Biderman, De Mello, and Schneider 2010; Heaton 2012; Carpenter and Dobkin 2015; Anderson, Crost, and Rees 2016), teen childbearing and abortions (Dee 2001), child abuse (Markowitz and Grossman 2000), and work days lost due to industrial injuries (Ohsfeldt and Morrisey 1997), while educational attainment goes down (Carrell, Hoekstra, and West 2011) as well as employment (Terza 2002).[3] Whether such effects are driven by alcohol consumption in general or HED in particular remains to be seen. This qualification is part of our work. In addition, only few studies focus on the peer effect side of binge drinking (Duncan et al. 2005; Kremer and Levy 2008).

Our analysis is perhaps more closely related to two existing studies. The first is the work by Carpenter, Dobkin, and Warman (2016), which like ours, considers the whole distribution of alcohol consumption and focuses on binge drinking. Using the universe of Canadian mortality records, this work documents that MLDA reduces mortality rates of young men (especially of extreme drinkers) but has much smaller effects on women. Carpenter and colleagues take these results as an indication that alcohol control policies can be effective, particularly if they are designed to reduce the number of episodes of heavy drinking, as opposed to lower intensities of drinking. The second study is the already mentioned work by Lindo, Siminski, and Swensen (2018). Using crime data from the National Incident Based Reporting System, their results indicate that Division 1 college American football

---

[2]Nelson (2015) provides a comprehensive review of studies about the effect of alcohol prices (or tax surrogates) on binge drinking by gender and age groups. The key result of this work is that binge drinkers are not highly responsive to increased prices and, as a result, increased alcohol taxes or prices are unlikely to be effective as a means to reduce HED, regardless of gender or age group.

[3]For almost each outcome domain there are exceptions. For instance, using a two-sample instrumental variables procedure similar to ours, Dee and Evans (2003) find that teenage drinking has small and statistically insignificant effects on educational attainment, although students who face an MLDA of 18 are more likely to binge as opposed to their counterparts with an MLDA of 21. Another example is the work by Lindo, Siminski, and Yerokhin (2016), which applies an age-based regression discontinuity design to Australian data and finds no evidence that legal access to alcohol has an impact on motor vehicle accidents.

games increase reports of rape involving college-aged victims by 28%. They find similar effects on crimes associated with excessive partying that is likely to be associated with football games, such as DUI, drunkenness, and liquor-law violations.

Restrictions on alcohol availability and minimum purchase ages, however, may have unintended consequences. One is that they could lead youths to switch from alcohol to illicit drugs. DiNardo and Lemieux (2001) estimate that raising the state MLDA from 18 to 21 increases marijuana consumption among youths. Similarly, Crost and Guerrero (2012), Anderson, Hansen, and Rees (2013), and Dragone et al. (forthcoming) find evidence that marijuana and alcohol are substitutes, and Conlin, Dickert-Conlin, and Pepper (2005) show that alcohol access and illicit-drug-related crimes (and not just those related to marijuana) are substitutes. We shall return to this point when we discuss the identification strategy in the next section.

Policy Relevance — We use our estimates of the effect on road accidents to monetize the externality of binge drinking.[4] A conservative estimate of this externality is £1.52 billion per year, which is equivalent to 5.3 pence per mile driven. This is about one-third of the Levitt and Porter's (2001) estimated externality due to drunk driving in the United States of 17 pence per mile driven expressed in 2015 prices (or 15 cents in 1993 prices as in their published work). At current arrest rates for drunk driving in England, the Pigouvian tax that internalizes our estimated externality is approximately £22,800 ($35,000 in 2015 prices) per drunk driving arrest. This is about eight times more than the maximum fine currently set by the UK government on drunk driving. If instead one considers the most frequent type of arrests in England during our sample period, i.e., arrests due to violent offenses, the punishment that internalizes the HED externality is equivalent to a fine of £5,000 per arrest.

The remainder of the paper is structured as follows. Section II presents the econometric method and discusses the key identification issues. Section III describes the variety of data sources used in the analysis, discusses the definitions of heavy episodic drinking, and shows the first-stage results. After the reduced form estimates, Section IV presents the benchmark estimates, shows the results from several robustness checks and overidentification exercises, and reports evidence on accident-related externalities associated with binge drinking. Section V concludes. Supplementary material on the data and additional results discussed throughout the paper are available in the Online Appendix.

II. Empirical Framework

II.A. Setup

Most data sets that measure drinking status do not collect information on individual outcomes, while most data sets that have information on economically salient outcomes do not record drinking measures. We therefore employ a two-sample procedure which requires only one data set with information on binge drinking status and another data set with information on outcomes. Based on a variant of

---

[4]In this analysis we focus on road accidents only because the road accident data are administrative records with full national coverage.

the two-sample instrumental variables (TSIV) methodology developed by Angrist and Krueger (1992) and Arellano and Meghir (1992), this procedure combines the first-stage and reduced form results to generate estimates of the effect of binge drinking on outcomes.[5]

Let Y_{it} be a given outcome for individual i at time t, B_{it} denote the endogenous binge drinking status, and z_{it} be our instrument. The limited information form representation of the model is given by

$$Y_{it} = \beta_1 B_{it} + \mathbf{X}'_{it}\lambda + \varphi_t + \epsilon_{it} \tag{1}$$
$$B_{it} = f(z_{it}; \phi) + \mathbf{H}'_{it}\gamma + \varphi_t + \nu_{it}, \tag{2}$$

where $f(\cdot)$ is a function of the instrument, $\mathbf{X}_{it}$ and $\mathbf{H}_{it}$ are vectors of (possibly different) covariates (such as age, sex, education, region of residence, and number of licenses per 1,000 individuals), φ_t are time (year and quarter of the year) fixed effects, and ϵ_{it} and ν_{it} are random shocks. The possibility of different variables in $\mathbf{X}$ and $\mathbf{H}$ reflects the fact that they are drawn from two different samples. We will be more specific on them in Sections III and IV. Our parameter of interest is β_1, which measures the effect of binge drinking on Y.

The first-stage data set is the Health Survey for England (HSE), which has information on B_{it}, $\mathbf{H}_{it}$ and z_{it}, but not on Y_{it}. The instrument used in (2) is the interaction between the difference in age among consumers (young and old) and the difference in drinking times (day of the week and time of the day). Assuming away the role of covariates $\mathbf{H}_{it}$ and time effects φ_t for simplicity, assuming $f(\cdot)$ is additive and linear in parameters, and letting $z_{it} = a_{it} \times w_{it}$, the first stage equation (2) is then given by

$$B_{it} = \phi_1(a_{it} \times w_{it}) + g(A_{it}; \phi_2) + h(W_{it}; \phi_3) + \nu_{it}, \tag{3}$$

where a_{it} is equal to 1 if i's age at time t is between 18 and 30 years ("young") and equal to zero if the individual is aged 50 or more ("old"), and w_{it} is equal to 1 if individual i has the heaviest drinking experience of the week during the "weekend" at time t and zero otherwise.[6] A_{it} is the age (in years) of individual i at time t, while W_{it} is the day of the week (not just weekend days) in which i reports drinking at time t, and $g(\cdot)$ and $h(\cdot)$ are flexible functions of age and day of the week, respectively. These last two terms are specified in order to emphasize that age and day of week are allowed to affect bingeing directly: that is, habitual ethanol seeking can occur in any day during the week and at all ages.

As such, the estimated effect of bingeing on outcomes β_1, which is driven by differences between the young and the old *and* between weekends and weekdays, should be thought of as identifying the effects of social activities associated with young adults in weekends, above and beyond what is expected on

---

[5]Dee and Evans (2003) perform another application of the TSIV method related to alcohol consumption and its effect on teen education. See Inoue and Solon (2010) for a discussion of the links between the TSIV estimator and its two-sample two-stage least squares variant and their standard errors.

[6]In the baseline analysis, individuals aged 31–49 are excluded. But these (and all the data) are used in robustness checks in a variety of combinations. See subsection IV.C.

each specific day of the week and for each separate age group. This is similar to the estimation strategy adopted by Lindo, Siminski, and Swensen (2018), who identify the effect of American football games on the incidence of rape by comparing reports of rape on game days to reports on non-game days, while controlling for differences expected across different days of the week and across different times of the year.[7]

Medical and social science research strongly supports the arguments behind the choice of our instrument, besides heaps of anecdotal evidence, literary references, popular movies, and everyday news accounts. Recent work in developmental and social psychology and neuroscience emphasizes that early adulthood is a critical, transitional period from the late teens through the twenties characterized by identity explorations, self-focus, and instability (Arnett 2000). In most advanced societies, substance (in particular, alcohol) abuse is one of the experiences that many young adults go through to explore their identities, affirm their personality, and cope with change and uncertainty (over relationships, friendships, residential location, and work). Furthermore, drinking is a powerful social dimension that relies on group coordination. In "wet" environments, where alcohol is readily accessible in a variety of ways and most people drink, young adults will tend to drink more than they otherwise would or may be initiated into HED relatively easily (Becker 1996; Cook and Moore 2000).[8] Weekends, when most people do not work or do not have college classes to attend, offer a convenient, natural device for time coordination over leisure activities that may involve heavy drinking (e.g., partying and clubbing). With fewer responsibilities and commitments (other than work or study), young adults will disproportionately use weekends to synchronize their leisure schedules.[9]

The interaction between early adulthood and weekends, $a_{it} \times w_{it}$ in (3), identifies what Bernheim and Rangel (2004) call "hedonic forecasting mechanism" that is responsible "for associating environmental cues with forecasts of short-term hedonic (pleasure/pain) responses" (p. 1562) among some individuals. In line with the cue-theory of consumption (Laibson 2001), as environmental cues — e.g., approaching 5 p.m. on Fridays and the smell and sight of alcohol — are paired with social stimuli — e.g., going to the pub on Friday evenings and clubbing on Saturday nights — alcohol seeking may become habitual (Barker and Taylor 2014; Barker et al. 2015). That is, alcohol seeking is no longer exclusively mediated by the reinforcing properties of alcohol, such as the taste or enjoyable components of drinking, but rather by exposure to cues and social milieux previously associated with alcohol. If

---

[7]Ours is also similar to the approach used, among others, in Bleakley's (2007) work which evaluates the economic consequences of the eradication of hookworm disease from the American South at the beginning of the 20th century. The central variable in that study mapped into ours becomes $(\widetilde{a_{it}} \times w_{it})$, where $\widetilde{a_{it}}$ is the age difference between the young whose heaviest drinking experience of the week is in the weekend and the young who drink most in any of the weekdays, and w_{it} is the same weekend indicator variable as described above. In our case, therefore, our identification takes further into account of the differential drinking experience among the old between weekends and weekdays. For similar setups, see also Acemoglu and Johnson (2007) and Nunn and Qian (2014).

[8]There is mounting evidence that contextual conditioning may have a long-term impact on ethanol seeking behavior, even after ethanol withdrawal. See Zironi et al. (2006).

[9]Although the precise identification of the neurobiological substrates of bingeing and habitual ethanol seeking remains to be clarified, recent medical research emphasizes that specific neuroanatomical substrates within the prefrontal cortex and the striatum have been identified as playing a key role in the expression of alcohol habits (e.g., Barker et al. 2015). Age and context are thought to be key alcohol reinforcers, which may affect corticostriatal network activity and differentially engage the neurocircuits that induce a loss of flexible control over behavior, thus triggering binges and addiction (Zironi et al. 2006; Corbit, Nie, and Janak 2012).

the time approaching 5 p.m. on Fridays has reliably predicted the ingestion of beer in the past, then approaching that time will elevate the current marginal utility of beer and will increase one's desire for a pint. With repeated alcohol abuse, cues associated with past alcohol consumption are expected to cause the hedonic forecasting mechanism to anticipate (exaggerated) pleasure responses, inducing a powerful impulse to consume more. This leads to bingeing behavior.

To document the extent to which age and day of the week are related to bingeing in England, we use data from the HSE. Figure 1 shows the proportion of individuals drinking 10 or more units of alcohol in one episode — one of the many possible definitions of HED (see subsection III.D) — over the life cycle. The figure distinguishes between having the heaviest drinking day during the weekend (Fridays and Saturdays) from having the heaviest drinking day in the other days of the week. About 50% of the individuals aged 18–30 binge drink at the weekend, when fewer than 20% of those aged 50+ do so. In weekdays, both age groups drink substantially less, with 30% of the young and less than 10% of the old having a binge.[10] Similar clear patterns emerge in Figure 2, where we look at the bingeing distribution over the life cycle distinguishing consumption between weekends and weekdays.

Equation (3) captures the arguments put forward earlier and reflects the regularities shown in Figures 1 and 2, stressing out that our source of identification takes advantage of the *double difference* between age groups (young and old) and days of the week (weekends and weekdays). Moreover, both figures document that, although the differences between groups are substantial, the relationships between B and age (in Figure 1) and between B and day of the week (in Figure 2) are smooth around the cut-offs, i.e., ages 30 and 50, and each adjacent day of the week. This is why we include the polynomials $g(\cdot)$ and $h(\cdot)$ in (3), besides the fact that age and day of the week could be determinants of the outcomes under study.[11]

In the second stage we use a variety of data sets, each of which contains information on a given outcome Y, z, and $\mathbf{X}$, but not on binge drinking status. The data come from administrative records on road accidents, accident and emergency attendances, and arrests. They will be described in more detail in the next section, along with the HSE. Substituting (3) into (1) and assuming away the $g(\cdot)$ and $h(\cdot)$ functions as well as the covariate vectors $\mathbf{X}_{it}$ and $\mathbf{H}_{it}$ and the time fixed effects φ_t for simplicity, yields the following reduced form relationship between outcome Y and the instrument

$$Y_{it} = \pi_1(a_{it} \times w_{it}) + u_{it}, \tag{4}$$

where $u_i = \beta_1\nu_{it} + \epsilon_{it}$ and $\pi_1 = \phi_1\beta_1$. Expression (4) has a simple interpretation in which π_1 is the reduced form effect estimate of being a young adult in weekend nights on outcome Y_{it}, and the

---

[10]The same qualitative results are observed with different definitions of HED, e.g., 6+ through to 15+ units in one session per week. The figure also shows the life-cycle pattern when we consider the average number of alcoholic units drunk on the heaviest drinking day (whether this is a weekend day or weekday) rather than the proportion of individuals drinking a given number of units. The same features emerge. The (smaller) increases observed on Wednesdays and Thursdays seem to be driven mainly by students but, as the results in our sensitivity exercises show, this increase is not strong enough to lead to significant changes in the outcomes under analysis (see subsection IV.C).

[11]It is worth noting that age differentials and differences in the timing over the day/week have also been used to identify behavioral responses to risks other than drinking, such as watching violent movies (Dahl and DellaVigna 2009).

treatment effect estimate β_1 is the familiar ratio between reduced form and first stage parameters.[12]

Knowledge of π_1 is important in and of itself because it informs us about the differential propensity of young people to be admitted to A&E departments, or be involved in road accidents, or be arrested during weekend nights *relative to* older individuals. Testing that $\pi_1 = 0$ tests the hypothesis that $\beta_1 = 0$. This procedure is robust to weak instruments since no information about the correlation between control variables and $a \times w$ is required to test that there is no relationship between the outcome and the instrument.

Estimating (3) with the HSE data permits us to retrieve the first stage parameters, ϕ. With the outcome data sets, (4) can be estimated to identify the reduced form parameters, π. With $\widehat{\phi}$ and $\widehat{\pi}$ at hand and using the fact that $\pi_1 = \phi_1 \beta_1$, β_1 can then be identified using a two-sample two-stage least squares (TS2SLS) variant of the Angrist and Krueger's (1992) TSIV estimator.[13] Following Inoue and Solon (2010), standard errors of the TS2SLS estimates are obtained using a bootstrap method with 1,000 replications.

II.B. Identification Issues

Exclusion Restriction

Since $(B, a, w, \mathbf{H})$ are observed in one sample and $(Y, a, w, \mathbf{X})$ in others, and since we cannot take advantage of the introduction of dry/wet laws or changes in the minimum legal drinking age, our model can only be identified provided suitable exclusion restrictions are available (Arellano and Meghir 1992). In particular, we assume that differences in age among drinkers (young versus old) and in drinking times (weekends versus weekdays) do not enter (1) directly, but affect outcomes through their effect on bingeing. We also assume that the error terms, ϵ and ν, are uncorrelated. These restrictions are sufficient to identify β_1.

As explained above in this section, our instrument is the interaction between such two differences, $a_{it} \times w_{it}$, where a is equal to 1 if individual i is aged between 18 and 30 (and zero if the individual is aged 50+) at time t, and w_{it} is equal to 1 if i has the heaviest drinking experience of the week during the weekend (Friday and Saturday nights) at t and zero otherwise. We shall perform several sensitivity tests in which we change the definition of both a and w as well as their reference (control) categories. We will also consider the whole distribution of alcohol consumption and focus on different definitions of B.

Although leisure lifestyles are constantly changing, young people in many cultures around the world, and especially in advanced societies, devote a larger share of leisure time to social activities

---

[12]Notice also that the inclusion of the $g(\cdot)$ and $h(\cdot)$ polynomials in estimation brings our framework close to the model introduced by Carneiro, Løken, and Salvanes (2015).

[13]An alternative method to retrieve the β parameters in (1) consists of minimizing

$$G(\beta) = [\widehat{\pi} - \widehat{\phi}\beta]' \ \Omega \ [\widehat{\pi} - \widehat{\phi}\beta]$$

with respect to β, where Ω is a positive definite (optimal) weighting matrix. This two-sample minimum-distance estimation procedure leads to point estimates that are identical (and very similar standard errors) to those found with the TS2SLS method. They are therefore not reported.

such as partying and clubbing than their older counterparts. In countries like Britain, pubs and bars remain key social hubs in cities as well as in smaller communities (Furlong and Cartmel 2007). Going to the pub and partying are activities that typically involve social drinking, and weekend evenings and nights are a natural coordination device that intensify partying and social drinking among youths. These arguments, which echo the reasoning put forward by Lindo, Siminski, and Swensen (2018), combine to provide our instrument.

Earlier we also motivated our instrument choice using the notion that cue-triggered responses tend to raise the marginal utility of alcohol consumption for the young more than for the old. It should be emphasized that life cycle models of crime imply that crime and age are directly related: in particular, arrest rates decline with age, simply reflecting the increasing opportunity cost of crime due to the upward sloping age profile of earnings and to life cycle wealth accumulation (e.g., Fella and Gallipoli 2014). Therefore, exclusion restriction notwithstanding, each outcome (including A&E attendances and road accidents, and not just arrests) is allowed to depend on a smooth, quadratic function of age. Higher order polynomials do not have additional explanatory power.

Economic theory instead does not suggest a specific relationship between day of the week or time of the day and any of the outcomes under analysis.[14] There are however empirical studies that find an unambiguous association of weekends and holidays with arrests or police reports (Jacob and Lefgren 2003; Dahl and DellaVigna 2009; Card and Dahl 2011), and others find a strong relationship between time of the day and car crashes (Levitt and Porter 2001). There is also evidence that A&E attendances have a clear (and smooth) cycle over the week as well as within each day (Baker 2015). Like in the case of age, then, we allow each outcome to depend on a quadratic function of day of the week.[15]

The validity of our results depends on the exclusion restriction that $a_{it} \times w_{it}$ does not affect Y_{it} in (1). Besides the discussion below in this subsection, we formally account for the possibility that the exclusion restriction does not hold, employing the method described by Conley, Hansen, and Rossi (2012). In particular, we estimate β_1 under various priors for the influence that the $a \times w$ interaction effect might have on Y directly (see also Fletcher and Marksteiner [2017] and Satyanath, Voigtländer, and Voth [2017]). A natural prior is to assume that the age group–weekend night interaction effect is a fraction, δ, of the reduced form parameter π_1. The size of the violation of the exclusion restriction increases as δ increases, so that $\delta = 0$ corresponds to the baseline case in which the exclusion restriction holds, while, say, $\delta = 0.5$ implies that the violation is half the size of the reduced form estimate. Related to this same issue is the concern about omitted variable bias, which we assess using the methodology developed by Altonji, Elder, and Taber (2005) (see also Oster [forthcoming]). To compute the AET statistic, we use the TS2SLS estimate found from a restricted model without controls (β_1^R) and the corresponding β_1 estimate from the full model. The AET ratio $\frac{\beta_1}{\beta_1^R - \beta_1}$ tells us how much stronger the selection on unobservables must be with respect to the selection on observables to explain away the entire effect of HED on Y.

---

[14] Although it is clear that people's synchronization of their work activities is important (Weiss 1996), we know much less about group time coordination of leisure activities (Winston 1982; Hamermesh, Myers, and Pocock 2008).

[15] Since hours-of-day information is not available to the same level of detail in each of the outcome datasets and is unavailable in the HSE, hours-of-day polynomials are excluded from estimation.

Threats to Inference

We focus on three potential threats that stem from the two-sample estimation procedure. The first refers to the possibility that the effect we estimate is attributable not to HED but to other risky behaviors, such as illicit drug use and cannabis use in particular, or a combination of drug use and alcohol consumption.[16] The second explores the chance that young adults may be more likely than older people to engage in moderate drinking (and not just bingeing) which in turn could have spillover impacts on the outcomes of interest. The third threat refers to the possibility that young adults are involved in all sorts of evening and night activities (not just drinking) during the weekend more than older individuals, and this greater involvement may trigger more accidents, hospitalizations, and crimes.

1. Illegal drug use — To support our strategy we draw attention to three important regularities highlighted in previous research. First, there is clear evidence that marijuana and alcohol are substitutes among youths (DiNardo and Lemieux 2001; Crost and Guerrero 2012; Anderson, Hansen, and Rees 2013; Dragone et al., forthcoming). If we do find an effect among young adults, to a large extent then this should reflect *either* ethanol seeking behavior *or* cannabis use, but not both. Second, if, as we expect from the cue-theory of consumption, impacts on outcomes are observed during the weekend through social consumption, they are more likely to come from alcohol abuse rather than cannabis consumption. This is because, differently from alcohol, cannabis is known to have *no* systematic pattern of consumption during the week, that is, weekends are not characterized by increased use as opposed to weekdays (e.g., Thomas et al. 2012). Third, the medical literature provides strong evidence that brain abnormalities such as hippocampal volume loss and asymmetry — which may lead to memory impairment and reduced balance, spacial memory and navigation skills, and these in turn to injuries and accidents — are more likely to be observed among heavy episodic drinkers than cannabis users or joint users of marijuana and alcohol (e.g., Lisdhal Medina et al. 2007).[17]

Neither the HSE nor the outcome data sets have information on illicit drug consumption. To assess the direct extent of the role played by illicit drugs in youth's behaviors, we use data from the only source available for our purpose, i.e., the Arrestee Survey collected in three repeated cross-sections between 2003 and 2006. Although the data give us information on a selected subpopulation (e.g., arrestees might be inclined to take on more risks than the general population), they are likely to be useful. The Online Appendix describes the data and presents our findings in greater detail.

We underline three results that are relevant to the identification of β_1. First, excessive episodic drinking is considerably more prevalent than taking drugs among British arrestees. A staggering 36%

---

[16]After alcohol, marijuana is the second most commonly used intoxicant by youths in Britain, the United States, and other advanced economies (Smart and Ogborne 2000).

[17]Some studies find that marijuana's impairing effects on driving are moderate when taken alone, but can be severe when combined with alcohol abuse, indicating there might be a "potentiating effect" of multi-drug use (e.g., Robbe 1998). Young people also use other illicit drugs. Much medical research on the concurrent use of alcohol and cocaine documents an offsetting effect of cocaine on alcohol induced behavioral deficits. A common finding is that the effect of the alcohol/cocaine combination on violent behaviors and motor vehicle accidents is almost entirely driven by alcohol alone (Pennings, Leccese, and de Wolff 2002). Similar evidence emerges among concurrent users of alcohol and other types of drugs, such as heroin, hallucinogens, and non-medical painkillers (Midanik, Tam, and Weisner 2007).

of all individuals had a binge in the 24 hours preceding their arrest, while the rate of marijuana use is estimated to be less than 15%,[18] and that of all other illegal drugs together less than 9%. Second, about two out of 10 arrestees combine cannabis and bingeing. If, as suggested by the medical literature, most of the negative effects among co-users are driven by HED, then the estimates we present in Section IV are likely to be underestimates of the actual effect of alcohol abuse alone. Excessive alcohol consumption combined with other illicit drugs is much more uncommon, with prevalence rates below 3%. We therefore do not expect our results to be affected much by this latter group of co-users. Third, bingeing and drug use are generally substitutes or unrelated at the time of arrest. The only exception is cocaine, which however involves an extremely small proportion of the arrestee population and, quite possibly, an even smaller fraction of the general population.

The evidence that surfaces from these findings and the regularities documented by earlier research give us a strong indication that the effects we estimate (focusing on young people's behavior during weekend evenings and nights) are attributable to binge drinking rather than to the use of illegal substances.

2. Moderate drinking — If young adults disproportionately binge drink on weekends, they may also engage in higher levels of moderate drinking, which could equally have impacts on outcomes. If this were the case, then it would be hard to believe that our results are driven by bingeing, but rather by a combination of drinking and other behaviors. We address this concern in two ways. First, the analysis will be performed at all levels of alcohol consumption across the entire distribution. This permits us to assess whether moderate drinking has similar impacts to HED. Anticipating some of the results illustrated in the next two sections, we find that the first stage analysis is generally weak for 3 or fewer units of alcohol (which correspond to slightly more than one pint of beer or one glass of wine in a given episode). This indicates that the impacts on road traffic accidents, hospitalizations, and crime cannot be attributed to small or moderate quantities of alcohol consumption.

Second, and differently from what we will do later on, we focus our attention on measures of moderate drinking, that is 1 to 4 alcohol units, with 4 units corresponding to just more than one and a half pints of beer. The results are in the Online Appendix. About 50% of the 18–30 year old individuals in the Health Survey for England consume between 1 and 4 units of alcohol during weekdays and only 25% on Fridays and Saturdays. This consumption reversal goes in the opposite direction of what we described earlier for larger amounts of consumption. At the same time, about 50% of the old in the HSE consume 1–4 units during weekends and 70% of them do so during weekdays. Young adults, therefore, engage in higher levels of moderate drinking neither on weekends nor on weekdays. The corresponding first stage analysis (which defines moderate drinking so to include 5 and 6 units of alcohol) confirms that the young are 15 to 25 percentage points less likely to engage in moderate drinking than their older counterparts on weekends. Most of the outcome effects of moderate drinking should then be observed among the old in weekday evenings or nights. But this is not what the outcome data reveal (see the next section).

---

[18]This is likely to be an overestimate because the information on cannabis consumption in the survey refers to the month (instead of the 24 hours) before the arrest.

Taken together these two pieces of evidence suggest that moderate drinking is unlikely to explain the outcomes and the alcohol consumption patterns we observe in the data.

3. Selectivity — Are young individuals more likely to engage in evening/night activities than older individuals during the weekend? For instance, is there a disproportionate presence of young drivers in nighttime traffic accidents because youths have a greater likelihood to drive on weekend nights than their 50+ counterparts?

To address such questions, we analyze fine-grained, highly detailed time use data. The only nationally representative data currently available for the UK come from the Time Use Survey (UK-TUS), which was conducted in 2000/2001 on about 11,600 individuals aged 8 years or more.[19] We first focus on 10 specific activities. For each activity, we estimate a series of models as specified in (3) including covariates and time fixed effects, in which the dependent variable is the number of minutes spent in the activity at five different time blocks of the evening and night (covering the time from 6 p.m. to 3 a.m.). The key predictor is $a \times w$ and its parameter estimates are displayed in Figure 3. Further details of this analysis are discussed in the supplementary material in the Online Appendix.[20]

In seven of the 10 activities, we cannot detect statistically different patterns in the time spent by individuals aged 18–30 over weekend evenings and nights relative to individuals aged 50 or more. These activities include driving a car as well as being a car passenger (panels A and B). Extending the analysis to other vehicle types (such as bicycles, bikes, and lorries) does not change this result.[21] There is no evidence therefore that our outcome effect estimates would reflect a greater propensity of young people to be on the road than their older counterparts. The other activities for which we detect no statistical difference by age over the weekend are: going to the cinema, cooking, watching TV (although in some time windows the young spend less time than the old on this activity), visiting (or being visited by) friends in private dwellings (and the young engage more in this at later time blocks), and spending time at home with other household members (panels D–H, respectively).

There are two activities in which younger and older individuals spend statistically and substantially different amounts of time during Friday/Saturday evenings and nights. The young sleep less (10–45 fewer minutes) than the old (panel I), and appear to substitute their shorter sleeping time with a greater amount of time devoted to 'social activities' not included in the previous panels (between 15 and 20 more minutes; see panel J). As shown in panel J.1, about 75–90% of this additional time devoted to social activities is accounted for by the time spent in pubs, bars, restaurants, and night clubs.[22]

---

[19]Respondents fill up two 24-hour diaries, one to be completed on a weekday and one on a weekend day, giving us a total of about 18,000 diaries. Each day is broken down into ten minute sections with both activity and location recorded. For more detail on the UK-TUS data, see the Online Appendix.

[20]As a robustness check, we performed the analysis using the indicator variable that takes value one if an individual spends a positive amount of time in any of the 10 activities and zero otherwise as a dependent variable. The estimates obtained from this exercise are consistent with those discussed here and are available in the supplementary material.

[21]Individuals aged 18–30 do spend 1–2 more minutes using public transports (panel C). This is arguably a safer mode of transportation as opposed to cars or bikes, which might reduce the likelihood of accidents among the young. This is also consistent with the evidence found on commuting to work and mentioned below.

[22]The young spend also 5–10 minutes more than the old on the residual category, 'Social Activities not in Pubs' (panel J.2). The data however do not allow us to categorically exclude all of the time spent in clubs and pubs from these other activities. This might explain the results in panel J.2, particulary the estimates found at the later time blocks.

We also analyze double differences (among young and old and on weekends versus weekdays) in other 12 (risky) activities, which young adults may disproportionately engage in and for reasons that could be completely unrelated to alcohol abuse. We look at the same five time windows between 6 p.m. and 3 a.m. as before. These activities include jogging, biking, playing ball games, doing gym and fitness activities, walking, swimming, engaging in all sorts of sport related activities, gardening, dog walking, and commuting to work using any means of transportation (public or private). The estimates, reported in the Online Appendix, detect no differential between young and old on weekend evenings and nights.

Such results therefore show quite unequivocally that, when compared to their 50+ counterparts, young adults are unlikely to be more engaged in behaviors that may lead to hospitalizations or accidents for reasons that are unrelated to HED. They are also unlikely to be over-represented in weekend evening/night activities, except for the time they spend in exactly those locations, such as pubs and bars, where people meet to drink. This should lessen concerns related to selectivity and lend credibility to our empirical strategy. The discussion of other related exercises, and in particular the issue of spillover effects involving either drinking companions (but non-bingers) or other (non-drinking) individuals, is deferred to Section IV.

III. Data

We examine three outcome domains using three different data sources. First, we present national administrative data collected by the Department of Transport on road accidents as well as records on breath test screening after collisions. Second, we describe Accident and Emergency attendances with unique hospital data obtained from one care trust in England. Third, the crime data come from the two largest police forces, the West Midlands Police and the Metropolitan Police Service. Finally, we describe the Health Survey of England, which contains data on drinking status and gives us the information needed to estimate the first stage.

III.A. Road Accidents

The Road Accidents Data (RAD) are collected by the police for the Department of Transport whenever an accident involves at least one personal injury. We have all the RAD administrative records from 2008 to 2010 on over 1.2 million vehicles in England. Each record contains details about the accident and the individuals involved, including their age and sex, the exact time and location of the accident, and its severity, which in turn is distinguished into fatal, serious, and slight.

A graphical cut of the data is given by the tempograms in Figure 4.A (breakdowns by accident type are reported in the Online Appendix). This shows the number of road accidents by hour of the week for individuals aged 18–30 and for individuals aged 50 or more averaged over the whole sample period. On the horizontal axis, 0 corresponds to the first hour of Monday (00:00 to 00:59) and 168 refers to the last hour of Sunday (23:00 to 23:59). The darker portion of the line represents attendances that occur during the night (from 00:00 to 06:59). The vertical lines indicate the baseline

definition of weekend (Friday and Saturday nights), when the gap between the accidents faced by the young and the accidents experienced by the old is largest. The difference is even clearer in the case of fatal injuries. For the most common type of road accidents, those slight in nature, we observe similar week-night patterns with the same divergence at weekends.

The RAD records do not collect data on alcohol consumption. The only source for England with information on alcohol involvement is the Breath Test Screening (BTS) data released by the Department of Transport on all breath tests performed by police authorities from 2009. We use the 2009–2011 BTS records on over 600,000 road traffic collisions, excluding the records collected in relation to moving traffic violations and other unspecified reasons that do not result in an accident.[23] For each record, the BTS data report the Breath Alcohol Level (BAL; measured in mcg/100ml) of all drivers involved, year and month of the collision, while the information on the day and time of the accident is available in six four-hour blocks from 00:00–03:59 to 20:00–23:59 and permits us only to separate weekdays from weekends. Drivers' age is reported in eight categories, allowing us to distinguish individuals aged 16–29 from those aged 50+. Also, there is no information on the severity of the accident, hence we cannot determine whether a collision is fatal, serious, or slight.

We use the BTS data for robustness, as they enable us to assess whether the estimates obtained with the RAD records are corroborated or not. Furthermore, since they contain a measure of alcohol involvement, the BTS data allow us also to estimate two-stage least squares (2SLS) models. We thus can check how the TS2SLS estimates (obtained from either RAD or BTS records) compare with the corresponding 2SLS results (obtained from the BTS data).[24]

III.B. Accident and Emergency Attendances

Data on A&E records come from the Solihull Care Trust (SCT). SCT was one of 152 primary care trusts in England, which were abolished in March 2013 as part of the UK Health and Social Care Act 2012. Solihull is a town in the West Midlands of England approximately 10 miles away from the city of Birmingham with a population of about 210,000 in 2010. With a median population size across primary care trusts of around 285,000 nationally, SCT is smaller. But if Greater London and Birmingham are excluded, SCT is in line with the national average. When comparing SCT to the rest of the country, we do not find any statistically significant difference in education and health measures (including bingeing, smoking, BMI, and healthy eating habits), although Solihull has a slightly greater fraction of 50+ residents (37% versus 34% nationally) and a slightly smaller fraction of 18–30 year olds (15% versus 18%).[25]

We have over 150,000 attendance records from midnight on the 1st of April 2008 to midnight on the 21st of January 2011. Attendances are recorded using the Tenth Revision of the International

---

[23]Such records are excluded because the BTS results are compared to those found with the RAD records, which pertain only to road accidents.

[24]To perform this comparison we need to convert blood alcohol levels into alcoholic units. A brief description of the conversion is provided in subsection IV.B. More details are in the Online Appendix, where we also show the distributions of units obtained from the two data sources.

[25]The supplementary material in the Online Appendix presents the comparison between SCT and the rest of the country's other care trusts. Balance exists essentially along all variables.

Classification of Diseases, ICD-10 (World Health Organization 2007), which specifies the exact cause of attendance. Among all records, we select those that medics report as being related to episodic alcohol abuse (Brismar and Bergman 1998).[26]

Figure 4.B shows the number of injury related A&E attendances by hour of the week for individuals averaged over the sample period distinguishing individuals aged 18–30 from individuals aged 50 or more. The pattern of attendances across the two age groups is almost identical during weekday nights, although during the day it is the 50+ who experience a greater number of attendances on average. As in the case of road accidents, however, a night gap emerges as the weekend approaches, with the young experiencing more attendances.

We can disaggregate the SCT data by the nature of injury and body region injured. This is important because falls are known to be a common consequence of excessive alcohol consumption, and head, hands, and elbows are typically the most affected body parts (e.g., Savola, Niemelä, and Hillbom 2005). With the nature of injury we can distinguish open wounds from superficial injuries (see the supplementary material in the Online Appendix). In the case of attendances due to head injuries, there are more spikes in comparison to Figure 4.B, and these appear most prominently for the young during the early hours of Sunday morning.[27] Off-weekend days instead have a very similar trend for treatment and control groups. Similar patterns are found for hand and elbow injuries, open wounds, and superficial injury attendances.

III.C. Arrests

Information on arrests comes from the West Midlands Police (WMP) and the London Metropolitan Police Service (MPS), obtained through Freedom of Information requests. These two forces operate in the two most populated areas in the UK, with a population of 2.6 and 7.8 million covered by the WMP and the MPS, respectively.[28]

We have counts of offences for each day of one week in February, May, August, and November for three consecutive years from 2009 to 2011. None of the twelve weeks includes a public holiday. Since information on the exact time of arrest is not available, each day is split into four 6–hour blocks. As our focus is on the effect of binge drinking on arrests we concentrate on the 00:00–05:59 block, allowing for some delay between drinking and the time of arrest. The data do not have exact information on the arrestees' age but are stratified into three age groups, i.e., below age 30, between 30 and 50, and over 50.

We identify two broad categories of arrests. The first category comprises arrests that, according to the police records, are directly related to alcohol. These in turn distinguish 'drunk' (which is a

---

[26]Despite the link to alcohol intoxication, the SCT data do not include a direct measure of alcohol involvement. Thus, we cannot differentiate between varying levels of alcohol consumption.

[27]As in the case of injury related attendances, the gap between the two age groups begins to open up before Friday and Saturday, with a gap already emerging on Thursday. We shall perform robustness checks of the baseline estimates by including Thursday (or Monday) as part of our treatment period.

[28]The WMP area (which includes Birmingham, the second largest city in England) is comparable to the national average on health behaviors, education, and age profiles. The population in London instead is significantly younger and has a smaller proportion of binge drinkers and smokers and a greater proportion of healthy eaters.

combination of drunk and incapable, drunk and disorderly, and drunk in a public place) from 'drunk driving' (a combination of drunk in charge of a motor vehicle, accidents with a positive breath test, and accidents with a refusal on breath test). The second category refers to arrests that are classified as indirectly related to alcohol, that is, crimes where police officers report a presumption that alcohol consumption might have played a role in the offence. These arrests comprise violent crimes (which include actual bodily harm, grievous bodily harm, violent disorder, and affray), common assault, sexual assault, criminal damage, robbery, theft, and burglary. As in Levitt and Porter (2001), however, the data do not contain quantitative information on alcohol involvement.

Figure 4.C shows the number of arrests for all alcohol related incidences (both directly and indirectly related to alcohol) for each 6–hour block for individuals aged under 30 and individuals aged more than 50 averaged over the twelve weeks under analysis. The top two lines correspond to the 18–30 age group. The solid darker line indicates arrests occurred in the night (00:00–05:59), while the dashed lighter line represents arrests recorded in the other three time blocks (06:00–11:59, 12:00–17:59, and 18:00–23:59). In weekdays, more arrests are made during the day, but the opposite occurs over the weekend when more arrests are made during the night. The bottom two lines show the average number of arrests for those aged over 50. During the working week the patterns (albeit not the levels) are similar across the two age groups. But the night pattern of arrests for older individuals remain low and flat during the weekend, while there is a large increase in arrests among young adults, almost trebling the mean for workday nights. An even sharper age/time difference emerges in the case of arrests that are directly related to alcohol abuse.

III.D. First Sample Data and HED Measurement

To estimate the first stage equation (3), we use nationally representative data from the Health Survey for England (HSE). This is a cross-sectional survey collected annually since 1991, with around 12,000–20,000 respondents each year. We use the HSE data and match both geographic areas and years of analysis on the second-stage outcome measures. Besides information on a wide range of demographic variables, the HSE has detailed self-reported questions on alcohol consumption. The survey asks which day the respondent drank most in the past seven days, and on that occasion the type and number of drinks. Following precise medical guidelines, the survey converts drinks into alcoholic units, so we know the number of units drunk on the heaviest day over the previous week.

In 2011 the HSE gathered additional information on drinking from a seven-day drinking diary.[29] In the Online Appendix we compare the alcohol consumption information from the standard HSE collection to that obtained from the 2011 diary and find patterns that are strongly consistent. This lends support to the standard HSE alcohol data we use in most of the analysis even if it does not have the distribution of alcohol consumption over the entire week. For the TS2SLS effects on outcomes we shall also perform a battery of sensitivity checks using this diary information rather than the standard

---

[29]For each of the seven days, respondents recorded whether they had drunk alcohol or not. If they did, they had to report the amount drank, the times of day when they drank and the location. They did not report how much they drank in each 10-minute period however, nor how much they drank in each location. Individuals were asked to complete the diary even if they drank no alcohol in the week.

HSE consumption data in the first stage. The discussion of such results is deferred to subsection IV.D.

To have a full picture of alcohol consumption, including moderate drinking, we consider the entire distribution of alcoholic units. In addition, we focus on three possible definitions of heavy episodic drinking, i.e., 8 or more, 10 or more, and 12 or more alcoholic units. To put these quantities into context, three pints of a 4% alcohol by volume (ABV) beer (each pint being 568 milliliters) are equivalent to three glasses of a 13% ABV wine (each glass being 175 milliliters). Both correspond to 6.9 units of alcohol. Six pints of beer, or six glasses of wine, correspond to 13.8 units. By looking at these specific definitions we can separate out HED (and its effect on outcomes) from "just" drinking smaller amounts of alcohol.

We do not commit to one specific definition for a number of reasons. Firstly, by looking at the whole distribution we can capture possible effect nonlinearities, which have been shown to be important in the context of ethanol seeking behavior (e.g., Cook and Moore 2000; Taylor et al. 2010).

Secondly, there is no officially accepted definition of binge drinking, for this varies across studies and regulatory agencies. Much of the US literature defines binge drinking as 5+ and 4+ *drinks* on a single occasion for men and women respectively (e.g., Naimi et al. 2003; Cawley and Ruhm 2012). In Britain, instead, the notion of alcoholic *units* is more frequently used. For instance, the British professional body of doctors defines heavy episodic drinking as 10+ units in a single session (Royal College of Physicians 2001); others define it as more than 8 units in one day (Wright and Cameron 1997), or at least half the weekly recommended units on a single occasion, i.e., 14+ units for men and 12+ units for women (Norman et al. 1998), or 12+ units in a row (Measham 1996). Since the HSE converts the information on drinking into units, our analysis is based on alcoholic units. Nevertheless, using the parameters published in the NIH Clinician's Guide (National Institute on Alcohol Abuse and Alcoholism 2005), according to which one drink is equivalent to two alcoholic units, we can convert one measurement into the other. So the 5+ drink definition used in most US studies corresponds to our 10+ unit definition (i.e., just over four glasses of wine or four pints of beer).

Thirdly, the HSE contains only self-reported measures of alcohol consumption and does not collect objective biological markers, such as blood alcohol concentration. According to some studies, self-reported measures of drink participation and intensity are likely to be subject to underreporting (Midanik 1988). Others, instead, argue that alcohol involvement is overstated in surveys (e.g., Ekholm et al. 2008). Using the entire distribution permits us to limit the influence of misreporting issues linked to recall bias or strategic reporting behavior and provides us with a more accurate picture of the range of possible effects.

Finally, we explore the data in a number of different ways that help us strengthen our identification strategy. We briefly report on three exercises.[30] First, we estimate the probability that an individual's heaviest drinking day is a weekend day, using both the HSE standard alcohol information and the 2011 drinking diary data. In both cases, we find that individuals aged 18–30 are at least 25% more likely to drink heavily in weekends than their 50+ year-old counterparts, even after controlling for a

---

[30]For the sake of space limitation, the results mentioned here are not reported in the text but are available in the supplementary material provided in the Online Appendix.

large set of observables. This reiterates the relevance of our instrument.

Second, using the 2011 diary information, we check whether individuals binged in weekends rather than engaging in HED frequently over the week and not just during the weekend (and thus had potential alcohol dependence problems). The data confirm that weekend HED is far more prevalent than alcohol dependence among the whole population (40.7% versus 3.3% among young and old when we look at 10+ units) and in particular among the young (53.3% versus 2.7%). Dependence is thus more likely to be observed among the 50+, even though older adults drink generally much less. Our implicit identifying assumption of relative alcohol abstinence on weekdays and heavy episodic drinking concentrated in the weekend among the young seems therefore to be borne out by the data quite strongly.

Third, we go back to the discussion in subsection II.B on the selection issues that may arise while using $a \times w$ as instrument. A related aspect is that the young and the old who go out and drink in the weekend may have different characteristics relative to their respective counterparts who go out and drink in weekdays. If this is the case, we might overstate the effects of HED on outcomes if, for instance, "reckless" young and "ordinary" 50+ year olds self-select to go out in the weekend. We cannot isolate (unobserved) taste-based characteristics along which young and old individuals might differ in their ethanol seeking behavior over the week. But we can check if weekend heavy drinkers are observationally different from weekday drinkers within each age category (i.e., within the young and within the old). We find that alcohol consumers as well as binge drinkers in weekends are essentially identical to weekday consumers/bingers along a wide range of observables (including sex, education, health measures, BMI, smoking habits, and geographic location). We detect differences only in terms of age: among those aged 18–30, weekend bingers are slightly older (24.6 versus 22.6 years), while among those aged 50+, weekend bingers are slightly younger (58.1 versus 60.2 years). By and large, therefore, this evidence suggests that selection on observables into weekend HED *within* the young and the old is unlikely to play a role.

III.E. First Stage Estimates

We conclude this section with the estimation of the first stage on the HSE. The estimation is performed on grouped level data for a total of 168 cells defined on year (3 years), quarter of the year (4 quarters), day of the week (7 days), and age group (2 groups). Since the three outcome data sets refer to different geographic units and have a slightly different temporal coverage, we perform three different first stage estimations, each of which is relevant to a different outcome using the following variant of (3):

$$B_{it} = \phi_0 + \phi_1(a_{it} \times w_{it}) + \phi_{21}A_{it} + \phi_{22}A_{it}^2 + \phi_{31}W_{it} + \phi_{32}W_{it}^2 + \mathbf{H}_{it}'\gamma + \varphi_t + \nu_{it}, \tag{5}$$

where the dependent variable is the cell-specific mean of the dummy variable equal to 1 to indicate consumption of a given number of alcohol units, and 0 otherwise. As mentioned in Section II, because the data show a smooth relationship between binge drinking status and both age and day of the week, we include quadratic polynomials in both A and W. The vector $\mathbf{H}$ includes cell-average controls for

years of education, and indicator variables for sex, ethnicity, highest attained qualification, region of residence, whether individuals ever smoked or not, and whether they have long-standing illnesses or not. To account for variation in alcohol availability, we also control for the cell-average number of licenses per 1,000 individuals of the population. The term φ_t denotes year and quarter of the year fixed effects. Each first stage regression is weighted by cell size.

Figure 5 presents the results. Each panel (one for each of the outcome domains) reports the point estimate and 95% confidence interval for ϕ_1 found at each separate level of alcohol consumption, from 1+ units (i.e., less than half a pint of beer or more) to 25+ units (almost 11 pints or more). In every panel we also report the value of the F test on the excluded instrument.

Panel A of the figure shows the first stage results which will be used to assess the effect on road accidents. The instrument is generally highly relevant. Between 2+ and 20+ units, the F-test values are well in excess of the standard critical threshold of 10, and the estimates suggests that, compared to the old, the young have a 10–25 percentage point greater likelihood of drinking any of the given amounts of alcohol on weekends as opposed to weekdays. The first stage results relevant for the A&E outcomes are presented in panel B. Again, $a \times w$ is a strong determinant of drinking status over a large part of the distribution of alcohol consumption (typically between 15 and 30 percentage point increases). The F-test statistics reveal strong first stage results from 3+ to 18+ units. Finally, panel C reports the results that will be used for the estimation of the arrest outcomes. As before, $a \times w$ plays a key role (reflecting a greater risk for the young on weekends of about 10–20 percentage points), with the F-test statistics indicating powerful first stage results between 3+ and 19+ units. The same pattern of results emerges if we use the 2011 drinking diary information.

Another piece of evidence from each of the three panels of Figure 5 is that $a \times w$ is a good predictor of neither low amounts nor very large quantities of alcohol consumption. As documented in Section II, the former might reflect the fact that the old are more likely to engage in moderate drinking than young adults on weekends (as well as during weekdays). The latter result instead can indicate that extremely high levels of alcohol consumption cannot be explained only by differences in age and day of the week, but may be linked to factors that goes beyond the scope of this paper, such as addiction and mental disorders.

IV. Results

IV.A. Reduced Form Estimates

As in the case of the first stage, each of the outcome data sets is aggregated into cell means stratified by year, quarter of the year, day of the week, and age group, for a total of 168 cells. For each outcome variable, we estimate the following augmented version of equation (4):

$$Y_{it} = \pi_0 + \pi_1(a_{it} \times w_{it}) + \pi_{21}A_{it} + \pi_{22}A_{it}^2 + \pi_{31}W_{it} + \pi_{32}W_{it}^2 + \mathbf{X}_{it}'\pi_4 + \varphi_t + u_{it}, \tag{6}$$

where the vector $\mathbf{X}$ includes the same controls for alcohol demand shifters and alcohol supply used in the first stage (see subsection III.E). Notice that our first stage sample and each of the reduced

form samples match in terms of the area under analysis. Table 1 reports the π_1 estimates by outcome domain.

Road Accidents — The results from the RAD records are in the top part of panel A. The fit of the data is remarkably good, with R^2 between 0.82 and 0.98. We find an increase of 16.4 accidents per weekend for the treated group at the treated time (first row of the panel). When the analysis is broken down by severity (next three rows), the ϕ_1 estimates are always positive and statistically significant. On average, there are 0.8 additional fatal accidents and 3.2 additional accidents with serious injuries every weekend. The largest reduced form effect of bingeing is on accidents with slight injuries, for which we find an increase of 12.4 additional accidents. The lower part of the panel reports the results obtained with the BTS data. We find a significant increase of almost 23 collisions for the young on weekend nights.

A&E Attendances — Panel B of the table shows there is a positive and highly significant estimate of 1.28 additional injury-related attendances each weekend for the 18–30 age group relative to the 50+ age group. The next two rows report the estimates on attendances for which the primary diagnosis is an injury to the head (second row) or to hands and elbows (third row), which A&E doctors expect to observe primarily as a result of episodic alcohol abuse. In both cases we find positive and highly significant increases in attendances. So we do in the last two rows of the panel, which refer to the estimates regarding the nature of the injury. The largest increase in attendances is found in the case of open wounds (with 0.61 additional attendances), but the effect is statistically significant also for superficial wounds. The fit of the data is good across all sorts of attendances, with R^2 ranging from 0.55 to 0.83.

Arrests — The results on arrests are reported in panel C. The fit of the data is generally strong, with R^2 between 0.75 and 0.96, except for sexual assaults and robberies where the R^2 statistic is 0.41 and 0.69 respectively. The reduced form estimates imply an additional 80 arrests every weekend for all alcohol-related crimes. About 28.8 additional arrests are directly related to alcohol (second row), with 13.3 arrests involving individuals who are drunk and disorderly and 15.5 involving people who are drunk drivers. Nearly 51 additional arrests are indirectly related to alcohol, with 24.4 extra arrests for violent crimes, 4.7 for common assault, 6.1 for criminal damage, and 3.8 for robbery. The ϕ_1 estimate is not significant for the remaining three types of crimes (sexual assault, theft, and burglary).

IV.B. Benchmark Estimates

For each outcome domain, the TS2SLS estimates of β_1 and 95% confidence bands are presented in Figures 6.A–6.C, which report the effects observed between 3+ and 18+ or 19+ units. The estimates for the bottom and top of the consumption distribution are not reported because their corresponding first-stage F-test values suggest that the instrument is not likely to be relevant. Tables 2–4 report the results using three different definitions of HED (i.e., 8+, 10+, and 12+ units), together with the daily mean and the night mean of the dependent variable. To ease interpretation, each table also reports

an average effect for each outcome at 10+ units, which is given by the value of the estimated β_1 for 10+ units expressed as a fraction of the daily mean of the dependent variable, and a marginal effect, which is computed as the change in β_1 between 8+ and 10+ units expressed as a fraction of the daily mean of the dependent variable. Finally, the last column of each table shows the AET statistic and in square brackets the lower and upper bounds of the union of the 95% confidence intervals around the TS2SLS estimate at 10+ units assuming treatment effects bounded above by zero and below by δ. As in Fletcher and Marksteiner (2017) among others, we contextualize the restrictions on δ by their percentage on the reduced form parameter π_1 and, in the table, present the results for the case in which $\delta = 0.50$. In the Online Appendix we show the results found for other unions of CI with δ set at 0.20, 0.40, 0.60 and 0.80. These alternative unions are strongly supportive of the estimates we discuss below. The specification of the $g(\cdot)$ and $h(\cdot)$ polynomials, the control variables in the $\mathbf{X}$ vector, and the fixed effects terms in φ_t are the same as in (6).

Road Accidents — The TS2SLS effect of drinking on all road traffic accidents over the alcohol consumption distribution is displayed in Figure 6.A for both the RAD records (solid line) and BTS data (dotted line). All estimates from both data sources are positive and statistically significant, and do not statistically differ from each other. The relationship between alcohol consumption and accidents is monotonically increasing from 5+ units up, with higher units consumed leading to an increase in recorded accidents. The same features emerge when we distinguish accidents by severity type (see the Online Appendix).

Figure 6.A also shows the 2SLS estimates obtained from the BTS data (solid gray line).[31] The figure displays the estimates only if the corresponding first-stage F-test values are above 10. The number of accidents increases as the number of units goes up from 11+ units, while at lower levels of consumption we find a negative relationship between units and collisions.[32] The key finding is that, across the whole consumption distribution, the 2SLS estimates are significantly greater than the corresponding TS2SLS estimates. This is not driven by the fact that the BTS data allow us to include fewer control variables than the RAD records do. Restricting the set of controls in the RAD to those available in the BTS data in fact leads to the same TS2SLS results found in the benchmark analysis (see the dashed line with squares in Figure 6.A). Thus, the TS2SLS model yields estimates of the effect of bingeing on accidents that are quite conservative as compared to those found with a standard

---

[31]To convert breath alcohol levels into alcoholic units we use a commonly used procedure that depends on individual body weight and the time elapsed since the last drink. Since the BTS does not report information on either of such measures, we impute weight with the HSE gender-specific means (84.3 kilograms for men and 71.2 kilograms for women) and we use different values of elapsed time (ranging from 2 to 6 hours). In the figure we report the estimates obtained using 4 hours of elapsed time. These estimates are very similar to those found with the other measures of time duration. They are also robust to alternative measures of individuals weight (e.g., when the gender-specific means are replaced with gender-specific medians). See the Online Appendix for more details.

[32]This nonmonotonic relationship can have a number of explanations. For instance, the mapping between blood alcohol levels and alcoholic units may be sensitive to body weight and time elapsed since the last drink, which are likely to be different for different individuals (e.g., alcohol absorption times may be different for individuals with equal body weight). Moreover, by using information on individuals involved in collisions, the BTS results refer only to a selected subpopulation, whereas the alcohol involvement measure used for the RAD estimates comes from a nationally representative sample.

2SLS method.[33] Of course, we cannot verify whether this is generalizable to the other outcomes.

Table 2 shows that defining binge drinking as 8 or more units of alcohol leads to an additional 66.1 road accidents per weekend day among individuals aged 18–30 as opposed to drinking fewer than 8 units. The effect increases to 97.3 accidents for 12+ units. There were an average of almost 163,000 road accidents per year over the relevant sample period, approximately 446 per day, of which 33 occurred each night. Therefore the figures for all accidents reported in the table imply a 15% increase in the daily average number of accidents if we use the 8+ unit definition and a 22% increase in the case of 12+ units. The estimates from the BTS data (reported in panel B) confirm the same patterns, although they are always slightly greater than the effect estimates found with the RAD records. Likewise, the BTS-based 2SLS estimates in panel C are at least twice as large as the TS2SLS estimates obtained with the RAD, regardless of whether we compare them to the benchmark results in panel A or the other TS2SLS estimates in panel C that include the same set of controls as in the BTS data. From now on, therefore, we focus exclusively on the more conservative RAD-based TS2SLS effects in panel A.

The average impact obtained in the case of 10+ units is 18.6%. This accounts for 4.1 additional fatal accidents (72% of the fatal road accident mean per day and 405% of the night mean), 16.3 additional accidents with serious injuries, and 62.8 additional accidents with slight injuries, with all the estimates being highly statistically significant. The estimated increase of 3.8% associated with the marginal change from 8+ to 10+ units corresponds to 6,100 additional road traffic accidents (including 300 extra fatalities) every year. These effects are substantial.

It is worth stressing that Levitt and Porter (2001) find that on average 18.5% of all car crashes in the US between 1983 and 1993 were attributable to drinking drivers. Our estimate for 10+ units at 18.6% is remarkably close, despite the fact it refers to different country and different time periods and even though not all road accidents involve cars and not all drinking drivers are binge drinkers.

The last column of Table 2 presents the AET statistics and in square brackets the unions of the 95% CI as proposed by Conley, Hansen, and Rossi (2012). For all accident types, datasets and estimation methods, the AET ratio is always negative. This means that our estimated treatment effect is likely to be a downward biased estimate of the true HED impact, as long as the selections on observables and unobservables are positively correlated. In addition, for $a \times w$ direct effects in (1) up to 50% of the size of the reduced form estimate, the union of confidence intervals never includes zero and delivers estimates that always contain the corresponding TS2SLS estimates of β_1. This indicates that our baseline results are robust to fairly sizeable violations of the exclusion restriction. The figures in the Online Appendix confirm that this is the case even when the violation gets greater and up to 80% the size of π_1. Therefore, our TS2SLS results do not get attenuated even if the direct impact of the instrument on all types of road accidents were allowed to be implausibly large.

We conclude this analysis by looking at possible effect heterogeneity by road type. This is meant

---

[33]The RAD-based TS2SLS estimates are also smaller than the corresponding least squares estimates obtained on the BTS data, which in turn are below their 2SLS counterparts. All such results are reported in the supplementary material of the Online Appendix.

to provide further support to our identification strategy. In subsection II.B, in fact, we showed that young and old individuals spend the same amount of time in a wide range of activities, including driving in weekend evenings and nights. They could however be involved in different types of driving experience, and it is this difference that brings about our results. For example, young adults might be driving in more densely populated urban centers to move around the places where they socialize, whereas older adults might be driving on less congested roads.[34] To check whether this is the case, we use data from the RAD, which distinguish different road types, e.g., classified roads (including motorways/freeways and major roads) and unclassified roads that are intended for local traffic, and repeated our benchmark analysis.[35] The TS2SLS estimates, reported in the Online Appendix, strongly support the results of Table 2. Across all types of roads and all accidents, the average effects obtained in the case of 10+ units is consistently around 18–19%. At each of the other definitions of HED and for each accident type, we also find estimates that are closely comparable to those reported in Table 2. Our results therefore cannot be attributed to different combinations of young and old drivers in different types of roads.

A&E Attendances — Figure 6.B shows the effect of drinking on all injury related attendances across the whole distribution of alcohol consumption. The effect is statistically significant over the entire distribution. From 7+ units, we find that, as alcohol consumption goes up, so does the number of attendances. But at lower drinking levels the relationship between consumption and attendances is mildly negative. These features hold up across all different types of attendances, distinguished either by the nature of injury or by the body part injured (see the Online Appendix).

Table 3 indicates that the TS2SLS effect of drinking 8 or more units on all injury related attendances is 3.8 per weekend day. The effect rises to 4.7 and 6.5 additional attendances in the case of 10+ and 12+ units respectively. The estimate found for 10+ units corresponds to 6.6% of the mean number of all injury related visits over the day, and to 123% of the mean number of injury related visits per night.

The impact on head injuries is even more substantial. The estimate of 2.4 additional attendances in the case of 10+ units implies a 22% increase in the average number of head injury visits during the entire day and a 244% increase over the night. Quite large is also the effect estimated in the case of A&E attendances due to open wounds, with the 2.2 estimate corresponding to an another increase of 22% in the mean number of attendances per day.

To put these estimates into perspective, consider the marginal increase of 0.87 attendances per day in all injury related attendances when moving from 8+ to 10+ units. This corresponds to a 1.2% rise in attendances. If applied to the whole country, such a figure would account for one-third of the annual growth in A&E attendances between 2003/04 and 2013/14, or about 63,000 extra attendances

---

[34]Public transports could offer an alternative to older people, especially in urban areas. The evidence in subsection II.B however shows that this is not the case.

[35]Major roads are known as A roads. Besides A roads and motorways, classified roads include also other smaller roads, known as B roads (intended to connect different areas and feed traffic between A roads and smaller roads on the network) and C roads (intended to connect unclassified roads with A and B roads, often linking housing estates or villages to the rest of the network).

every year (NHS England 2015). These are large effects.

All AET ratios are positive (except that for superficial wounds, which is negative) and extremely large. This suggests that, to attribute the entire effect of HED on hospital attendances to selection, the selection on unobservables would have to be between 18 and 72 times stronger than the selection on observables, which seems unlikely. Moreover, the figures from the union of CI (including those in the Online Appendix) comprise zero only in the case of superficial injury attendances and indicate that the $a \times w$ instrument on A&E attendances would have to be more than three-fifths of the overall effect to render our TS2SLS results insignificant, a magnitude that seems implausible, given that we control for $g(A)$ and $h(W)$ in our structural equation. We take all these checks as providing strong evidence in support of our identification strategy.

Arrests — Figure 6.C displays the results for all arrests that the police attribute to alcohol abuse. The effect is always positive and statistically significant, except at 15+ and 18+ units, and it is increasing from about 5+ units up. As shown in the supplementary material, virtually the same features are found for the two broad categories of arrests we have in the data, both those that are directly related to and those that are indirectly related to alcohol abuse.

Table 4 shows that a binge of 8+ units leads to 415 additional arrests per weekend day, which implies a 63% increase in the daily average number of arrests associated with all alcohol related incidences. The impact goes up to 471 and 577 extra arrests for 10+ and 12+ units respectively, corresponding to 71% and 87% increases in the daily mean.

Almost two-fifths of the effect on all arrests is due to arrests that are directly related to alcohol abuse. The largest fraction of this effect is attributable to drunk driving. According to the results in Table 4, this amounts to 54% (=80.6/149.4, where 149.4 is the estimate on directly alcohol-related arrests, and 80.6 is the coefficient on drunk driving using the 8+ unit definition). This proportion remains stable even when we use the other definitions of HED. Bingeing 10+ units in one session more than doubles the average effect on arrests due to drunk driving and more than triples the average effect on arrests due to being drunk.

Arrests that are indirectly related to alcohol abuse account for the remaining three-fifths of the effect of bingeing on all arrests. A substantial role is played by arrests due to violent behavior. For example, a binge of 8+ units causes 126.5 additional arrests for violence related incidences per day, about 83% of the daily mean of violent arrests and 3.7 times more than the mean at night. This effect rises to 143.6 arrests for a binge of 10+ units (95% of the daily mean of arrests due to violent offenses and 22% of the daily mean of *all* arrests) and to 176 for a binge of 12+ units. The estimated impacts of HED on arrests for criminal damage, common assault, and robbery are smaller but always quantitatively important and statistically significant. In contrast, but in line with the reduced form estimates presented in Table 1, the TS2SLS effects on arrests due to theft, burglary, and sexual assault are never statistically significant at conventional levels.[36]

---

[36] Hansen and Waddell (2018) also find that legal access to alcohol is associated with no significant increase in crimes due to rape or robbery. This similarity could lend further credibility to our approach.

To assess the size of our estimates from another perspective, consider the marginal impact on the number of arrests moving from 8+ to 10+ units. This change corresponds to an increase of almost 9% in the number of arrests. Applying this estimate to the whole country implies 100,000 additional arrests per year (Home Office 2015).

As in the case of road accidents, the last column of Table 4 reveals that the AET ratios are always negative, suggesting that our TS2SLS treatment effects can be downward biased estimates of the true HED impact on arrests. The union of CI figures in square brackets are also strongly consistent with the findings so far. Our benchmark TS2SLS estimates are robust to violations of the exclusion restrictions up to 60% or 80% the size of the reduced form estimate for all types of arrests, except for sexual assault, robbery, theft, and burglary, which — apart from robbery — are exactly those crime categories that seem to be insensitive to bingeing in our analysis as well as in previous research (e.g., Hansen and Waddell 2018).

Summary — We find significant impacts of binge drinking on all outcomes. When estimated in relation to the marginal increase from 8+ to 10+ units, the effects are large, implying 6,100 extra road accidents, 63,000 additional A&E attendances every year, and 100,000 additional arrests. Considering the change from, say, 5+ units to 10+ units will lead to even greater effects. Considerably greater effects also emerge if we use a standard 2SLS estimation method rather than the TS2SLS technique in the case of road accidents. We therefore focus on the more conservative picture arising from the TS2SLS estimates. To gain greater confidence in these estimates, we next present and discuss further results from a wide range of sensitivity checks. Adding to the tests we have already presented, these checks will confirm that the benchmark evidence documented so far is robust.

IV.C. Sensitivity Analysis

Figures 7.A and 7.B show the TS2SLS estimates of β_1 for road accidents (RAD records) and A&E attendances respectively, obtained from 11 different specific checks in which each time we change one of the assumptions used in the benchmark estimation. Essentially, in each check, we redefine the inputs of the hedonic forecasting mechanism that associates environmental cues with the anticipation of (near-term) pleasure. In both figures, a bar represents the estimate obtained using 10+ units as the measure of binge drinking and whiskers show the 95% confidence interval.[37] For every outcome domain, the first bar on the left in each block displays the benchmark results for ease of comparison.

The exercises are as follows. First, we perform three changes to the age of the individuals in the control group, using individuals aged 40+ in one case, individuals aged 60+ in another, and individuals aged 31+ in the last (and thus using the whole sample, without dropping any age group). Second, we change three times the definition of age in the treatment group, considering only individuals in the 18–24 range in one case and individuals in the 25–30 in another, and expanding the age range to include individuals aged 18–40 in the last case. Third, we modify the treatment time window twice, examining outcomes recorded over either a shorter period (00:00–05:59) or a longer period (00:00–

---

[37]The same results are found with other measures of heavy episodic drinking, ranging between 6+ and 14+ units.

07:59). Fourth, in order to capture the heterogeneity of drinking patterns observed in Figures 1–2, we change the definition of weekend to include Friday mornings, to then include Monday mornings, and as a further change we include both Thursday and Friday mornings. Finally, to test the validity of our findings further, we report the results from one falsification test in which we redefine Mondays, Tuesdays and Wednesdays as our "placebo" weekend (excluding Saturdays and Sundays) *and* change the treatment age group to individuals aged 31 to 43.[38]

For road accidents (Figure 7.A) — irrespective of whether we look at all accidents, fatal, serious or slight accidents — the benchmark estimate is essentially identical to those found when we change the age bounds in the control group. Interestingly, when the control group comprises individuals aged 31+, the effect estimate is the same as in the benchmark case. Splitting the treatment group into two age groups allows us to determine that about two-thirds of the benchmark effect on accidents is attributable to those aged 18–24, and the remaining one-third to those aged 25–30. Expanding the definition of treatment age up to age 40 leads to slightly (albeit not significantly) larger effects, suggesting that bingeing is likely to occur also among individuals aged above 30. We take our more conservative benchmark results as a better reflection of the effect of HED on road accidents for they mitigate the ambiguities as to where age cutoffs should be drawn.

Reducing the treatment time period to 05:59 leads to larger effects, while or increasing it to 07:59 leads to smaller (but still substantial) effects than those found in the benchmark case.[39] These departures are mainly driven by slight accidents. Changing the definition of weekend by including Friday mornings does not affect our baseline results, while including both Thursday and Friday mornings does reduce the impact somewhat. Including Monday mornings instead leads to estimates that are always statistically insignificant, which suggests that most of the excessive social drinking and its consequences are circumscribed to the official weekend. Finally, the estimates obtained from the falsification test are always statistically insignificant, often small and wrong-signed. Although placebo tests cannot be definitive, these results provide additional support to the identifying assumptions about our instrument.

Similar evidence emerges for A&E attendances shown in Figure 7.B. For each attendance type, the magnitude of the benchmark estimate is quantitatively close to the magnitude of all the other estimates. Again, about two-thirds of all attendances can be attributed to the very young, aged 18–24; likewise, the effect estimates are identical to the benchmark results when the control group age is reduced to 31+. Without exception, the falsification test estimates are reassuringly small and either

---

[38]The reason for this age interval is to match the 13 years in the original treatment age group. We also performed other falsification tests in which we redefined the placebo weekend as Mondays and Tuesdays or Tuesdays and Wednesdays and also changed the treatment time, from night times to day times. The TS2SLS estimates from these alternative tests are similar to those reported in the text and are thus not shown. Moreover, the F-test statistics from the first stage estimation are much lower than those reported in Figure 5, indicating that we cannot reject the null hypothesis that the instruments are uncorrelated with binge drinking status in most cases. Finally, in an additional falsification exercise for the A&E outcomes we replaced the dependent variable with asthma related attendances, which are not expected to depend on HED. The results (presented in the supplementary material) strongly confirm this expectation.

[39]Further perturbations to the time window are carried out in the supplementary material of the Online Appendix, in which we show the estimated effects for five additional time blocks, i.e., 22:00-00:59, 22:00-01:59, 22:00–02:59, 22:00–03:59, and 22:00–04:59. These stricter definitions of time are likely to rule out commuters (see also the discussion in Section II.B). They all strongly support the baseline results.

insignificant or wrong-signed (or both).

Figure 7.C presents the sensitivity results on arrests. Given the data, we are more limited with the number of checks we can perform. In this case, we change the definition of control group (to include individuals aged 30–50, rather than individuals aged 50 or more as in the baseline analysis), the treatment time (18:00–23:59 rather than 00:00–05:59), and we change the definition of weekend as before (including either Friday mornings, Thursday and Friday mornings, or Monday mornings), and we perform the same falsification test.

Lowering the age of the control group generally increases the effect on all types of arrests compared to the benchmark case, but this increase is never statistically significant. The benchmark estimates are robust to changes in the definition of weekend that includes Friday mornings, and Thursday and Friday mornings albeit lower in magnitude. Bringing the treatment time forward, instead, usually reduces the effect on arrests more substantially, perhaps because the impact of binge drinking on arrests becomes evident only later in the night. And so does the extended definition of weekend that includes Monday mornings. Finally, the estimates obtained from the falsification test are always statistically insignificant, much smaller than the baseline estimate, and often of the wrong sign.

IV.D. Further Robustness and Other Exercises

Spillover Effects on Road Accidents

Spillover effects in excessive social drinking are part and parcel of the problem. This is why so far we have not been concerned in separating out the effects of own HED from the effects of others' HED. For instance, young adults on weekend nights may not get hurt on account of their own drunkenness but on that of their peers or those around them. Similarly, in the case of road traffic accidents, those accidents that involve more than one vehicle in which some of the drivers (or passangers) are not inebriated but others are according to police records are all defined as being part of our treatment. This is not the case in Levitt and Porter's (2001) application, because they are interested in the estimating the consequences caused by drinkers, while we are interested in assessing the overall impact associated with bingeing.

It is however useful to have an idea of the extent of spillovers mentioned above. We do this for road accidents using detailed information contained in the RAD records according to which we can distinguish between accidents involving only one vehicle (but possibly more than one casualty) and accidents involving two or more vehicles. The data permit us also to isolate accidents in which there is only one casualty (but possibly more than one vehicle) from accidents in which two or more casualties are recorded. Our reasoning is that if most of the effects are concentrated among cases with multiple casualties or multiple vehicles, then spillovers are likely to play a key role and we might over-inflate the role of HED.

Table 5 summarizes the estimates for the 10+ unit case (the estimates for 8+ and 12+ units are in the Online Appendix). Of the 83 additional accidents reported in Table 2, we find that 52 (nearly 63%) are accounted for by cases in which there is only one vehicle per accident, while the other 31 are

related to cases with two vehicles or more. When we focus on casualties, the results are essentially the same: about 65% of the whole effect on accidents is attributable to accidents that involve one casualty only, and the remaining 35% refer to accidents with multiple casualties. These findings hold true across accident types, expect for fatalities by number of casualties.

Bingeing is a social phenomenon, whereby young adults associate with other drinking peers during weekends. But the simple message from the estimates in Table 5 is that, although spillover effects on road traffic accidents play a role, a large fraction of the effects on outcomes seems to be driven by own drinking decisions rather than to the drinking behavior of others. Better data and more research on this important issue are needed.

First Stage Estimation with Time Use Data

An interesting sensitivity exercise is to see how our estimates change when we use a first-stage data source different from the HSE. To this purpose, we employ the UK-TUS data which has detailed diary information. This allows us to explore whether there are differences in the drinking time patterns between the young and the old during the weekend, which could have consequences on the outcomes under analysis.

The UK-TUS data however has two shortcomings. The first is that UK-TUS respondents do not report whether or not they consume alcohol while engaging in a particular activity. With no information on alcohol involvement, we cannot define heavy episodic drinking on the basis of the number of alcoholic units on a single occasion, as we did so far. A binge is therefore now defined to occur when an individual is in a pub, restaurant or bar while *not* eating at specific times of the evening and night. Excluding eating is an attempt to increase the probability that our measure does identify an activity that is predominantly drink related. Moreover, focusing on specific public places allows us to add to the group time coordination considered so far (weekends) the exposure to spatial alcohol cues (such as pubs and bars) that may play a role in ethanol seeking behavior and bingeing.

A second shortcoming is that the UK-TUS refers to an earlier period than that covered by our second-stage data sets, and it also dates before the 2003 Licensing Act that affected opening and closing hours of public houses and clubs. To account for potential differences in behavior associated with changes in the legal environment, we perform the analysis at different times of the day and night.

The TS2SLS effects are reported in the Online Appendix, where we also show the corresponding first-stage results. Looking at the estimates across different hour blocks, we cannot find evidence that the old binge earlier in the day and the young binge later at night. In addition, compared to the benchmark estimates the new effects are substantially greater. Even the smallest estimates, typically found when the time window for a binge is restricted to the 20:00–23:00 period, are twice as large in the case of road accidents and arrests, and almost four times greater in the case of A&E attendances. When the time window includes hours after midnight, the estimates for all outcomes become much larger. The lack of information on actual alcohol involvement might inflate the effect that we attribute to binge drinking, as it could capture other aspects that are not necessarily related to ethanol seeking behavior. Based on these considerations, we favor our more conservative estimates found from the

benchmark analysis.

Using the 2011 HSE Drinking Diary Data

First stage estimation — In this check we use the 2011 drinking diary information from the HSE rather than the standard HSE alcohol consumption data used so far in the first stage. We run this exercise with different cutoffs for drinking times (e.g., after 6 p.m., or after 9 p.m., or any time), and adding further month controls in the analyses for road accidents and A&E attendances.

The TS2SLS results are strongly consistent with (albeit substantially higher than) those shown in the benchmark analysis. For instance, using the 9 p.m. threshold, we find that drinking 10+ units of alcohol leads to an additional 173.1 accidents per weekend day among individuals aged 18–30, as opposed to 83.1 extra accidents estimated using the standard HSE information. This implies an average effect of 41.7% rather than 18.6% as we have from the benchmark analysis. Similar overstatements of the impact of HED on all accidents emerge across the different definitions of bingeing, for all drinking time thresholds, and across most of the accident types. The same is found in the case of the other two outcome domains. We take these results as further evidence that our benchmark estimates are not affected by the fact that we use information on the heaviest drinking day of the week and not total drinks consumed in each day of the week, and, if anything, they are likely to be considerably more conservative.

Location — So far we have emphasized the importance of time-of-day drinking patterns, although we have also recognized the importance of group coordination over places where alcohol is consumed. Our concern about location is motivated by the observation that the young could be more likely to engage in HED while being out, while the old could do the same while at home. If most of the outcomes we observe pick up behaviors that are only (or predominantly) associated with drinking out, then we may wrongfully attribute them to how much people drink when actually they depend on where people drink.

To better account for location therefore we use data from the 2011 HSE drinking diaries, which contain details not only on the times of alcohol consumption but also the places where such consumption occurred. Drinking at home or at a neighbor's home is defined as "drinking in", while drinking in all other locations (mainly pubs, bars, restaurants, night clubs, and friends' places not nearby) is defined as "drinking out".

We then repeat our entire analysis distinguishing such two broad locations. For both location types, we find impacts of bingeing on outcomes that are large, significant, and fully consistent with (and always greater than) those shown in Tables 2–4. Importantly, the estimates reported in the Online Appendix indicate that, across all outcome domains and subcategories as well as across all HED definitions, there is no statistically significant difference in the effects of bingeing by location. That is, the effects of HED found for young adults drinking out on weekends are the same as the effects found for drinking in. It is thus unlikely that our analysis fails to account for episodes of excessive drinking at home and incorrectly attributes them to how much, rather than where, people drink.

Police Deployment

An issue of selection other than those discussed in subsections II.B and III.D may arise because of police presence. For example, there might be more policy deployment in the areas where young adults are known to congregate in weekend evenings and nights, e.g., city centers around pubs and nightclubs. This may contribute to produce the patterns observed in the case of arrests, since most criminal activities (even if unrelated to excess drinking) could be detected more promptly and arrests effected more readily. However, police presence could deter individuals (including young heavy drinkers) from engaging in criminal behavior. The net effect on arrests therefore is ambiguous, and we leave the analysis of this interesting issue to future research.[40]

Here we use data on the number of police officers on duty provided by two police forces (the Durham Constabulary and the London Metropolitan Police Service) to document whether there is more police deployment in weekend evenings/nights or not. Unfortunately, the data are not detailed enough for us to see where the deployment takes place within each geographic area.[41] In both Durham and London the patterns are very similar. Most officers are deployed in the middle of the day during weekdays and in the evenings/nights during weekends. This finding therefore provides evidence that policing may track binge drinking.

In turn, this suggests that HED behavior leads to police allocation decisions (i.e., police deployment could be seen as an outcome of bingeing), while the opposite relationship is arguably groundless. We thus analyze police deployment as an outcome, for which the dependent variable is given by the average number of police officers on duty during the night (00:00–05:59) in the two police forces per 10,000 residents, and following the specifications used in the main analysis of the paper as closely as possible. Notice that for this outcome we can only consider the effect of weekend nights. From our data, in fact, we cannot identify if (and the extent to which) police officers are deployed on the basis of the public's age, and thus we cannot exploit any variation in individual age for identification purposes.

We find significant reduced form estimates, according to which we have an additional 3.2 officers per 10,000 residents at night during the weekend in both locations (corresponding to a 8% increase over the mean). The TS2SLS results imply a much larger increase (26 additional officers on duty per 10,000 residents in the case of 10+ units). But this is driven by the impact observed in Durham, while the estimates for London are not statistically significant. This might be due to the fact that the first stage estimates found with the London MPS data are weak, with F-test values always below 10. From the bulk of all results, we infer that there appears to be some evidence that HED increases police deployment, but more research is needed with better data.

Gender

---

[40] Arguably, the effect of endogenous police allocation on road accidents and A&E attendances instead is likely to be negligible.

[41] The data, which were obtained through Freedom of Information requests, cover the same 16 weeks in four 6–hour daily blocks as those in the arrest data. For each hour of the day we know the number of police officers who were recorded on duty. More information on the data, the analysis, and the results is presented in the supplementary material of the Online Appendix.

Although binge drinking has been usually seen as an issue for men, recent research and media commentaries point out the increasing prevalence of HED among young women (e.g., Motluk 2004; Young et al. 2005; Miller et al. 2007; Lyons and Willott 2008). It is interesting, therefore, to see whether our TS2SLS benchmark estimates vary by sex, although more work is left for future research. The results of this exercise are reported in the Online Appendix.

We find statistically significant differences in all road outcomes for 8+ and 10+ units, however for higher numbers of alcoholic units the differences in fatal and serious road accidents are no longer statistically significant. Bingeing leads to more road accidents involving men. For instance, drinking 10+ units of alcohol implies 20 additional accidents per day caused by women and 70 extra accidents caused by men. Similar gender differences emerge for each type of accident. This suggests that about 22% of all road accidents can be associated with female HED behavior.[42] Turning to the effect on A&E attendances, the gender gap is much less pronounced and statistically significant only for open wounds, in which case the estimates are larger for men. Finally, there are large and significant gender differences in the effect of bingeing on arrests, both directly and indirectly related to alcohol, with female bingers representing about one fifth of all arrests.

These results therefore confirm the view that binge drinking is primarily associated with men (Holmila and Raitasalo 2005; Rahav et al. 2006). But, with almost one quarter of road accidents caused by female heavy drinkers, one fifth of arrests accounted for by women, and fairly similar numbers of hospital emergency attendances, they also lend support to the growing relevance of episodic alcohol abuse amongst young women. This deserves more attention in future research.

IV.E. Externalities and Implications

A full cost-benefit analysis of heavy episodic drinking is infeasible, essentially because the utility that individuals obtain from bingeing and the value to alcoholic beverage for producers and retailers cannot be easily accounted for (Levitt and Porter 2001; Cawley and Ruhm 2012). In what follows, therefore, to ascertain some of the social cost associated with bingeing, we compute the externalities that are attributable to road accidents. The RAD, in fact, are administrative records that cover the whole country allowing us to make no assumption to scale estimates up to the national level. We use an approach similar to that proposed by Levitt and Porter (2001) who assess the externality generated by deaths due to drunk driving. But we do not assume away the cost borne by binge drinkers, as this is arguably part of the total cost of bingeing that society must face and it is unclear whether binge drinkers take the risk of accidents and injuries (on others and themselves) fully into account.

Let k label a type of accident (fatal, serious, and slight). For each k, the weekly cost of bingeing is $\beta_{1k}^{(q)} \mu^{(q)} c_k$, where $\beta_{1k}^{(q)}$ is the TS2SLS effect on the number of accidents of type k per weekend when drinking at least q units of alcohol, $\mu^{(q)}$ is the observed fraction of individuals aged 18–30 who drink q units or more at weekends, and c_k is the official unit cost associated with accidents of

---

[42]For Canada, Carpenter, Dobkin, and Warman (2016) find large increases in fatal road accidents at the MLDA which are statistically significant for men but not for women. Their results indicate that the MLDA is effective in reducing extreme drinking (which can be correlated, but does not coincide, to HED among young men.

type k as published by government statistical agencies. The total cost at q units or more is then $C^{(q)} = \sum_k \beta_{1k}^{(q)} \mu^{(q)} c_k$. Repeating this computation at all units gives us an estimate of the cost at different points of the distribution of alcohol consumption. Let q_0 be a baseline (or reference) unit level, below which binge drinking does not occur. Our externality measure, which is given by

$$\Delta(q_0, q) = \sum_k \beta_{1k}^{(q)} \mu^{(q_0)} c_k - C^{(q_0)},$$

captures what the additional cost of drinking q units would be with respect to the reference case, provided the fraction of young people who drink q units is the same as at baseline, $\mu^{(q_0)}$. For example, $\Delta(5, 8)$ indicates the externality when the reference cost is computed using 5+ units and the counterfactual cost is calculated with 8+ units keeping the share of people drinking 8 or more units at $\mu^{(5)}$.

In our exercise we set q_0 at 5+ units. We have two reasons to do so. First, it is close to the maximum intake of 3 units a day recommended by practitioners (Royal College of Physicians 2001): five units correspond to about two pints of beer or two glasses of wine and do not represent a binge. Second, the estimates found at lower levels of alcohol intake are often characterized by low F-test values in the first stage.

Table 6 summarizes the results. We present three different estimates, namely $\Delta(5, 8)$, $\Delta(5, 10)$, and $\Delta(5, 12)$, and show the figures by type of accident separately. Around each point estimate, we report the 95% confidence interval obtained using the corresponding bootstrap standard error of the TS2SLS effects. Further details are presented in the Online Appendix.

Between 2008 and 2010, 58.1% of 18–30 year-old individuals drank 5 or fewer units of alcohol every day. Using the estimates in Figure 6.A and the published unit costs c implies a total cost $C^{(5)}$ of £1.4 billion per year (in 2015 prices). If the same proportion of young people were to have a binge with at least 8 units, Table 6 suggests we expect a negative externality of £920 million. The externality more than doubles to £2.02 billion if heavy episodic drinking is defined on 12+ units. About 60% of the total burden is accounted for by accidents that involve at least one fatality, and more than one-quarter by accidents with serious injuries.

Consider the £1.52 billion estimate. According to industry estimates, each adult consumed 9.9 liters of pure alcohol in 2011 (Sheen 2013), which translate into a total of 51 billion of alcoholic units drunk in England in that year. Our estimate then implies a negative externality of 3p per alcoholic unit, representing an increase of more than 10% in the current alcohol duty rate per unit of pure alcohol. This in turn is equivalent to an increase in the duty rate of 14% for a pint of beers of 4.5% ABV and of 14% for a bottle of wine of 14% ABV.

Taxing alcohol consumption, however, is a blunt policy instrument and is likely to introduce distortions into consumption decisions which this calculation does not account for.[43] We then look at the same issue from a different angle. According to the 2008–2010 National Travel Surveys, 157.1

---

[43]On this issue, see the interesting insights and findings in Griffith, Leicester, and O'Connell (2013) and Griffith, O'Connell, and Smith (2017).

billion miles were driven each year during the 2006–2009 period in England. Of these, 18.3% (which translates into 28.8 billion miles) were driven by individuals aged 20–29 (close to our group of 18–30 years old). The $\Delta(5, 10)$ estimate in Table 6 thus implies a negative externality of about 5.3 pence per mile driven by young people who drink 10+ units of alcohol in weekends. This is about one-third of the Levitt and Porter's (2001) estimated externality due to drunk driving of 18 pence per mile driven (measured in 2015 prices, or 15 cents in 1993 prices as in their published work). At current arrest rates for drunk driving, the Pigouvian tax that internalizes our estimated externality is approximately £22,800 per arrest, which is nine times more than the maximum fine currently set by the UK government on drunk driving. If we consider arrests that are indirectly related to alcohol and due to violent offenses (which are the most frequent type of arrests in our sample), the corresponding Pigouvian tax is approximately £5,000 per arrest.

Alternative Unit Costs — Although the unit costs available from statistical agencies provide an official benchmark, they are reported as point estimates without accompanying measures of sampling and nonsampling errors that may be nonrandom and large. Moreover, there are alternative cost sources that could be used. To account for the uncertainty in the measurement of unit costs and examine the sensitivity of the estimates in Table 6, we recalculated our estimates using a wide array of alternative unit costs.

The details of these computations are in the Online Appendix, while the results obtained from this new analysis are shown in Figure 8. We rank all the 27 alternative values from the lowest on the left to the highest on the right of the figure and distinguish the results based on the 10+ unit definition from those based on the 8+ and 12+ unit definitions. We report both the point estimates and, for the 10+ unit case, their 95% confidence bands. The confidence bands for the other two definitions overlap with those shown in the figure. To ease comparisons we also present the benchmark externalities described before, which are represented by the three horizontal lines and the 95% confidence interval in correspondence to the 10+ unit case.

Focusing on the estimates based on the 10+ unit definition, we find that the road accident externality ranges from £700 million to £5.40 billion, for an average point estimate of £1.95 billion per year. The majority of these alternative values lie within the 95% confidence interval around the benchmark estimate. We have five estimates that fall significantly outside this range, one below and four above, which are found when we apply lower and higher unit cost estimates of fatal accidents, respectively. We thus take the results in Figure 8 as evidence that our £1.52 billion estimate is a conservative measure of the annual HED externality arising from road accidents.

V. Conclusion

A key contribution of this paper is to develop a strategy for estimating the impact of binge drinking relying on the joint double differences between the young and the old and between weekends and weekdays, which are expected to intensify partying and drinking among young adults. Using a variety

of unique data and a two-sample instrumental variables estimation procedure, our study is the first to provide a detailed picture of the consequences of heavy episodic drinking on road accidents, accident and emergency attendances, and arrests. We find that bingeing increases the number of road accidents by 18.6% and fatal road accidents by 72%, injury-related A&E attendances by 6.6%, and the number of arrests for all alcohol related incidences by 71%. These results imply nearly 6,100 extra road accidents (including 300 more fatalities), roughly 63,000 additional A&E attendances every year, and 100,000 additional arrests. The externality per mile driven by a binge drinker is about 5.3 pence, and the punishment that internalizes this externality is equivalent to a fine of £22,800 per drunk driving arrest. The magnitude of such effects is substantial, considering that they are obtained using rather conservative definitions of HED.

To assess the robustness of our results we performed several sensitivity checks, allowing for violations of the exclusion restriction, changing definition of binge drinking, varying treatment and control age groups and times, and using alternative first stage data. We have also considered the impacts of moderate drinking, selectivity into different types of weekend activities by age, the role played by illicit drugs, and the possibility of spillover effects (in the context of road accidents). The results of all such additional exercises provide evidence that confirms our estimated benchmark effects. In the case of road accidents, we also estimated 2SLS models which yield effect estimates that always exceed those obtained with the TS2SLS procedure.

These findings add to the literature on the impact of risky behaviors on health and crime. Alcohol has been shown to influence a variety of behaviors, including schooling, employment, and teenage fertility. Binge drinking, which brings young people together for short episodes of heavy drinking during the weekend, is likely to elevate the risks of offenses and lower performance. Such negative consequences are the argument of intense policy debates in several industrialized economies, with proposals that range from minimum unit pricing policies (as introduced by the Scottish Government in May 2018) to reforms of the whole system of alcohol excise taxes and restrictions on alcohol availability by age, day of the week, and premise.

The methodology we present provides a simple, but powerful, tool for analyzing the causal effect of binge drinking when reliable measures of alcohol involvement are not available in the same data set that has information on outcomes. The simultaneous measurement of alcohol consumption and outcomes may be problematic if the data cover only small or selected populations, such as college students or arrestees, since their generalizability is questionable. It may also be problematic if systematic misreporting is an issue, or if interviewee's response behaviour is influenced by the nature of the interview setting. The two-sample approach instead avoids such problems. It is likely to be useful in several other substantive applications where concerns about data availability are similar to ours and information on risky behaviors (e.g., illicit drug use, smoking, and unprotected sex) is not collected with outcomes, such as salaries, teen pregnancies, and sexually transmitted diseases.

References

Acemoglu, Daron, and Simon Johnson. 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy*, 115(6): 925–985.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling." *Journal of Human Resources*, 40(4): 791–821.

Adams, Scott, McKinley L. Blackburn, and Chad D. Cotti. 2012. "Minimum Wages and Alcohol-Related Traffic Fatalities among Teens." *Review of Economics and Statistics*, 94(3): 828–840.

Anderson, D. Mark, Benjamin Crost, and Daniel I. Rees. 2016. "Wet Laws, Drinking Establishments, and Violent Crime." Unpublished Manuscript, May.

Anderson, D. Mark, Benjamin Hansen, and Daniel I. Rees. 2013. "Medical Marijuana Laws, Traffic Fatalities, and Alcohol Consumption." *Journal of Law and Economics*, 56(2): 333–369.

Angrist, Joshua D., and Alan B. Krueger. 1992. "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples." *Journal of the American Statistical Association*, 87(418): 328–336.

Arellano, Manuel, and Costas Meghir. 1992. "Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets." *Review of Economic Studies*, 59(3): 537–557.

Baker, Carl. 2015. *Accident and Emergency Statistics.* London: House of Commons Library, Briefing Paper Number 6964.

Barker, Jacqueline M., Laura H. Corbit, Donita L. Robinson, Christina M. Gremel, Rueben A. Gonzales, L. Judson Chandler. 2015. "Corticostriatal Circuitry and Habitual Ethanol Seeking." *Alcohol*, 49(8): 817–824.

Barker, Jacqueline M., and Jane R. Taylor. 2014. "Habitual Alcohol Seeking: Modeling the Transition from Casual Drinking to Addiction." *Neuroscience and Biobehavioral Reviews*, 147: 281–294.

Becker, Gary. S. 1996. *Accounting for Tastes.* Cambridge, MA.: Harvard University Press.

Bernheim, D. Douglas, and Antonio Rangel. 2004. "Addiction and Cue-Triggered Decision Processes." *American Economic Review*, 94(5): 1558–1590.

Bleakley, Hoyt. 2007. "Disease and Development: Evidence from Hookworm Eradication in the American South." *Quarterly Journal of Economics*, 122(1): 73–117.

Biderman, Ciro, De Mello, João M.P., and Alexandre Schneider. 2010. "Dry Laws and Homicides: Evidence from the São Paulo Metropolitan Area." *Economic Journal*, 120(543): 157– 182.

Brismar, Bo, and Bo Bergman. 1998. "The Significance of Alcohol for Violence and Accidents." *Alcoholism: Clinical and Experimental Research*, 22(7, Supplement): 299S–306S.

Card, David, and Gordon B. Dahl. 2011. "Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior." *Quarterly Journal of Economics*, 126(1): 103– 143.

Carneiro, Pedro, Katrine V. Løken, and Kjell G. Salvanes. 2015. "A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children." *Journal of Political Economy*, 123(2): 365–412.

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

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

Carpenter, Christopher, and Carlos Dobkin. 2015. "The Minimum Legal Drinking Age and Crime." *Review of Economics and Statistics*, 97(2): 521–524.

Carpenter, Christopher, and Carlos Dobkin. 2017. "The Minimum Legal Drinking Age and Morbidity in the United States." *Review of Economics and Statistics*, 99(1): 95–104.

Carpenter, Christopher, Carlos Dobkin and Casey Warman. 2016. "The Mechanisms of Alcohol Control." *Journal of Human Resources*, 51(2): 328-356.

Carrell, Scott E., Mark Hoekstra, and James E. West. 2011. "Does Drinking Impair College Performance? Evidence from a Regression Discontinuity Approach." *Journal of Public Economics*, 95(1): 54–62.

Cawley, John, and Christopher J. Ruhm. 2012. "The Economics of Risky Health Behaviors." In Mark V. Pauly, Thomas G. McGuire, and Pedro P. Barros (eds.) *Handbook of Health Economics*, vol. 2. New York, NY: Elsevier, pp. 95–199.

Chaloupka, Frank J., Henry Saffer and Michael Grossman. 1993. "Alcohol-control policies and motor-vehicle fatalities." *Journal of Legal Studies*, 22(1): 161–186.

Conley, Timothy G., Christian B. Hansen, and Peter Rossi. 2012. "Plausibly Exogenous." *Review of Economics and Statistics*, 94(1): 260–272.

Conlin, Michael, Stacy Dickert-Conlin, and John Pepper. 2005. "The Effect of Alcohol Prohibition on Illicit-Drug-Related Crimes." *Journal of Law and Economics*, 48(1): 215–234.

Cook, Philip J. and Michael J. Moore. 2000. "Alcohol." In Anthony J. Culyer and Joseph. P. Newhouse (eds.), *Handbook of Health Economics*, vol. 1, part B. New York, NY: Elsevier, pp. 1629–1673.

Corbit, Laura H., Hong Nie, and Patricia H. Janak. 2012. "Habitual Alcohol Seeking: Time Course and the Contribution of Subregions of the Dorsal Striatum." *Biological Psychiatry*, 72(5): 389–395.

Crost, Benjamin, and Santiago Guerrero. 2012. "The effect of alcohol availability on marijuana use: Evidence from the minimum legal drinking age." *Journal of Health Economics*, 31(1): 112–121.

Dahl, Gordon, and Stefano DellaVigna. 2009. "Does Movie Violence Increase Violent Crime?" *Quarterly Journal of Economics*, 124(2): 677–734.

Dee, Thomas S. 1999. "State Alcohol Policies, Teen Drinking and Traffic Fatalities." *Journal of Public Economics*, 72(2): 289–315.

Dee, Thomas S., 2001. "The Effects of Minimum Legal Drinking Ages on Teen Childbearing." *Journal of Human Resources*, 36(4): 823–838.

Dee, Thomas S., and William N. Evans. 2003. "Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates." *Journal of Labor Economics*, 21(1): 178–209.

DiNardo, John and Thomas Lemieux. 2001. "Alcohol, Marijuana, and American Youth: The Unintended Consequences of Government Regulation." *Journal of Health Economics*, 20(6): 991–1010.

Domjan, Michael, Brian Cusato, and Ronald Villarreal. 2000. "Pavlovian Feed-Forward Mecha- nisms in the Control of Social Behavior." *Behavioral and Brain Sciences*, 23(2): 235–270.

Dragone, Davide, Giovanni Prarolo, Paolo Vanin, Giulio Zanella. Forthcoming. "Crime and the Legalization of Recreational Marijuana." *Journal of Economic Behavior & Organization*.

Duncan, Greg J., Johanne Boisjoly, Michael Kremer, Dan M. Levy, and Jacque Eccles. 2005. "Peer Effects in Drug Use and Sex Among College Students." *Journal of Abnormal Child Psychology*, 33(3): 375–385.

Ekholm, O., K. Strandberg-Larsen, K. Christensen, and M. Gronbaek. 2008. "Comparison of Assessment Methods for Self-Reported Alcohol Consumption in Health Interview Surveys." *European Journal of Clinical Nutrition*, 62(2): 286–291.

Fella, Giulio, and Giovanni Gallipoli. 2014. "Education and Crime over the Life Cycle." *Review of Economic Studies*, 81(4): 1484–1517.

Fletcher, Jason, and Ryne Marksteiner. 2017. "Causal Spousal Health Spillover Effects and Implications for Program Evaluation." *American Economic Journal: Economic Policy*, 9(4): 144–166.

Furlong, Andy, and Fred Cartmel. 2007. *Young People and Social Change.* McGraw-Hill: Open University Press.

Grant, Bridget F., Deborah A. Dawson, Frederick S. Stinson, S. Patricia Chou, Mary C. Dufour, Roger P. Pickering. 2004. "The 12-Month Prevalence and Trends in DSM-IV Alcohol Abuse and Dependence: United

States, 19911992 and 20012002." *Drug and Alcohol Dependence*, 74(3): 223–234.

Griffith, Rachel, Andrew Leicester, and Martin O'Connell. 2013. "Price-Based Measures to Reduce Alcohol Consumption." IFS Briefing Note BN138. London: Institute for Fiscal Studies, March.

Griffith, Rachel, Martin OConnell and Kate Smith. 2017. "Design of Optimal Corrective Taxes in the Alcohol Market." IFS Working Paper W17/02. London: Institute for Fiscal Studies, January.

Gul, Faruk, and Wolfgang Pesendorfer. 2007. "Harmful Addiction." *Review of Economic Studies*, 74(1): 147–172.

Hamermesh, Daniel S., Caitlin Knowles Myers, and Mark L. Pocock. 2008. "Cues for Timing and Coordination: Latitude, Letterman, and Longitude." *Journal of Labor Economics*, 26(2): 223–246.

Hansen, Benjamin. 2015. "Punishment and Deterrence: Evidence from Drunk Driving." *American Economic Review*, 105(4): 1581–1617.

Hansen, Benjamin, and Glen R. Waddell. 2018. "Legal Access to Alcohol and Criminality." *Journal of Health Economics*, 57(1): 277–289.

Heaton, Paul, 2012. "Sunday Liquor Laws and Crime." *Journal of Public Economics*, 96(1–2): 42–52.

Holmila, Marja, and Kirsimarja Raitasalo. 2005. "Gender Differences in Drinking: Why Do They Still Exist?" *Addiction*, 100(12): 1763–1769.

Home Office. 2015. *Police Powers and Procedures, England and Wales.* London: Home Office, Statistical Bulletin 07/15, November.

Inoue, Atsushi, and Gary Solon. 2010. "Two-Sample Instrumental Variables Estimators." *Review of Economics and Statistics*, 92(3): 557–561.

Jacob, Brian A., and Lars Lefgren. 2003. "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime." *American Economic Review*, 93(5): 1560–1577.

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

Laibson, David, 2001. "A Cue-Theory of Consumption." *Quarterly Journal of Economics*, 116(1): 81–119.

Levitt, Steven D., and Jack Porter. 2001. "How Dangerous Are Drinking Drivers?" *Journal of Political Economy*, 109(6): 1198–1237.

Lindo, Jason M., Peter Siminski, and Isaac D. Swensen. 2018. "College Party Culture and Sexual Assault." *American Economic Journal: Applied Economics*, 10(1): 236–265.

Lindo, Jason M., Peter Siminski, and Oleg Yerokhin. 2016. "Breaking the Link Between Legal Access to Alcohol and Motor Vehicle Accidents: Evidence from New South Wales." *Health Economics*, 25(7): 908–928.

Lisdhal Medina, Krista, Alecia D. Schweinsburg, Mairav Cohen-Zion, Bonnie J. Nagel, and Susan F. Tapert. 2007. "Effects of Alcohol and Combined Marijuana and Alcohol Use During Adolescence on Hippocampal Volume and Asymmetry." *Neurotoxicology and Teratology*, 29(1): 141–152.

Lyons, Antonia C., and Sara A. Willott. 2008. "Alcohol Consumption, Gender Identities and Women's Changing Social Positions." *Sex Roles*, 59(9–10): 694–712.

Markowitz, Susan, and Michael Grossman. 2000. "The Effects of Beer Taxes on Physical Child Abuse." *Journal of Health Economics*, 19(2): 271–282.

Marcus, Jan, and Thomas Siedler. 2015. "Reducing Binge Drinking? The Effect of a Ban on Late-Night Off-Premise Alcohol Sales on Alcohol-Related Hospital Stays." *Journal of Public Economics*, 123(1): 55–77.
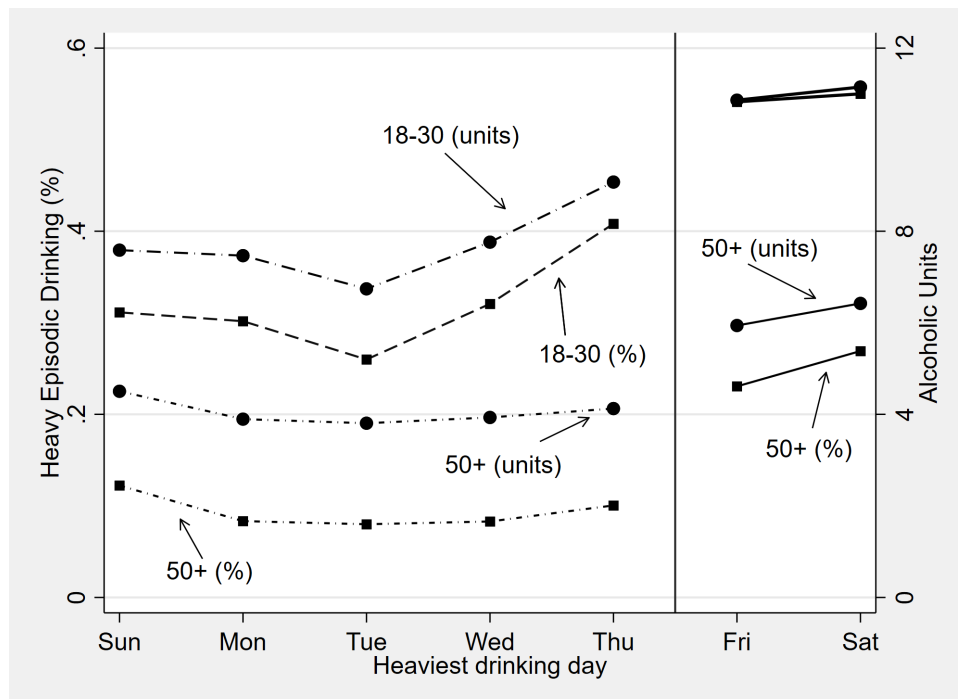
Measham, Fiona. 1996. "The 'Big Bang' Approach to Sessional Drinking: Changing Patterns of Alcohol Consumption amongst Young People in North West England." *Addiction Research and Theory* 4(3): 283–299.

Midanik, Lorraine. 1988. "Validity of Self-reported Alcohol Use: A Literature Review and Assessment." *British Journal of Addiction*, 83(9): 1019–1029.

Midanik, Lorraine T., Tammy W. Tam, and Constance Weisner. 2007. "Concurrent and Simultaneous Drug and Alcohol Use: Results of the 2000 National Alcohol Survey." *Drug and Alcohol Dependence*, 90(1): 72–80.

Miller, Jacqueline W., Timothy S. Naimi, Robert D. Brewer, Sherry Everett Jones. 2007. "Binge Drinking and Associated Health Risk Behaviors Among High School Students." *Pediatrics*, 119(1): 76–85.

Motluk, Alison. 2004. "Intemperate Society." *New Scientist*, 183(2461): 28–33.

Naimi, Timothy S., Robert D. Brewer, Ali Mokdad, Clark Denny, Mary K. Serdula, and James S. Marks. 2003. "Binge Drinking Among US Adults." *Journal of the American Medical Association*, 289(1): 70–75.

National Institute on Alcohol Abuse and Alcoholism. 2005. *Helping Patients Who Drink Too Much: A Clinician's Guide.* Bethesda, MD: National Institutes of Health.

Nelson, Jon P. 2015. "Binge Drinking and Alcohol Prices: A Systematic Review of Age-Related Results from Econometric Studies, Natural Experiments and Field Studies." *Health Economics Review*, 5:6, February.

NHS England. 2015. *A&E Attendances and Emergency Admissions.* Available at: http://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity.

Norman, Paul, Paul Bennett, and Helen Lewis. 1998. "Understanding Binge Drinking among Young People: An Application of the Theory of Planned Behaviour." *Health Education Research*, 13(2): 163–169.

Nunn, Nathan, and Nancy Qian. 2014. "US Food Aid and Civil Conflict." *American Economic Review*, 104(6): 1630–1666.

OECD. 2015. *Tackling Harmful Alcohol Use: Economics and Public Health Policy.* OECD Publishing. <http://dx.doi.org/10.1787/9789264181069-en>

Ohsfeldt, Robert L., and Michael A. Morrisey. 1997. "Beer Taxes, Workers' Compensation and Industrial Injury." *Review of Economics and Statistics*, 79(1): 155–160.

Oster, Emily. Forthcoming. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business and Economic Statistics.*

Parker, Howard, and Lisa Williams. 2003. "Intoxicated Weekends: Young Adults' Work Hard-Play Hard Lifestyles, Public Health and Public Disorder." *Drugs: Education, Prevention, and Policy*, 10(4): 345–367.

Pennings, Ed J.M., Arthur P. Leccese, and Frederik A. de Wolff. 2002. "Effects of Concurrent Use of Alcohol and Cocaine." *Addiction*, 97(7): 773–783.

Rahav, Giora, Richard Wilsnack, Kim Bloomfield, Gerhard Gmel, and Sandra Kuntsche. 2006. "The Influence of Societal Level Factors on Men's and Women's Alcohol Consumption and Alcohol Problems." *Alcohol and Alcoholism*, 41(Suppl. 1): i47–i55.

Robbe, H. 1998. "Marijuana's Impairing Effects on Driving are Moderate when Taken Alone but Severe when Combined with Alcohol." *Human Psychopharmacology: Clinical and Experimental*, 13(S2): S70–S78.

Royal College of Physicians. 2001. *Alcohol: Can the NHS Afford It? Recommendations for a Coherent Alcohol Strategy for Hospitals.* London, February.

Ruhm, Christopher J. 1996. "Alcohol Policies and Highway Vehicle Fatalities." *Journal of Health Economics*, 15(4): 435–454.

Satyanath, Shanker, Nico Voigtländer, and Hans-Joachim Voth. 2017. "Bowling for Fascism: Social Capital and the Rise of the Nazi Party." *Journal of Political Economy*, 125(2): 478–526.

Savola, Olli, Onni Niemelä, and Matti Hillbom. 2005. "Alcohol Intake and the Pattern of Trauma in Young Adults and Working Aged People Admitted after Trauma." *Alcohol and Alcoholism*, 40(4): 269–273.

Sheen, David. 2013. *Statistical Handbook 2013. British Beer and Pub Association.* London: Brewing Publications.

Simpson, T., Murphy, N., and D.F., Peck. 2001. "Saliva Alcohol Concentrations in Accident and Emergency Attendances." *Emergency Medicine Journal*, 18(4): 250–254.

Smart, Reginald G. and Alan C. Ogborne. 2000. "Drug Use and Drinking Among Students in 36 Countries." *Addictive Behaviors*, 25(3): 455–460.

Stewart, Bernard W., and Christopher P. Wild (eds.). 2014. *World Cancer Report 2014.* WHO Press.

Taylor, B., H.M. Irving, F. Kanteres, R. Room, G. Borges, C. Cherpitel, T. Greenfield, J. Rehm. 2010. "The More You Drink, The Harder You Fall: A Systematic Review and Meta-Analysis of How Acute Alcohol Consumption and Injury or Collision Risk Increase Together." *Drug and Alcohol Dependence*, 110(1): 108–116.

Terza, Joseph V. 2002. "Alcohol Abuse and Employment: A Second Look. *Journal of Applied Econometrics*, 17(4): 393–404.

Thomas, Kevin V., Lubertus Bijlsma, Sara Castiglioni, Adrian Covaci, et al. 2012. "Comparing Illicit Drug Use in 19 European Cities through Sewage Analysis." *Science of the Total Environment*, 432(1): 432–439.

Wechsler, Henry, Andrea Davenport, George Dowdall, Barbara Moeykens, and Sonia Castillo. 1994. "Health and Behavioral Consequences of Binge Drinking in College A National Survey of Students at 140 Campuses." *Journal of the American Medical Association*, 272(21): 1672– 1677.

Weiss, Yoram. 1996. "Synchronization of Work Schedules." *International Economic Review*, 37(1): 157–179.

Williamson, Richard J., Pak Sham, and David Ball. 2003. "Binge Drinking Trends in a UK Community-Based Sample." *Journal of Substance Use*, 8(4): 234–237.

Winston, Gordon C. 1982. *The Timing of Economic Activities.* New York, NY: Cambridge University Press.

Wright, Neil R., and Doug Cameron. 1997. "A Pilot Study of Prospectively Recorded Drinking Patterns among British Men Who Habitually Drink 1-4 Units of Alcohol per Day." *Alcohol and Alcoholism*, 32(6): 777–778.

World Health Organization. 2007. *International Classification of Diseases and Related Health Problems.* Geneva: World Health Organization, 10th revision.

World Health Organization. 2010. *Global Strategy to Reduce the Harmful Use of Alcohol.* Geneva: World Health Organization.

World Health Organization. 2014. *Global Status Report on Alcohol and Health 2014.* Geneva: World Health Organization.

Young, Amy M., Michele Morales, Sean E. McCabe, Carol J. Boyd, and Hannah D'Arcy. 2005. "Drinking Like a Guy: Frequent Binge Drinking Among Undergraduate Women." *Substance Use and Misuse*, 40(2): 241–267.

Zironi, Isabella, Costanza Burattini, Giorgio Aicardi, and Patricia H. Janak. 2006. "Context Is a Trigger for Relapse to Alcohol." *Behavioural Brain Research*, 167(1): 150–155.

Figure 1: Patterns of Binge Drinking and Alcohol Consumption over Days of the Week, by Age Group

Figure 2: Age Patterns of Binge Drinking and Alcohol Consumption, by Weekend and Weekday
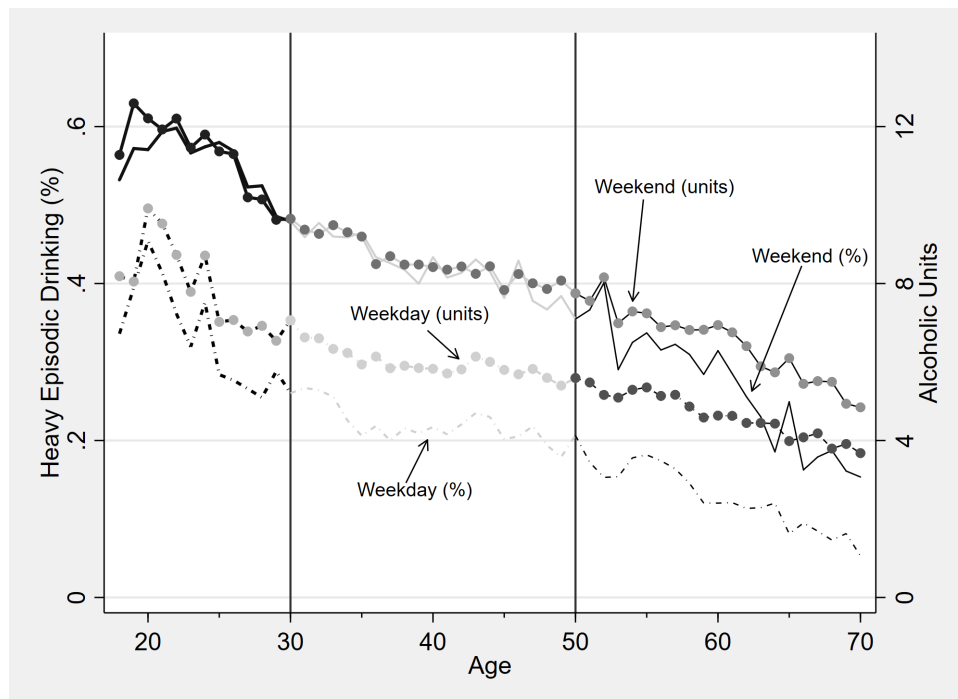
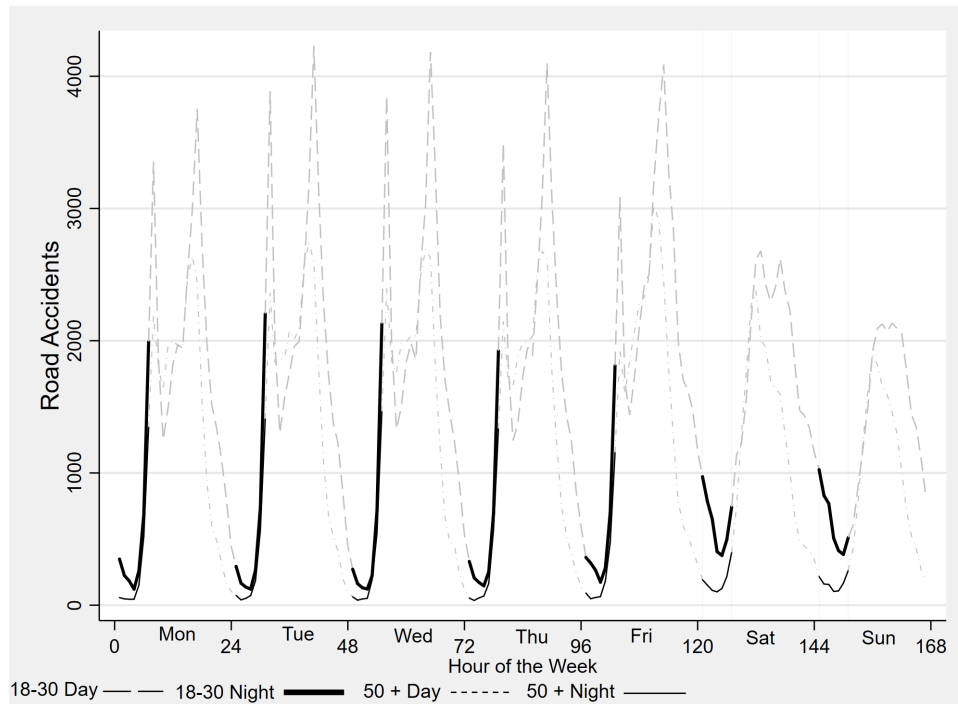Figure 3: Differences in Time Spent in Evenings/Nights at the Weekend between Younger (aged 18-30) and older (aged 50+) Individuals

Figure 4.A: Total Number of Road Accidents, by Hour of the Week

Figure 4.B: Total Number of Injury Related A&E Attendances, by Hour of the Week
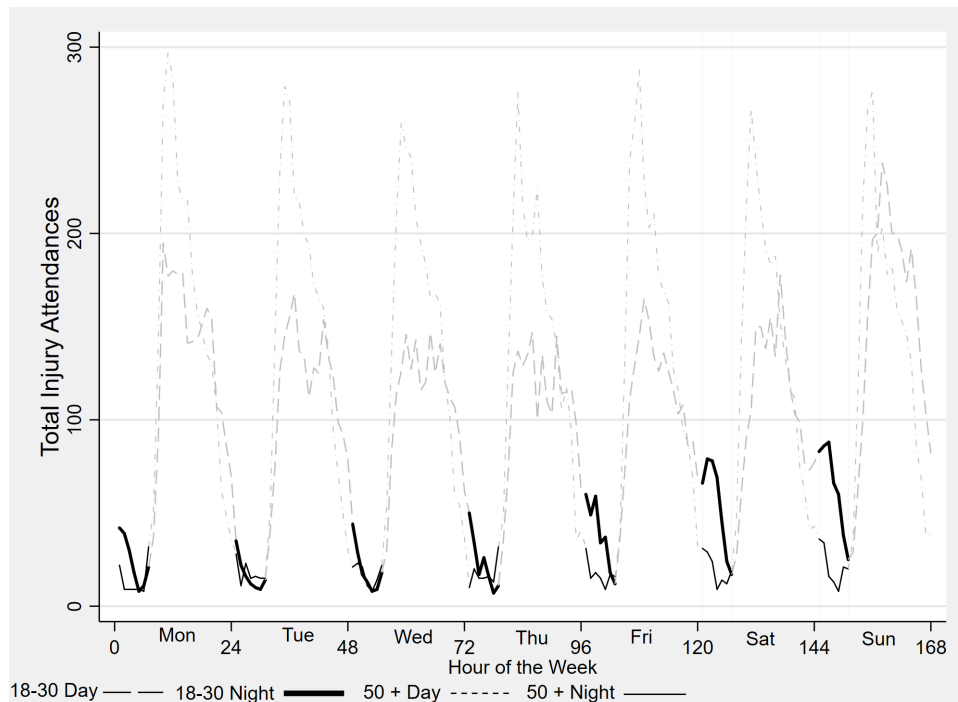
Figure 4.C: Total Number of Arrests per Six Hour Block
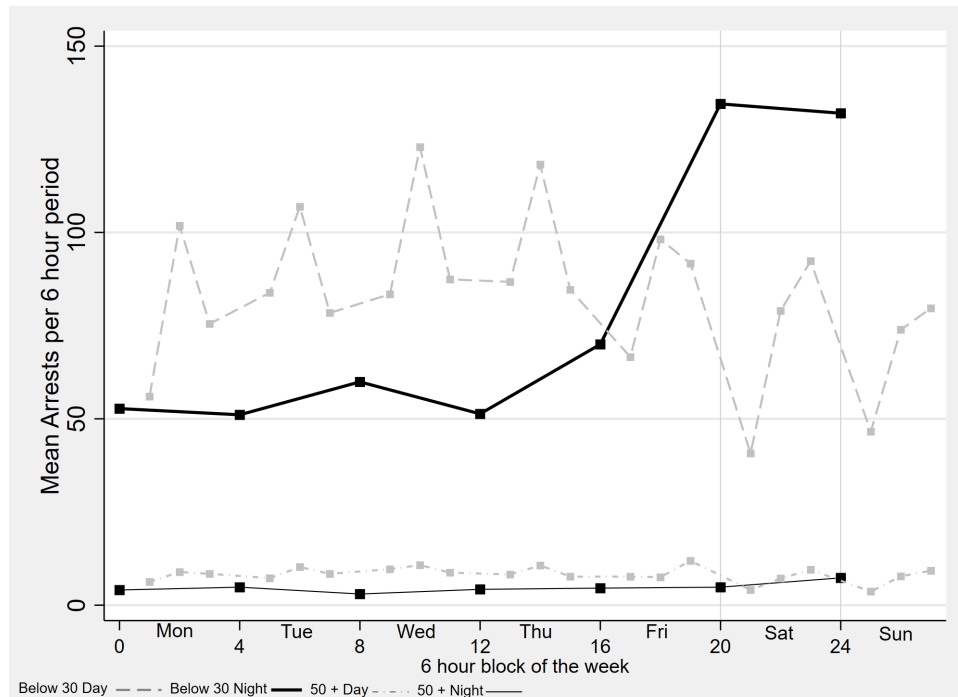
Note: Total numbers are averaged over the sample period. Along the horizontal axis, 0 denotes the first 6-hour block of Monday (00:00–05:59) and 24 denotes the last six-hour of Sunday (18:00–23:59). The vertical lines indicate the benchmark weekend nights.

Figure 5.A: First Stage Estimates by Level of Alcohol Consumption (Road Accidents)
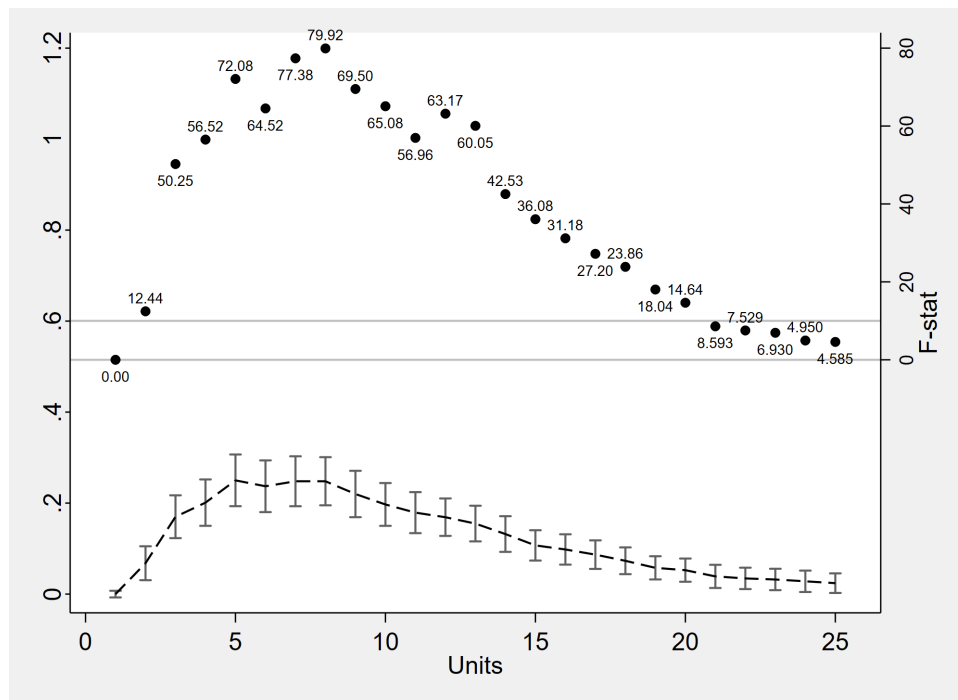
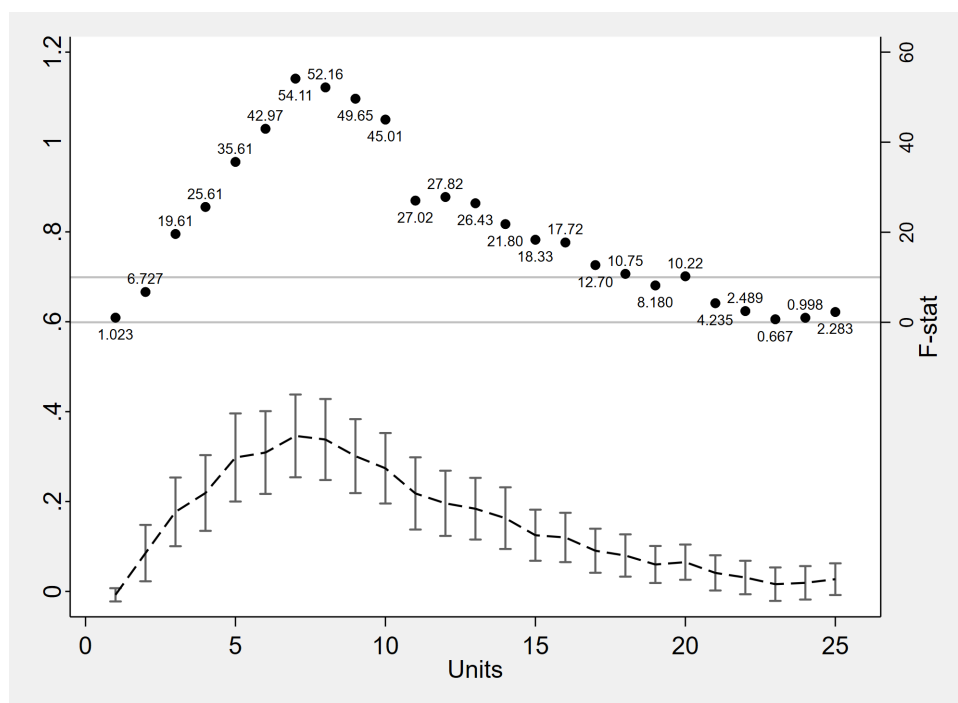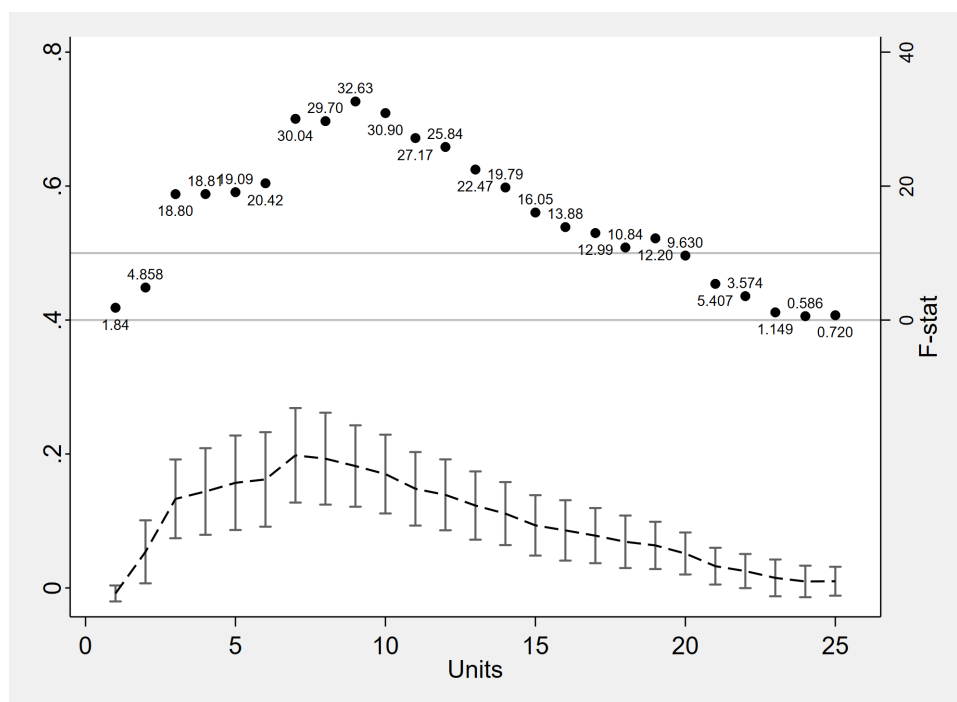Figure 5.B: First Stage Estimates by Level of Alcohol Consumption (A&E Attendances)

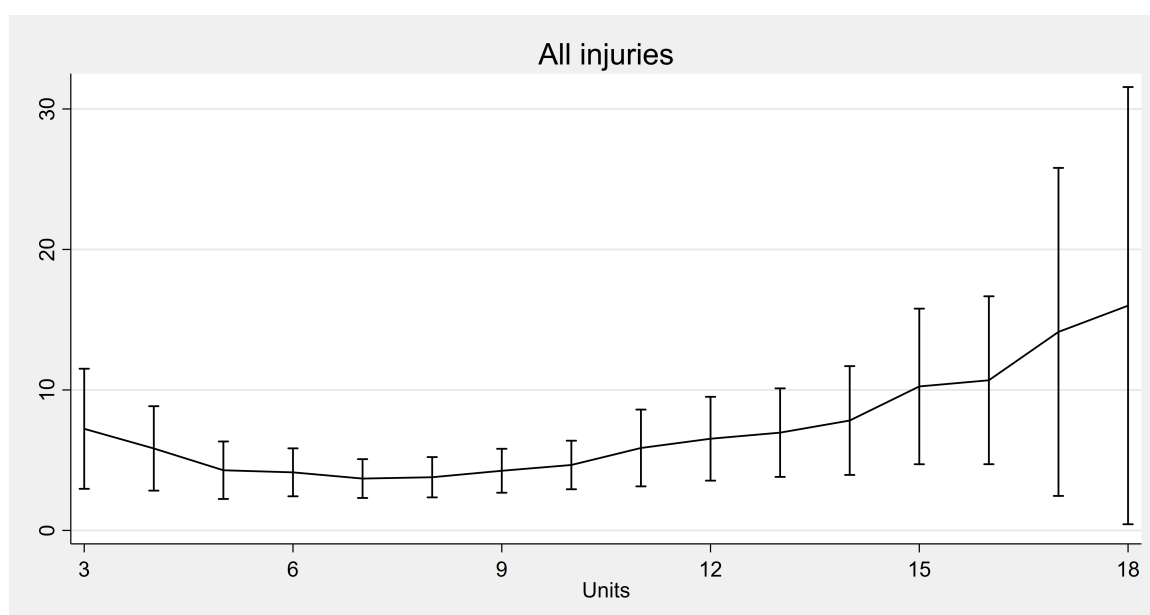Figure 5.C: First Stage Estimates by Level of Alcohol Consumption (Arrests)

Note: In each panel and for each alcoholic unit level, the dependent variable is an indicator for whether the amount of alcohol drunk on the heaviest day in the past week was at or above the specified amount. The estimates of ϕ_1 on $a \times w$ (see equation (5)) are reported at the bottom of each panel, where a is equal to 1 if an individual is between 18 and 30 years of age, and 0 if the individual is aged 50 or more, and w is equal to 1 if an individual drank most in last seven days on a Friday or Saturday, 0 otherwise. For each estimate, the confidence interval is represented by the whiskers. The dots at the top of each panel represent the value of the F-test for the significance of the instrument. Additional controls included in each regression are cell averages for years of education, indicators for gender (=1 if male), race (=1 if white), whether the respondent had a long standing illness (=1 if yes), whether the respondent had ever been a smoker (=1 if yes), the age at which the individual left full time education, the number of licenses per 1,000 individuals, a set of year and quarter dummy variables, and quadratic polynomials in age and day of the week.

Figure 6.A: Effect of Drinking on Road Accidents by Level of Alcohol Consumption
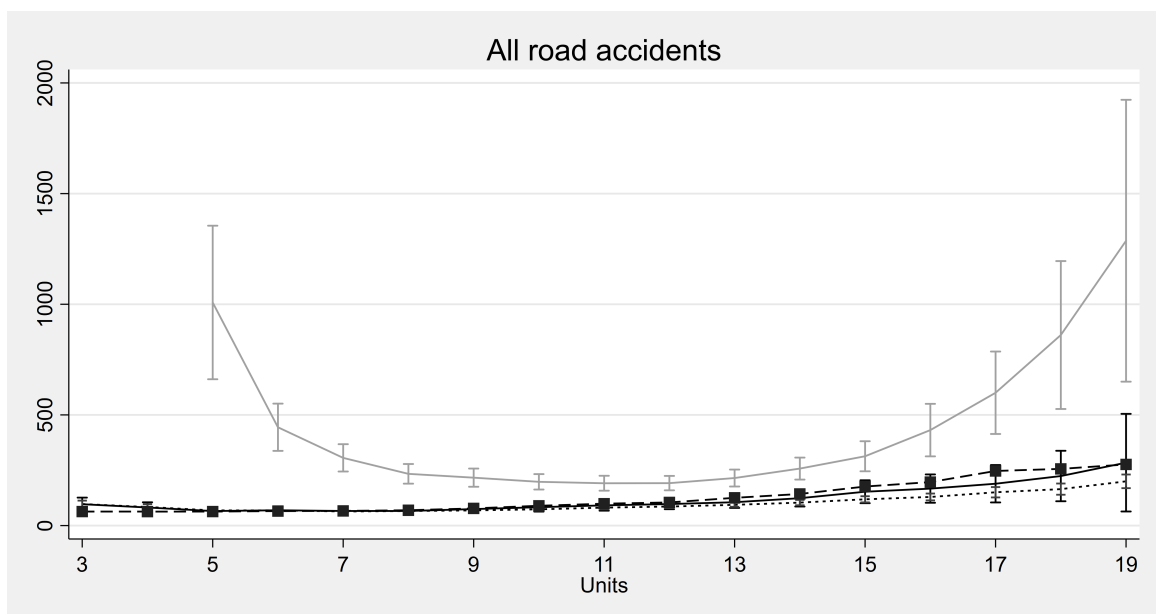
Note: Estimates of β_1 are obtained from two-sample two stage least squares (TS2SLS). At each unit level, the estimates come from a separate regression. Bootstrapped standard errors are estimated using 1,000 replications and are represented by the whiskers. First stage (first sample) estimation uses the estimates obtained from HSE data and reported in Figure 5.A. The solid line and corresponding whiskers denote the TS2SLS estimates obtained using RAD records, the dashed line with the squares denote the TS2SLS estimates obtained with RAD records but using only the controls that are available in the BTS sample (i.e., month, month squared, and sex). The dotted line and corresponding whiskers denote the TS2SLS estimates using BTS data, while the solid gray line and whiskers display the 2SLS estimates from BTS data. All first stage F-statistics are above 10.

Figure 6.B: Effect of Drinking on All Injury Related A&E attendances by Level of Alcohol Consumption

Note: First stage (first sample) estimation uses the estimates obtained from HSE data and reported in Figure 5.B. For all other estimation details see note to Figure 6.A.

Figure 6.C: Effect of Drinking on Alcohol Related Arrests by Level of Alcohol Consumption
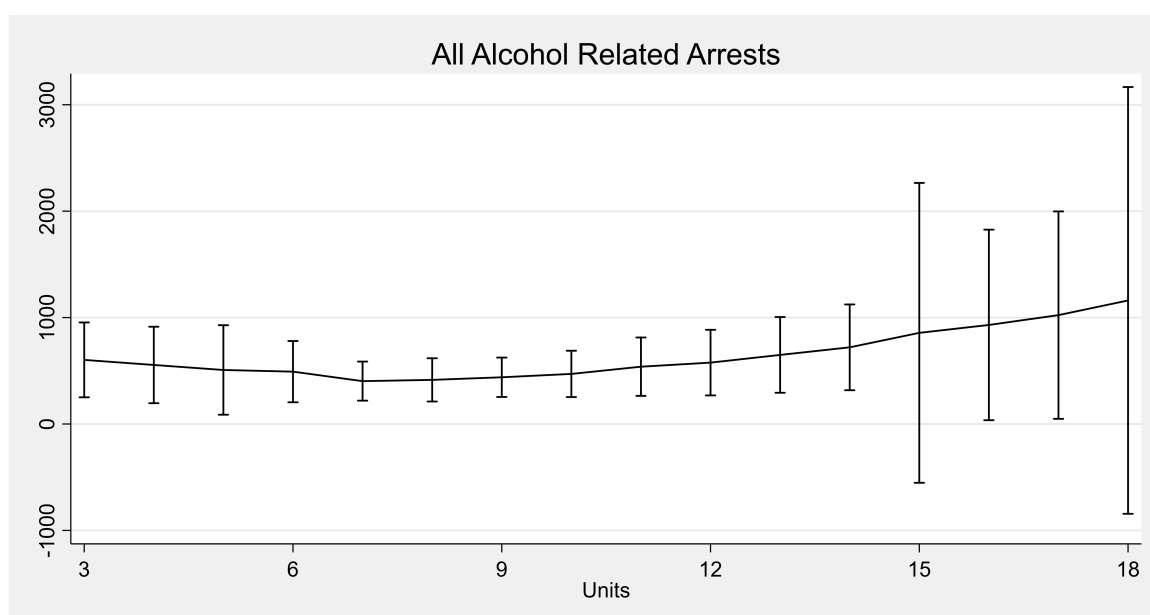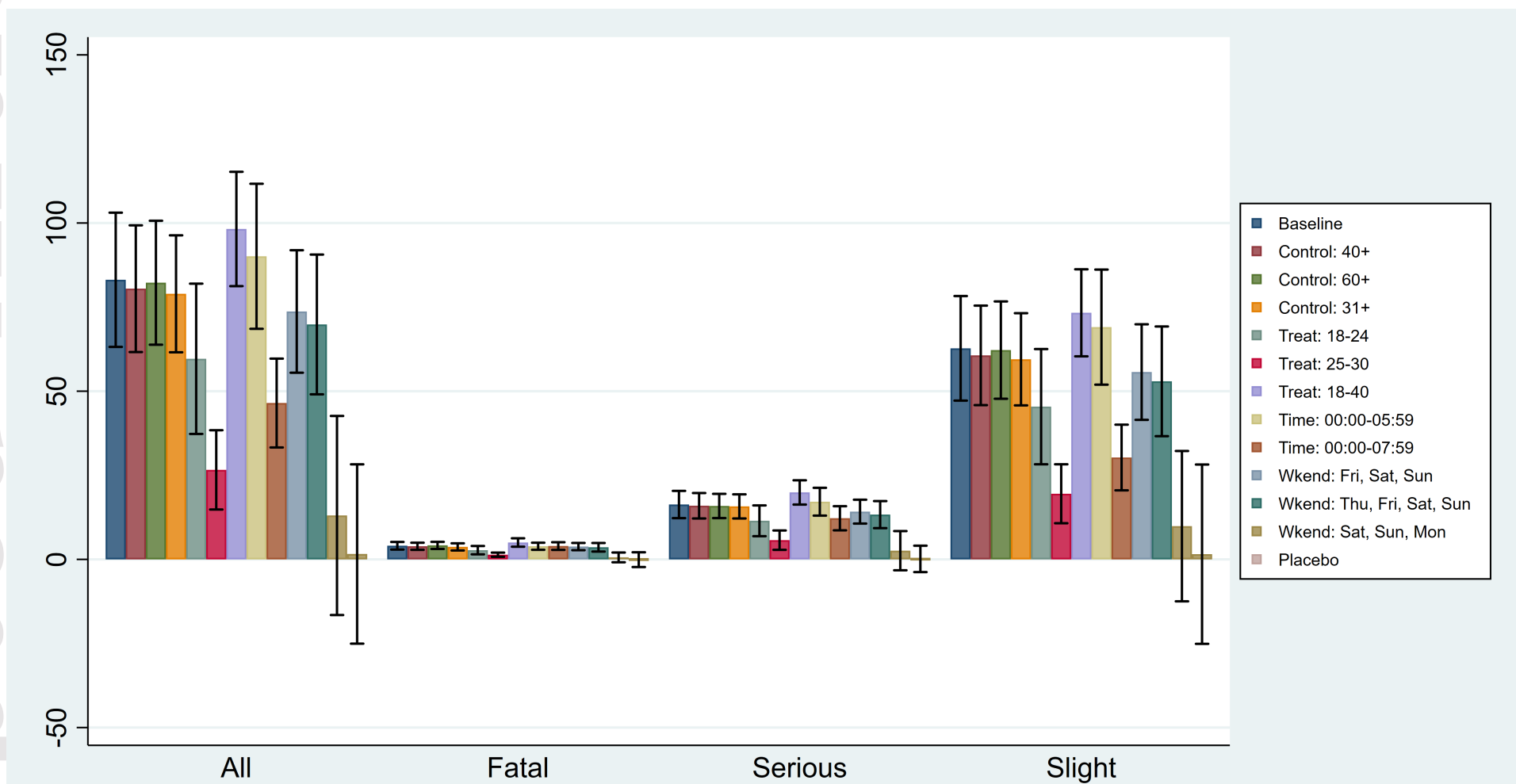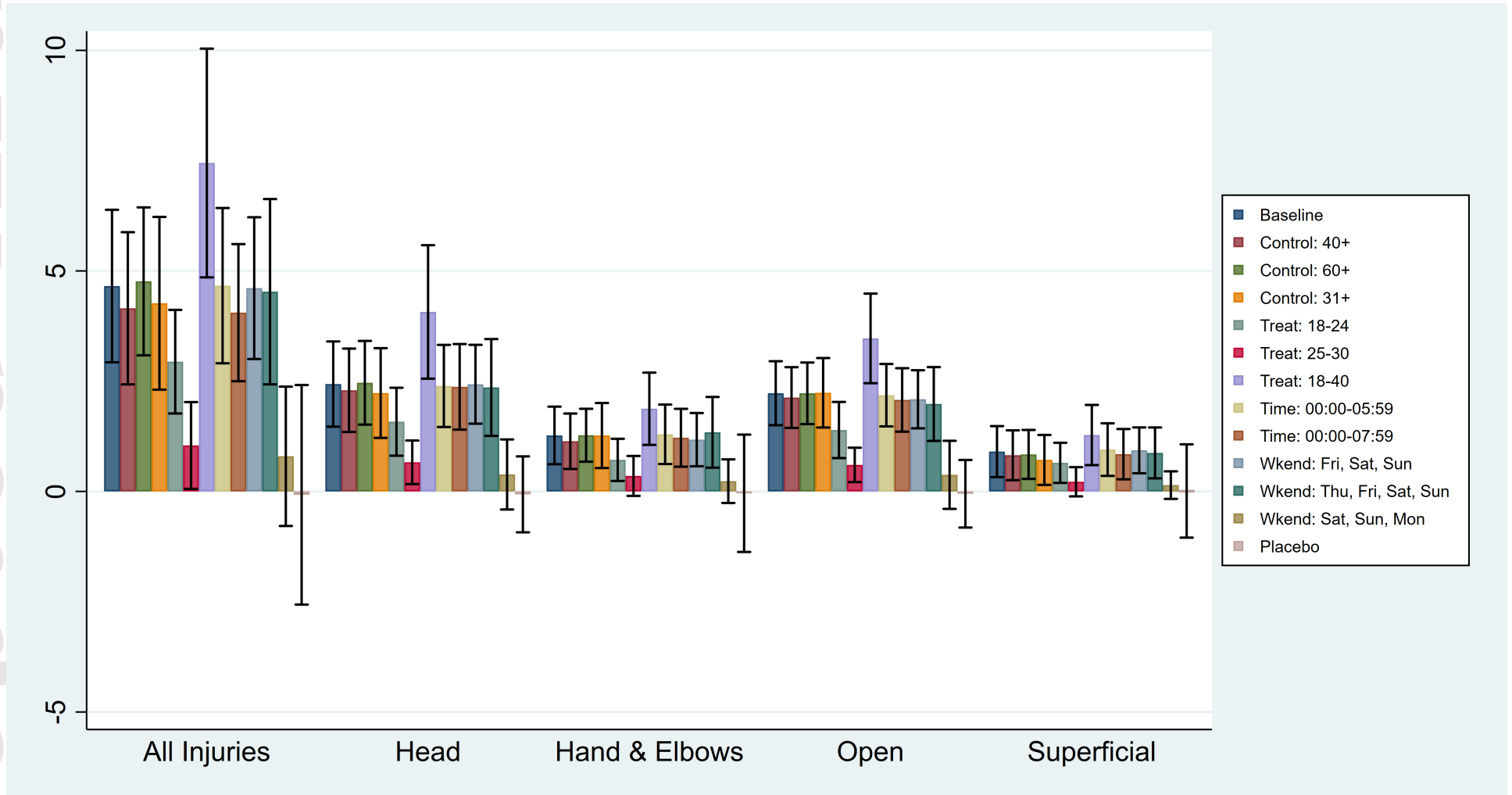

Note: First stage (first sample) estimation uses the estimates obtained from HSE data and reported in Figure 5.C. For all other estimation details see note to Figure 6.A.

Figure 7.A: Sensitivity Analysis (Road Accidents)

Figure 7.B: Sensitivity Analysis (A&E Attendances)

Legend:
- Baseline
- Control: 40+
- Control: 60+
- Control: 31+
- Treat: 18-24
- Treat: 25-30
- Treat: 18-40
- Time: 00:00-05:59
- Time: 00:00-07:59
- Wkend: Fri, Sat, Sun
- Wkend: Thu, Fri, Sat, Sun
- Wkend: Sat, Sun, Mon
- Placebo

Categories: All Injuries, Head, Hand & Elbows, Open, Superficial

Note: Each bar represents the outcome specific TS2SLS estimate of β_1. The whiskers depict the 95% confidence interval. All other details are as specified in the note to Figure 7.A, except that, in order to fit on the figure, the estimates and standard errors of 'Wkend: Sat, Sun, Mon' are divided by 10 and those of 'Placebo' are divided by 100.

Figure 7.C: Sensitivity Analysis (Arrests)
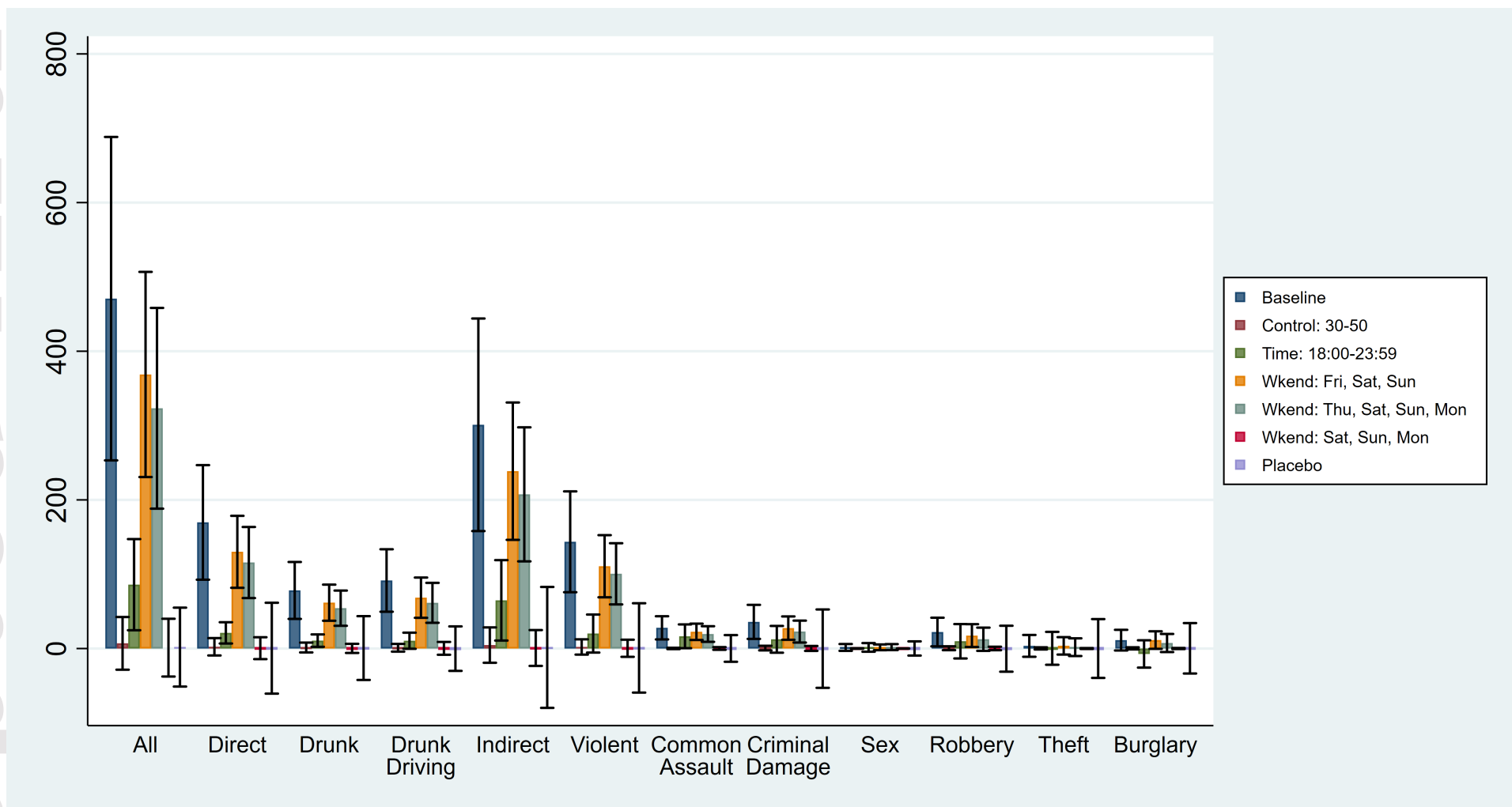
Note: Each bar represents the outcome specific TS2SLS estimate. The whiskers depict the 95% confidence interval. From left to right, the bars are as follows: 'Baseline' (treatment age: 18–30, control age: 50+; treatment time: 00:00–05:59 Saturdays and Sundays, control time: 00:00–05:59 Mondays–Fridays); 'Control: 30–50' changes control group age to 30–50; 'Time: 18:00–23:59' changes treatment time to 18:00–23:59; 'Wkend: Fri, Sat, Sun' changes the weekend definition to include Friday mornings; 'Wkend: Thu, Fri, Sat, Sun' changes the weekend definition to include Thursday and Friday mornings; 'Wkend: Sat, Sun, Mon' changes the weekend definition to include Monday mornings; 'Placebo' shows estimates found when Mondays, Tuesdays, and Wednesdays are defined as weekend (excluding Saturdays and Sundays) and treatment age group are changed to individuals aged 30–50. In order to fit on the figure, the estimates and standard errors of 'Placebo' and 'Control: 30-50' are divided by 100 and those of 'Weekend: Sat, Sun , Mon' are divided by 1000.

Figure 8: Distribution of Alternative Externalties of Binge Drinking Due to Road Accidents
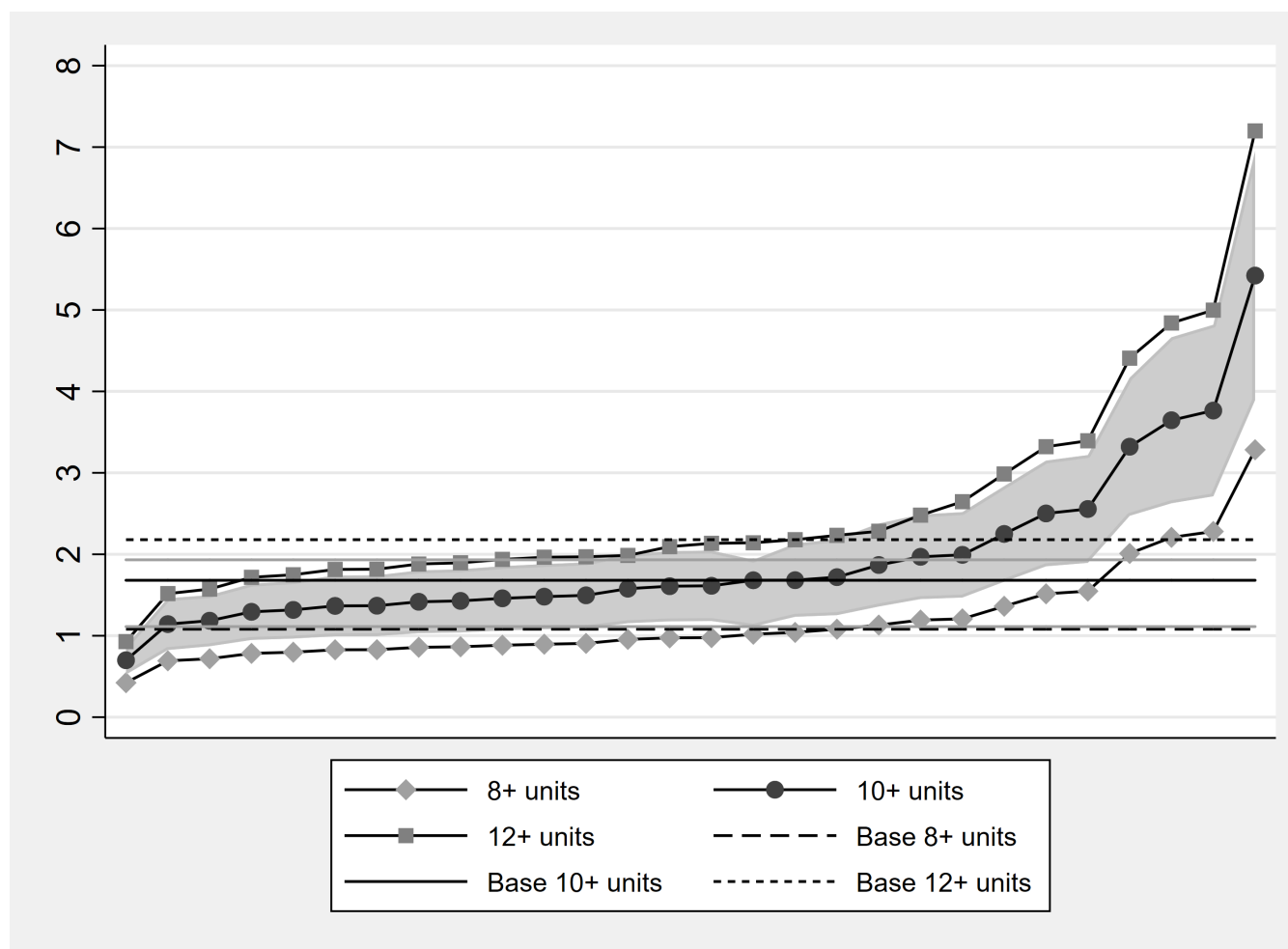
Note: Each dot represents the externality of road accidents associated with binge drinking with the unit cost of one specific component changed at a time. Alternative unit costs are described in the Online Appendix. Each estimate is ranked with the smallest on the left and the largest on the right. The grey area denotes the 95% confidence interval for the 10+ unit case. The three black horizontal lines (continuous and dashed) represent the externality values found with the benchmark estimates reported in Table 2. The grey horizontal lines are the 95% confidence interval around the benchmark estimate for 10+ units (continuous line).

Table 1: Reduced Form Effects of Binge Drinking

A. Road Accidents					**B. A&E Attendances**				
All (RAD)	Fatal	Serious	Slight	All (BTS)	All injuries	Head	Hand & Elbows	Open Wounds	Superficial wounds

	All (RAD)	Fatal	Serious	Slight	All (BTS)	All injuries	Head	Hand & Elbows	Open Wounds	Superficial wounds
π_1	16.39	0.800	3.217	12.38	22.65	1.278	0.668	0.348	0.611	0.248
	(0.737)	(0.0750)	(0.223)	(0.604)	(1.086)	(0.138)	(0.0741)	(0.0732)	(0.0661)	(0.0658)
R^2	0.977	0.816	0.938	0.977	0.978	0.831	0.779	0.636	0.749	0.549

C. Arrests

	Direct					**Indirect**							
	All	Direct	Drunk	Driving		Indirect	Violent	Common assault	Criminal damage	Sexual	Robbery	Theft	Burglary
π_1	79.94	28.82	13.27	15.54		51.13	24.39	4.727	6.089	0.2478	3.763	0.606	1.912
	(6.639)	(2.532)	(1.626)	(1.488)		(4.668)	(1.933)	(0.946)	(1.427)	(0.386)	(1.381)	(1.194)	(1.052)
R^2	0.960	0.915	0.843	0.883		0.958	0.930	0.807	0.801	0.402	0.686	0.794	0.750

Note: The estimates of π_1 are obtained from OLS regressions (see equation (6)). Robust standard errors are in parentheses. The data are aggregated into cell means by year, quarter of the year, day of the week, and age group. With these aggregations the number of observations are 168 (=7 days × 3 years × 2 age groups × 4 quarters = 168. All regressions are weighted by cell size. For the BTS data we have 144 cells (2 age groups × 3 years × 12 months × 2 day types). The coefficient π_1 is on $a \times w$, where a equals 1 if an individual is between 18 and 30 years of age, and 0 if the individual is aged 50 or more, and w equals 1 if an individual drank most in last seven days on a Friday or Saturday, and 0 otherwise. Additional controls that are not reported are indicators for gender (=1 if male), ethnicity (=1 if white), whether the respondent had a long standing illness (=1 if yes), whether the respondent had ever been a smoker (=1 if yes), the age at which the individual left full time education, years of education, the number of licenses per 1,000 individuals, quadratic polynomials in age and day of the week, and a set of year and quarter dummy variables.

Table 2: Effects of Binge Drinking on Road Accidents

	Mean (All Day)	Mean (Night)	8+	10+	12+	Average Effect	Marginal Effect	AET Ratio & Union of C.I.
			Alcoholic Units					
A. RAD								
All	445.8	33.2	66.05 (7.276)	83.12 (10.18)	97.27 (11.96)	0.186	0.038	<0 [31.73, 101.2]
Fatal	5.6	1	3.222 (0.419)	4.054 (0.579)	4.745 (0.704)	0.725	0.149	<0 [1.269, 5.095]
Serious	59.8	6.1	12.96 (1.553)	16.31 (2.064)	19.09 (2.584)	0.273	0.056	<0 [5.825, 19.99]
Slight	380.4	26.1	49.86 (5.650)	62.75 (7.930)	73.43 (9.124)	0.165	0.034	<0 [23.38, 76.89]
B. BTS								
All	96.7	41.1	69.19 (8.724)	89.48 (11.90)	104.6 (14.28)	0.925	0.210	<0 [31.10, 109.7]
C. BTS-RAD								
All, BTS (2SLS)	96.7	41.1	233.8 (22.63)	197.7 (17.81)	191.8 (16.62)	2.044	-0.373	<0 [79.67, 232.6]
All, RAD (TS2SLS)[a]	445.8	33.2	66.01 (3.558)	73.67 (3.669)	86.40 (4.083)	0.165	0.017	<0 [32.22, 80.48]

Note: Estimates obtained from two-sample two stage least squares (TS2SLS). Each coefficient represents a separate estimation. Bootstrapped standard errors obtained using 1,000 replications are in parenthesis. First stage (first sample) estimation uses the estimates in Figure 5.A for the estimates on the RAD records. First stage (first sample) F-test values for the BTS sample are 75.5, 69.7, and 59.0 for 8+, 10+, and 12+ units, respectively. The estimation on the BTS data is performed on 144 cells (=2 age groups × 3 years × 12 months × 2 day types). 'Average Effect' is given by the estimate of β_1 evaluated at 10+ units divided by the (all day) mean of the corresponding outcome variable. 'Marginal Effect' is calculated by the difference between β evaluated at 10+ units and at 8+ units divided by the (all day) mean of the corresponding outcome variable. 'AET ratio' is given by $\frac{\beta_1}{\beta_1^R - \beta_1}$ where β_1 is the estimate with the full set of controls (as presented in this table) and β_1^R is the estimate with a restricted set of controls (only including quadratic polynomials in age and day of the week). 'AET ratio' in panel B and the lower part of panel C are computed with no controls. 'Union of C.I.' in square brackets represents the 95% confidence interval based on the union of prior-weighted confidence intervals of the direct effect of the instrument on the outcome. The figures show the C.I. with a lower bound of the direct effect of zero and an upper bound of 50% of the reduced form effect.

[a] These estimates are obtained using the same controls (month and gender) as in the BTS (2SLS) sample.

Table 3: Effects of Binge Drinking on A&E Attendances

	Mean (All Day)	Mean (Night)	Alcoholic Units			Average Effect	Marginal Effect	AET Ratio & Union of C.I.
			8+	10+	12+			
All Injuries	70.937	3.8	3.786	4.657	6.527	0.066	0.012	17.77
			(0.732)	(0.882)	(1.522)			[1.280, 6.126]
Head	11.158	0.997	1.979	2.435	3.412	0.218	0.041	18.59
			(0.380)	(0.493)	(0.876)			[0.602, 3.291]
Hand and Elbows	14.059	0.492	1.032	1.270	1.779	0.090	0.170	40.97
			(0.289)	(0.333)	(0.519)			[0.144, 1.816]
Open	10.266	0.687	1.810	2.227	3.120	0.217	0.041	71.84
			(0.306)	(0.370)	(0.662)			[0.668, 2.848]
Superficial	6.894	0.33	0.734	0.903	1.266	0.131	0.025	<0
			(0.239)	(0.295)	(0.469)			[-0.029, 1.445]

Note: Estimates obtained from two-sample two stage least squares (TS2SLS). Each coefficient represents a separate estimation. Bootstrapped standard errors obtained using 1,000 replications are in parenthesis. First stage (first sample) estimation uses the estimates reported in Figure 5.B. For other estimation details see the note to Table 2.

Table 4: Effects of Binge Drinking on Arrests

	Mean (All Day)	Mean (Night)	Alcoholic Units 8+	10+	12+	Average Effect	Marginal Effect	AET Ratio & Union of C.I.
All	659.8	139.6	414.6 (103.7)	470.7 (111.0)	576.8 (157.3)	0.713	0.085	<0 [129.5, 642.6]
Direct	68.4	36.1	149.4 (37.70)	169.7 (39.32)	207.9 (55.98)	2.481	0.297	<0 [46.15, 231.6]
Drunk	23.9	10.4	68.82 (18.64)	78.14 (19.55)	95.75 (27.76)	3.269	0.390	<0 [17.68, 108.8]
Drunk Driving	44.6	25.7	80.62 (20.47)	91.53 (21.44)	112.2 (30.04)	2.052	0.245	<0 [23.55, 126.0]
Indirect	591.4	103.4	265.1 (67.15)	301.0 (72.96)	368.9 (102.8)	0.509	0.061	<0 [79.01, 413.7]
Violent	151.7	34.2	126.5 (33.16)	143.6 (34.61)	176.0 (49.88)	0.947	0.113	<0 [39.11, 197.0]
Common assault	83.3	13.8	24.51 (7.800)	27.83 (7.957)	34.10 (11.28)	0.334	0.040	17.07 [3.182, 40.74]
Criminal Damage	54	12.5	31.58 (10.46)	35.85 (11.68)	43.94 (15.82)	0.664	0.079	41.69 [2.171, 54.89]
Sex	23.8	3.5	1.284 (2.215)	1.458 (2.355)	1.787 (2.987)	0.061	0.070	1.373 [-3.465, 5.641]
Robbery	45.4	6.3	19.52 (8.463)	22.16 (9.822)	27.15 (12.95)	0.488	0.058	<0 [-4.795, 39.56]
Theft	145	13.2	3.143 (6.961)	3.569 (7.484)	4.373 (9.066)	0.025	0.003	2.494 [-11.26, 16.63]
Burglary	52.5	10.2	9.914 (6.788)	11.26 (7.092)	13.79 (9.199)	0.214	0.026	<0 [-6.187, 23.66]

Note: Estimates obtained from two-sample two stage least squares (TS2SLS). Each coefficient represents a separate estimation. Bootstrapped standard errors obtained using 1,000 replications are in parenthesis. First stage (first sample) estimation uses the estimates reported in Figure 5.C. For other estimation details see the note to Table 2.

Table 5: Effects of Binge Drinking on Road Accidents by number of Vehicles and Casualties

	Vehicles		Casualties	
	1	2+	1	2+
All	51.94	30.96	53.38	29.36
	(6.285)	(4.093)	(6.668)	(3.594)
Fatal	2.527	1.485	1.495	2.606
	(0.376)	(0.357)	(0.302)	(0.401)
Serious	11.82	4.465	9.407	6.979
	(1.408)	(0.931)	(1.317)	(1.113)
Slight	37.60	25.01	42.48	19.78
	(4.885)	(3.234)	(5.501)	(2.446)
First stage F-test	65.08	65.08	65.08	65.08

Note: Estimates obtained from two-sample two stage least squares (TS2SLS). Each coefficient represents a separate estimation and uses 10+ units as a definition of a binge. Bootstrapped standard errors obtained using 1,000 replications are in parenthesis. First stage (first sample) estimation uses the estimates reported in Figure 5.A. For other estimation details see the note to Table 2.

Table 6: Externality of Binge Drinking Due to Road Accidents

	$\Delta(5,8)$	$\Delta(5,10)$	$\Delta(5,12)$
All	0.92	1.52	2.02
	[0.72, 1.12]	[1.11, 1.93]	[1.44, 2.6]
Fatal	0.56	0.92	1.23
	[0.43, 0.69]	[0.66, 1.19]	[0.85, 1.6]
Serious	0.26	0.43	0.57
	[0.2, 0.31]	[0.32, 0.53]	[0.41, 0.72]
Slight	0.10	0.17	0.23
	[0.09, 0.12]	[0.13, 0.21]	[0.17, 0.28]

Note: Figures are expressed in billion pounds sterling (2014 prices). The computation of each estimate is described in the text. The 95% confidence intervals are presented in square brackets.