CPB Discussion Paper

No

May 6, 2009

The impact of schooling on problem drinking

Evidence from Australian Twins

Dinand Webbink and Nikolay Zubanov

CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14
P.O. Box 80510
2508 GM The Hague, the Netherlands

Telephone +31 70 338 33 80
Telefax +31 70 338 33 50
Internet www.cpb.nl

ISBN

Abstract

This paper uses three waves of a repeated survey of genetically identical twins from Australia to identify a causal impact of schooling on regular and problem drinking. The identification comes from controlling for identical twin fixed effects, whereby unobserved family and genetic factors affecting both schooling and drinking are isolated. Our main finding is that, while schooling and regular drinking are unrelated, more schooling leads to less problem drinking for men. Thus, extra schooling helps combat such aspects of problem drinking as alcohol dependence and health problems. No effect of schooling was found for women.

This version is preliminary and incomplete. Please do not quote.

1 Introduction

Alcohol abuse is among the top causes of deaths and poor health in the world, especially for men and high-income countries (see Lopez et al., 2006). In the European Union, alcohol disorders are responsible for 7.4% of all cases of disability and premature death, cause some 60,000 underweight births and 10,000 traffic accident deaths per year, and cost 125 bn Euros (1.3% EU's GDP, 2003 estimate) of direct damage and lost productivity potential (Anderson and Baumberg, 2006). Not only does it have direct negative consequences on health, but also it affects individual socioeconomic status, emotional well-being, and safety (WHO, 2007, p. 19-23). These significant negative externalities make a case for policy interventions to reduce problem drinking. Our study evaluates the effects of one such policy – promoting public education.

The role of schooling in reducing problem drinking is a topic of a lively academic debate. The results from comprehensive literature surveys in Cutler and Lleras-Muney (2006) and Furnée et al. (2008) suggest that better educated people lead healthier lifestyles. Yet, when it comes to the link between schooling and problem drinking, studies which report a reduction in problem drinking with extra schooling (e.g., Häkkinen et al., 2006; Cutler and Glaeser, 2005) are mixed with those showing no result (e.g., Park and Kang, 2008).

One of the key topics in the research on education and health outcomes is causality: whether more education induces healthier lifestyle, or healthier people can get more education, or there are third factors – often unobserved – which co-determine education and health, such as genes, family background and early upbringing. Having to cope with such unobservables has motivated two identification strategies: (i) using instrumental variables to identify the variation in schooling which is independent of the unobservables, and (ii) using data designs which are arguably free from confounding effects of unobservables. Such designs may be obtained from adoptees, whose genes are independent of those of their adopted parents, or from twins who share the same family background and (some) genes, so that these factors can be controlled away.

We use Australian Twin Register data from three waves of the survey of the same group of Australian twins: "Canberra", "Alcohol-1" and SSAGA, collected between years 1980 and 1993. These data have been used by psychologists and geneticists (the total number of studies on these data exceeds 600), and also by economists (e.g., Miller et al. (1995) on returns to schooling, Webbink et al. (2008b) on the effects of teenage pregnancy on health outcomes and Webbink et al. (2008a) on schooling and obesity). The first two surveys contain a large number of questions covering, among schooling and drinking patterns, biometric data, health, family, employment, personality and socio-economic background. Alcohol-1 also includes questions on problems with individual's health, personality and conduct perceived to be caused by drinking. The third survey, SSAGA, contains some additional questions related to problem drinking.

In addition to the possibility to control for family and genetic unobservables, our data offer further advantages. First, we can dinstinguish between monozygotic (identical) and dyzygotic (fraternal) twins. The former share family characteristics and half of the genes. The latter share much of the genotype and family charactersitics. So, monozygotic twins enable ultimate control over many unobservables. Second, our data are longitudinal in a sense that every individual in our sample was surveyed twice, in 1980 and 1988. This allows control over individual history of drinking which, as we will show, is an important determinant of problem drinking. Third, the data contain both own and twin reports on schooling and some aspects of drinking, which we use to control for measurement error in schooling (as in Ashenfelter and Krueger, 1994).

In sum, we find that, for men, schooling improves the ability to enjoy drinking alcohol responsibly, without adverse consequences for self and others. Although schooling does not seem to significantly affect alcohol intake, it does lower the risk of developing behavioural and health problems perceived to be caused by drinking, as well as addiction to drink. The effect of schooling on problem drinking by women is insignificant.

2 Previous studies

Better educated people tend to lead healthier lifestyles (Cutler and Lleras-Muney, 2006; Furnée et al., 2008). This association, labelled as “health gradient”, is also found in the literature on schooling and alcohol consumption. Some studies report significant and negative associations between the amount of alcohol consumed and years of schooling (Häkkinen et al., 2006; Sharpe et al., 2001; Baltagi and Griffin, 1995). This bird’s eye view ignores the distinction between the effects of schooling on “regular” and “problematic” drinking. In fact, more schooling may increase the probability of initiation into drinking (Yen, 1995; Van Oers et al., 1999) or even the total intake (for instance, Su and Yen (2000) found that in the U.S. more educated people drink less beer but more wine per week).

Focussing on problem drinking, significant negative correlations between its various symptoms and schooling are often reported. Thus, Kenkel (1991) using data from U.S. 1985 National Health Interview Survey found that schooling increases the total fortnightly alcohol intake but reduces heavy drinking defined as the number of days last year when an individual had five or more drinks. Cutler and Glaeser (2005) also found a negative correlation between education level and the probability of being a heavy drinker (three or more drinks per day) using the 1990 data from the same survey. Examples from a wider social science literature include studies by Van Oers et al. (1999) of excessive drinking and alcohol-related problems among Dutch adults and Droomers et al. (2004) on excessive drinking, also among the Dutch. The conclusion from both of these studies is that schooling is associated with significantly lower risk of excessive drinking and health problems for men and psychological dependence on drink for both sexes.

Correlations, though informative, are not enough to establish a causal impact of schooling on problem drinking because they may be susceptible to unobservables, such as genes or family background. One approach to establishing causality is to use instrumental variables to identify the variation in schooling which is independent on the unobservables. Thus, Park and Kang (2008) used availability of schools in the region of birth and the culturally-induced parental preference towards the first son as instruments for individual schooling attainment. They found some evidence that schooling induces healthy lifestyle, but none that it reduces heavy drinking (defined as drinking more than five times a month).

An alternative to instrumental variables is twins data which design allows to control for unobservables through twin pair fixed effects. Twins data have been used in economics for estimating the effects of schooling on earnings (Miller et al., 1995; Bonjour et al., 2003; Behrman and Rosenzweig, 1999; Isacsson, 1999) and some health outcomes (Webbink et al., 2008a), but not as yet on drinking. Outside economics, twins data studies on genetic and environmental determinants of health outcomes included schooling as one of the controls. It was found significantly to affect the development of clinically diagnosed alcohol dependence (Heath et al., 1997; Knopik et al., 2004). These studies also conclude that genes are responsible for about half of the variation in alcohol dependence, thus leaving another half to environmental factors, such as schooling, and the interactions between ‘nature’ and ‘nurture’.

Not all education is important for reducing problem drinking. Perhaps surprisingly, many special educational interventions to promote awareness of the harms of drinking were often found to be ineffective. Foxcroft et al. (2003) systematically surveyed fifty-six studies on education interventions designed for young (< 25) people, forty-one of which had a randomised controlled trials design allowing to identify causal effect of the intervention. They found that only a few had brought a significant reduction in problem drinking.

3 Data and measures

Our data come from the Australian Twin Register and include three waves of the survey of the same group of Australian twins: Canberra (collected in 1980-82), Alcohol-1 (collected in 1988-89) and Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) (collected in 1993). The first survey, Canberra, covers 3808 complete twin pairs aged over 18, of whom 2934 complete pairs participated in the second survey, Alcohol-1, and 2719 pairs participated in the third. Canberra and Alcohol-1 contain a large number of questions covering, among schooling and drinking patterns, biometric data, health, family, employment, personality and socio-economic background. Alcohol-1 also includes questions on problems with individual's health, personality and conduct perceived to be caused by drinking. SSAGA is the key source of data to produce the alcohol dependence indicator (e.g., Knopik et al. (2004)). Our key dataset, however, is Alcohol-1 because it contains most of the questions of our interest. It is well-suited for our inquiry because the minimum respondent age in Alcohol-1 is 25, so that most of the subjects will have completed their schooling and had some history of drinking alcohol by the time of the survey. Unless otherwise indicated, all of our descriptive statistics and regression results are based on Alcohol-1

In the following analysis we focus on active drinkers, thus excluding tea-totallers (296) and those who had not had any alcohol within at least a year before the survey (506). Tea-totallers are excluded because of the focus of our study on problem drinking, since they cannot by definition be problem drinkers. Including them with a zero alcohol intake would confuse the link between schooling and drinking because the role of schooling in choosing whether to drink or not and how much to drink may be different. As for people with a zero alcohol intake within a year, we assume that that they stopped drinking alcohol. Again, the effects of schooling on deciding to stop drinking altogether or drink very little may differ, and we do not want to confuse the two[1]. We describe the key variables for our research in the following subsections.

3.1 Schooling

In the original datasets, schooling is measured in levels (less than 7 years at school; 8-10 years; 10-12 years or apprenticeship or trade certificate; technical or teacher's college; university; post-graduate degree), which we convert into years, as is the standard procedure in studies on returns to education. The resulting variable ranges from 5 to 17 years. In addition to the continuous years of schooling variable, we use high school graduation dummy to capture potentially nonlinear effects of passing this important stage in one's education.

[1] Also, we use alcohol intake as one of the control variables for alcohol-related problems. It may be that an individual decided to stop drinking after having experienced an alcohol problem. So, non-active drinkers are excluded to prevent such reverse causality.

Most men and women attain 11.5 years of schooling which corresponds to a graduation from a comprehensive high school or its more vocational equivalent (apprenticeship, trade certificate etc.). On average, however, men have more schooling than women: 12.5 years versus 11.3. Women are twice more likely than men to have stopped at 9 years; and only half as likely to have progressed to a university or post-graduate degree.

The correlation between the twins' years of schooling is high: 0.81 for monozygotic (identical) and 0.59 for dyzygotic (fraternal) twins, similar to the estimates of the same reported in Behrman and Taubman (1989), 0.75 and 0.55. The higher correlation for monozygotic twins suggests an important role of genetic endowments in determining educational attainment. A simple back-of-the-envelope calculation suggests that genes account for 62% in the total variation in schooling [2]. A more elaborate model (Miller et al., 2001) puts this share at 50% or more. So, genes appear to play an important part in schooling attainment, and therefore we need to control for them by means of monozygotic twin fixed effects.

Average schooling differs between the Canberra and Alcohol-1 sample. In the Canberra sample, the average schooling is 12 years for men 11 years for women. Most of this difference is presumably due to people graduating from a college or university later in life. At the same time, 14.5% of respondents report *less* own schooling in 1988 than in 1980, indicating considerable measurement error in our schooling data. We take up the measurement error issue in sub-section 4.1.

3.2 Regular and excessive drinking

The most direct measure of regular drinking is the number of drinks taken within a given period of time, say a week. The Canberra and Alcohol-1 samples contain detailed data on the number of drinks of five different sorts – beers, wines, spirits, sherry, and other – consumed on each day of the last week. The respondents were instructed what a "drink" means, so we calculate the total, which we adjust for seasonality in the following regression analysis. This is our preferred measure of regular alcohol intake, but we consider alternatives in sub-section **??**.

Because regular drinking does not necessarily imply problem drinking, we develop three additional measures of alcohol intake which are meant to pick up excessive drinking: the maximum number of drinks had in a day of an average week, the maximum number of drinks ever had in a single day, and the standard deviation of daily alcohol intake over a week.

[2] Suppose the total variance in schooling (σ^2) is the sum of three independent components: genetic (σ_g^2), family (σ_f^2) and environmental (σ_e^2). For monozygotic twins, who share the genetic and family components, the share of the common component in the total variance in schooling is $\frac{\sigma_g^2+\sigma_f^2}{\sigma^2}$, which is equal to the squared correlation coefficient between years of schooling of monozygotic twins, $0.81^2 = 0.66$. For dizygotic twins, who share the family component and half the genes, the commont component accounts for $\frac{0.5\sigma_g^2+\sigma_f^2}{\sigma^2} = 0.59^2 = 0.35$. So, the fraction in the total variance in schooling due to genes is $\frac{\sigma_g^2}{\sigma^2} = 2 \times (0.66 - 0.35) = 0.62$.

Intuitively, having 7 drinks on Friday and Saturday constitutes a different, less healthy, type of drinking than consuming 2 drinks per day every day.

Table 3.1 Average alcohol consumption by years of schooling

MEN years of schooling	drinks last week mean	drinks last week obs.	max. drinks p/day, average week mean	max. drinks p/day, average week obs.	max. drinks p/day ever in lifetime mean	max. drinks p/day ever in lifetime obs.	std. deviation of alcohol intake mean	std. deviation of alcohol intake obs.
5	13.93	28	4.19	27	5.70	30	1.07	28
9	12.32	309	5.19	308	10.35	314	1.44	309
11.5	13.03	768	5.90	770	13.11	794	1.80	768
13	10.21	262	4.80	262	10.98	266	1.48	262
15	11.56	382	4.77	380	12.10	392	1.53	382
17	9.71	236	3.99	236	9.50	246	1.30	236
Total	11.88	1985	5.18	1983	11.67	2042	1.58	1985
Median	7		4		10			
WOMEN								
5	3.68	53	1.87	53	3.57	56	0.46	53
9	4.94	1135	2.53	1128	4.49	1200	0.64	1135
11.5	5.63	1297	2.91	1292	5.95	1367	0.82	1297
13	5.76	500	3.06	494	6.07	523	0.91	500
15	5.16	314	2.82	314	6.57	331	0.83	314
17	6.32	175	2.77	175	5.74	188	0.85	175
Total	5.39	3474	2.77	3456	5.50	3665	0.77	3474
Median	3		2		5			

Table 3.1 reports the means and medians of the alcohol consumption variables by years of schooling, separately for men and women. The distribution of weekly alcohol intake is skewed to the right by the presence of heavy drinkers. For instance, 10% of men (women) had last week in excess of 26 (13) drinks. These individuals put the average number of drinks last week up to 11.9 for men and 5.4 for women, while the medians are 7 and 3, respectively.

Weekly alcohol consumption by both sexes is uneven: as we can see from the maximum daily intake statistics, about half of the average weekly intake is consumed on a single day. Excessive drinking is noticeable: thus, the average maximum daily intake over a lifetime is comparable with the average intake for the whole week. There appears to be a weakly negative link between alcohol consumption and schooling for men and, curiously, positive for women.

3.3 Alcohol-related problems

Although alcohol intake, regular or excessive, is perhaps the most objective measure of individual exposure to potentially harmful effects of alcohol, this measure alone fails to account for all the consequences of drinking for the individual in question. How much alcohol is consumed is just one of the criteria according to which problem drinking is diagnosed. There are also psychological indicators, such as feeling guilty and neglecting usual responsibilities because of drinking, as well as incidents of deviant behaviour, for example driving while drunk and physical fights under the influence of alcohol.

We begin with the established measure of problem drinking – the alcohol dependence (AD) indicator used in Knopik et al. (2004), Slutske et al. (2002) and a number of other studies. This indicator is based on American Psychiatric Association's manual of mental disorders (APA, 2000) which lists nine symptoms of AD: (i) drinking more, and over longer periods of time, than intended; (ii) wanting but not being able to stop/cut down drinking; (iii) long time spent drinking; (iv) withdrawal symptoms that prevent fulfilling major obligations (e.g., work, school etc.); (v) giving up important activities because of drinking; (vi) carrying on drinking in the knowledge of its adverse effects; (vii) tolerance to alcohol (larger intake required to get drunk than when started); (viii) experiencing withdrawal symptoms (e.g., shaking, DTs); (ix) drinking to relieve withdrawal symptoms. An individual is diagnosed positively for alcohol dependence if he or she has three or more of the above symptoms. 27% of men and 9% of women in our sample reported symptoms consistent with the AD diagnosis.

Further in this section we construct a selection of drinking problem indicators, drawing on the Alcohol Use Disorders Identification Test (AUDIT) (Babor et al., 2001), as well as our own variables. We argue that these are useful additions to the well-known AD indicator as they will account for the severity of AD symptoms and focus on particular aspects of problem drinking such as inability to stop and behavioural and health complications.

The full set of questions we use to construct the drinking problem indicators is gven in table 3.2. This table also reports the occurrence of the symptoms described. A median man drinks 1-3 times a week, consuming 2-4 drinks on a typical day (not weekend). A median woman has 1-2 drinks a day once or twice a week. These figures are consistent with the median number of drinks last week (table 3.1), especially if we recall that much of the weekly alcohol intake falls on one day which is likely to be weekend and therefore not to be covered by the question on the number of drinks on a typical day.

Turning from the usual drinking patterns to the events which happened at least once in a lifetime, more than half (57%) of the respondents felt at some point that they should cut down their drinking. This statistic is potentially ambiguous, because the threshold from which drinking becomes problematic for the individual in question varies, and may be misleading if this threshold is positively related to the amount of alcohol usually consumed. Yet, drinking is

Table 3.2 Incidence of alcohol-related problems

	question	cases	%
1	How often do you have a drink containing alcohol?		
	Never	814	13.1
	Monthly or less	2,103	33.84
	2-4 times a month	1,566	25.2
	2-3 times a week	979	15.75
	4 or more times a week	752	12.1
2	How many drinks containing alcohol do you have on a typical day when you are drinking?		
	1 or 2	3,929	79.92
	3 or 4	687	13.97
	5 or 6	184	3.74
	7 to 9	44	0.9
	10 or more	72	1.46
	Have you ever...		
3	...felt you should cut down drinking?		
	No	2,356	43.15
	Yes	3,104	56.85
4	...tried, but failed, to cut down drinking?		
	No	5,169	96.17
	Yes	206	3.83
5	.. heard people objecting to your drinking?		
	No	4,291	79.86
	Yes	1,082	20.14
6	...felt bad or guilty about your drinking?		
	No	4,006	74.09
	Yes	1,401	25.91
7	...had a drink first thing in the morning?		
	No	5,294	95.92
	Yes	225	4.08
8	... been drinking heavily for 2 or more days in a row?		
	No	3,134	89.49
	Yes	368	10.51
9	...neglected your usual responsibilities because of drinking for 2 days or more?		
	No	5,137	95.45
	Yes	245	4.55
10	...had problems with friends and/or family because of drinking?		
	No	2,975	83.68
	Yes	580	16.32
11	...got into trouble at work because of drinking?		
	No	3,267	94.2
	Yes	201	5.8
12	...got into trouble because of drunk driving?		
	No	3,199	90.75
	Yes	326	9.25
13	... injured yourself or others when you were drinking?		
	No	3,015	85.27
	Yes	521	14.73
14	...approached anyone for help about your drinking?		
	No	3,768	96.79
	Yes	125	3.21
15	...been treated for a drinking problem, e.g. through Alcoholics Anonymous?		
	No	3,358	97.59
	Yes	83	2.41
16	...had delirium tremens, severe shaking or hallucinations after heavy drinking?		
	No	3,767	96.96
	Yes	118	3.04
17	...experienced other health problems because of drinking?		
	No	2,971	84.86
	Yes	530	15.14

Data samples: Alcohol-1 (questions 1-3, 5-7), SSAGA (14, 16), Alcohol-1 & SSAGA (the rest)

often regarded as excessive not only by the individual in question but also by those around him. Thus, 20% of the respondents report having heard other people objecting to their drinking, and 26% have felt bad or guilty about their drinking; 13% admitted both.

Recognising own drinking as problematic does not always lead to reducing or stopping it completely. Thus, a number of people report symptoms of (at least temporary) addiction to drink: trying, but failing, to cut down drinking (4%), drinking first thing in the morning (4%), and drinking heavily for more than two days in a row (11%). Statistics in table 3.2 also reveal that many people see their drinking as the reason behind a variety of personal problems – family and friends (16%), drunk driving (9%), health (3% and 15%), to name a few. The answers form a coherent picture, where alcohol intake is positively correlated with both inability to stop or reduce drinking and alcohol-related health and personal problems.

3.3.1 Standard tests for alcohol problems

We now construct the drinking problems indicators based on the AUDIT test, which we complement with three indicators of our own. The original version of the test includes ten questions concerning the amount, frequency of alcohol consumed, drinking patterns, (in)ability to stop drinking, its psychological and behavioural consequences. Data limitations prevent us from implementing the full version of it; however, there is a shorter version (the *five-shot test*, see Seppä et al., 1998), almost as accurate, which we can implement using the data from the Alcohol-1 sample.

We construct the five-shot test variable as the sum of the following indicators from table 3.2: how often do you have a drink containing alcohol? (“never” = 0 points, “monthly or less” = 1 point, “2-4 times a month” = 2 points, “2-3 times a week” = 3 points, “4+ times a week” = 4 points); how many drinks containing alcohol do you have on a typical day? (“1-2” = 0, “3-4” = 1, “5-6” = 2, “7-9” = 3, “10+” = 4); have people ever objected to your drinking? (“no” = 0, “yes” = 2); have you ever felt bad/guilty about your drinking (“no” = 0, “yes” = 2); have you ever had a drink first thing in the morning? (“no” = 0, “yes” = 2). A score higher than or equal to five indicates a potential alcohol problem needing further evaluation.

Another test which our data are enough to implement in full is the CAGE (cut down, annoy, guilty, ‘eye opener’) test (Mayfield et al., 1974). The CAGE test includes the last three questions of the five-shot test, plus the question “have you ever thought you should cut down on drinking?”. Each “yes” answer receives one point; a score of two points or higher indicates a potential alcohol problem.

Table 3.3 shows the distributions of scores on the five-shot and CAGE tests, separately for men and women, including the average years of schooling for each test score. According to the five-shot test, 45% of men and 20% of women from our sample show drinking patterns indicative of a potential alcohol problem. The CAGE test statistics give similar results: 45% for men and 22% for women. These statistics look surprisingly large; however these tests can only be used as a preliminary diagnostic tool to identify people whose behaviour warrants further investigation.

The data reveal a negative link between the five shot test score and mean years of schooling for men. The men whose score is below 5 (no alcohol problem) have on average 0.5 more years

Table 3.3 Five-shot and CAGE test results by gender

	men			women		
score	cases	yrs.school	%	cases	yrs.school	%
			five-shot test			
0	39	13.24	2.37	169	11.21	6.11
1	210	12.84	12.73	795	11.35	28.74
2	213	12.85	12.92	501	11.87	18.11
3	244	13.11	14.8	442	11.78	15.98
4	192	13.06	11.64	307	11.42	11.1
5	191	12.53	11.58	217	11.92	7.85
6	176	12.73	10.67	146	11.58	5.28
7	132	12.91	8	93	11.94	3.36
8	90	11.99	5.46	55	11.51	1.99
9	82	12.38	4.97	25	11.58	0.9
10+	80	11.80	4.84	16	11.66	0.58
			CAGE test			
0	495	12.50	26.51	1,601	11.09	46.3
1	525	12.82	28.12	1,107	11.64	32.01
2	369	12.79	19.76	485	11.76	14.03
3	409	12.46	21.91	246	11.87	7.11
4	69	12.07	3.7	19	10.39	0.55

of schooling than those with a score of 5 or more. There is a weaker negative link for men between the CAGE score and schooling. There are no statistically significant differences in schooling between women with different scores for either of the two tests.

3.3.2 Own problem drinking indicators

In addition to the known alcohol problem tests, we construct our own indicators, each focussing on a particular aspect of problematic drinking, as follows:

1. *Inability to reduce/stop drinking* (“yes” = 1, “no” = 0): (1) tried but failed to cut down drinking; (2) had a drink first thing in the morning; (3) drank heavily for 2 or more days in a row; (4) approached someone for help about drinking; (5) felt bad or guilty about your drinking.

2. *Behavioural problems caused by drinking*: (1) heard people objecting to your drinking; (2) neglected your usual responsibilities because of drinking for 2 days or more; (3) had problems with friends and/or family because of drinking; (4) had problems at work/study because of drinking; (5) got into trouble because of drunk driving; (6) injured yourself or others while drinking.

3. *Health problems caused by drinking*: (1) received treatment for a drink problem; (2) had delirium tremens, severe shaking or hallucinations after heavy drinking; (3) had other health problems because of drinking.

Table 3.4 reports the distributions of each of the three indicators by gender. The data reveal that a surprisingly large share of respondents have experienced problems caused by drinking. Thus, about half of men and a third of women report having difficulty in reducing the amount of alcohol they consumed at least on one occasion in their lifetime. Further, more than half of men and 20% of women have had behavioural problems caused by drinking. (Zooming in on particularly grave problems, such as drunk driving and injuring self or others while drinking, we see these happening with 29% of men and 7% of women.) There is also a relatively high prevalence of health problems perceived to be caused by drinking: 20% for men and 9% for women. All these stylised facts together once again illustrate the dangers of excessive drinking.

Table 3.4 Alcohol problem indicators by gender

	men			women		
score	cases	yrs.school	%	cases	yrs.school	%
		Inability to stop/reduce drinking				
0	719	12.82	51.69	1,417	11.60	70.71
1	439	12.73	31.56	487	11.89	24.3
2	144	12.72	10.35	69	12.09	3.44
3	60	12.41	4.31	22	12.55	1.1
4	20	11.03	1.44	4	8.63	0.2
5	9	11.39	0.65	5	10.50	0.25
		Behavioural problems caused by drinking				
0	666	12.95	48.12	1,589	11.67	79.09
1	305	12.85	22.04	270	11.57	13.44
2	180	12.35	13.01	90	12.10	4.48
3	105	12.29	7.59	27	11.40	1.34
4	69	12.01	4.99	10	12.55	0.5
5	33	11.91	2.38	4	11.00	0.2
6	26	11.50	1.88	19	11.66	0.95
		Health problems caused by drinking				
0	1,091	12.78	80.04	1,810	11.64	90.95
1	211	12.52	15.48	153	12.02	7.69
2	51	12.33	3.74	19	12.08	0.95
3	10	11.05	0.73	8	9.50	0.4

Again, as with the five shot and CAGE tests, the men with more schooling tend to report fewer alcohol-related problems. Thus, the difference in mean schooling between those who never had difficulty reducing own drinking and those who had this problem at least once is about 0.2 years. This difference s even larger for the behavioural and health problems scores: 0.5 and 0.35 years, respectively. On the other hand, more educated women appear to be more likely to have difficulty reducing drinking or develop a health problem, and just as likely to have a behavioural problem because of drinking.

3.4 Controls

Our most important control is of course twin pairs. Operationally, this control is achieved by introducing twins fixed effects in our regressions (see section 4). Twins share family background and (some of) genotype, which factors are likely to play a critical role in the formation of habits, intelligence and attitudes and therefore influence schooling and drinking later in life. So, comparing the results based on within twin pair comparisons with those drawing on the entire observable variation (within and between twin pairs) will show to what extent the link between schooling and drinking is driven by these often unobservable factors.

The data enable us to further distinguish between monozygotic (identical) and dyzygotic (fraternal) twins. Zygosity was determined based on self-reports and standard questions (such as physical likeness), which methods together give fairly accurate determination. Monozygotic twins make up for 41.3% and 50.55% of male and female twins in our data, respectively. Zygosity is important for isolating genetic factors affecting both educational attainment and alcohol use because, unlike dyzygotic, monozygotic twins have the same DNA and genes. So, within monozygotic twin par comparisons will offer the ultimate control over both family and genetic confounding factors.

Twin pair controls, though crucial for the identification of the causal effect of schooling on the variables of our interest, are “expensive” in terms of the variation in schooling and drinking between the twin pairs which is lost when we limit ourselves to within twin pair comparisons[3]. Therefore, as a side inquiry we investigate to what extent observable twins characteristics can proxy for the unobservable so that the true causal link between education and alcohol consumption may be revealed without having to limit ourselves to exploiting the within twin pair variation only. Accordingly, we include parents’ schooling and drinking habits. Parents’ schooling is measured in the same way as twins’. Drinking habits are measured as a dummy variable which takes the value of 1 if father (mother) drank more than 7 (4) drinks per week. These levels of alcohol consumption correspond to the top 25-30% of reported amounts of alcohol consumed per week by the twins’ parents.

In addition to the twin pair controls, we use birth weight, the month when the questionnaire was filled in, the age when the individual started to drink, and their past drinking. Birth weight has been used in the labour economics literature as a control for health status which determines a number of personal and labour market outcomes (Black et al., 2007). The average birth weight was 2.61 kg for men and 2.44 kg for women. The time of the year is important because alcohol consumption is seasonal. Most of the respondents (around 70%) return their questionnaires between November and January, the period when an increase in drinking during the festive

[3] For example, the “between” variation in the number of drinks last week and schooling makes about 55% and 70% of the total, respectively.

season is followed by a lull thereafter.

Age when started to drink controls for the time of exposure to drinking. The length of exposure to drinking may affect the degree of addiction to alcohol which is an important determinant of both alcohol consumption and the incidence of drinking problems. The median age when started to drink is 17 for men and 18 for women. 95% of men and 90% of women will have tried alcohol by age 25. Thus, with the average age in 1980 of just under 30, most of the subjects will have had some history of drinking alcohol.

Finally, apart from giving an additional degree of control for alcohol addiction, including past drinking in our regressions will also help isolate the effect of misreporting own drinking, assuming that the tendency to under- or overreport of a given individual does not change over time. Note that, as far as misreporting is concerned, this is a different control from that introduced by twins fixed effects. The twins fixed effects control for characteristics, common to both twins, that affect their schooling and (mis)reporting. Past drinking, on the other hand, is meant to control for *individual* propensity to misreport to the extent that misreporting depends on regular alcohol intake.

4 Estimation issues

We begin to model the statistical relationships of interest to this study by expressing a given dependent variable y (e.g., alcohol problem counts) as a linear[4] function of the observed (vector $\mathbf{x}$) and unobserved (α) variables, and the error term u:

$$y_{ij} = \mathbf{x}_{ij}\beta + \alpha_i + u_{ij} \tag{4.1}$$

where $i = 1, 2, \ldots$ is the twin pair index, $j = 1, 2$ is the individual twin index. The observed variables $\mathbf{x}_{ij}$ include years of schooling and all the applicable controls as outlined in sub-section 3.4 and are allowed to vary between twin pairs and individual twins. The unobserved α_i is the total result of the influences of family background, genes and other factors shared by the two twins; therefore, it is only allowed to vary between twin pairs. The α's may or may not be correlated with the $\mathbf{x}$'s. We do allow such correlation, however, treating the α's as fixed effects and not random. (For instance, family background may be correlated with both drinking behaviour and demand for education.)

Consistent estimation of equation 4.1 requires either a direct estimation of α or removing it from the original equation through a suitable transformation within each twin pair. Assuming additivity of its effect, α can be dispensed with by taking within twin pair differences and estimating

$$\triangle y_i = \triangle \mathbf{x}_i \beta + \triangle u_i \tag{4.2}$$

using maximum likelihood or least squares. If the effect of α is instead multiplicative, the required transformation is log-differencing, which possibility we will explore as a robustness check.

We estimate equation 4.2 separately for men and women. Starting with the regression without the twin pair controls, we proceed to the twins fixed-effects specification, which we then confine to the monozygotic twins sub-sample. To ensure comparability of regression results from different specifications, we limit our research sample only to cases for which all the control variables are non-missing, both for the individual in question and for his or her twin. Missing observations are frequent, so to try and and increase the number of available observations, we impute the missing values of non-key regression variables with their twin reports. Comparing the regression results on the data with and without imputations reveals no major differences in the regression results.

[4] We consider a nonlinear specification in sub-section 6.1.

4.1 Measurement error in schooling

Estimating equation 4.2 may still render inconsistent results if schooling is measured with error. Measurement error is an important feature of self-reported data such as ours (recall sub-section 3.1). In fact, the correlation between own- and twin-reported schooling (the "reliability ratio") in our data is about 0.85, close to 0.9 reported in (Ashenfelter and Krueger, 1994, p. 1160) for U.S. twins schooling, implying that about 15% of the observed variance in schooling is due to measurement error. Measurement error attenuates the estimated coefficient compared to its true value, especially in equation 4.2, because its share in the total variance of $\triangle$schooling will increase due to own and twin self-reported years of schooling being correlated (Griliches, 1979). As we noted in sub-section 3.1, this correlation is rather high.

Fortunately, our data offer a selection of proxies for $\triangle$schooling which can act as instruments for each other. For every twin pair, we have four measures of schooling, all being consistent estimates of the true schooling in the absence of systematic reporting bias: s_1^1, s_1^2, s_2^1, and s_2^2, where s_m^n is the years of schooling of twin m as reported by twin n. Out of these measures we can construct proxies for the true $\triangle s$ as follows: $\triangle s' = s_1^1 - s_2^2$, $\triangle s'' = s_1^2 - s_2^1$, $\triangle s^* = s_1^1 - s_2^1$, and $\triangle s^{**} = s_1^2 - s_2^2$.

Ashenfelter and Krueger (1994) consider two IV approaches to estimating equation 4.2: i) using $\triangle s'$ as a proxy for $\triangle s$ and instrumenting it with $\triangle s''$ (in words: use twin-reported schooling as the instrument for own-reported schooling); and ii) using $\triangle s^*$ instrumented by $\triangle s^{**}$ (use the difference in schooling as reported by one twin as the instrument for the same as reported by the other twin). The first approach assumes that the measurement errors in all four reported measures of schooling are uncorrelated. The second approach relaxes this assumption by allowing for correlation between the measurement errors in responses of the same twin to correlate through an individual-specific error component which is cancelled through taking the difference between the same individual's responses. In other words, if someone overstates his own education he is also likely to overstate his twin's education. We adopt the second approach because its assumptions seem more plausible.

5 Basic regression results

5.1 Schooling and drinking

Table 5.1 reports the regression results for the number of drinks last week in 1988 for men and women. The coefficients show the effects of a one-unit change in the respective variable on the number of drinks. Column 1 in table 5.1 reports linear regression results on the pooled sample. Twin fixed effects are introduced in columns 2-4. Column 2 reports results for the entire sample, while column 3 concentrates on the sub-sample of monozygotic twins only, thus controlling for both family background and genes. Column 4 contains regression results for monozygotic twins corrected for measurement error in schooling as discussed in sub-section 4.1.

Turning to the regression estimates, we observe that regular alcohol intake by men falls by 0.2-0.3 drinks per week with every extra year of schooling. That the estimates from the pooled and the fixed effects specifications are not far from each other implies that the family background variables available to us (parental education and drinking habits) can adequately capture the variance in drinking which is due to such unobservables. In particular, drinking by both parents is a significant predictor of alcohol intake in every specification. Schooling, on the other hand, is insignificant. Correcting for measurement error does increase the estimate for men (there is evidently a misspecification problem for women), but this increase is small.

In fact, current drinking cannot be explained by much else apart from previous drinking, with every drink had eight years ago increasing the current (1988) number of drinks per week by about 0.3-0.4[5]. This is a powerful covariate which might have obscured the effect of schooling. However, even in the absence of other controls, schooling is not significantly related to either current or past drinking. Thus we cannot conclude a causal relationship between schooling and regular drinking, either direct or mediated by other variables.

We now turn to discussing the effect of schooling on excessive drinking measured as (1) the maximum number of drinks had on a day of a typical week, (2) the maximum number of drinks ever had in a single day, and (3) standard deviation of weekly alcohol intake (see table 3.1). Table 5.2 reports. For men, schooling appears at first to affect significantly all three measures of excessive drinking (specification 1); however, when controlling for unobservables by means of twin fixed effects (specification 2–4), the estimates drop in magnitude and cease to be significant, even after correcting for measurement error in schooling. For women, controlling for unobservables increases the magnitude of the estimates for schooling, but its significance record is also poor. So, unlike for the number of drinks last week (table 5.1), the available family background variables fail to adequately approximate the effects of unobservables on both

[5] This is probably an overestimate because part of the correlation between current and past drinking runs through the common error of misreporting. However, there is also a downward bias in this estimate due to the measurement error in the past drinking variable.

Table 5.1 The impact of schooling on the number of drinks last week in 1988

	All twins				Monozygotic twins			
	(1) Pooled		(2) Twin fixed effects		(3) Twin fixed effects		(4) Same, corrected for measur. error	
MEN								
years of schooling	-0.194		-0.186		-0.258		-0.307	
	0.16		*0.24*		*0.43*		*0.63*	
high school	0.779		-0.130		-3.790		-3.730	
graduate	*1.02*		*1.44*		*2.43*		*2.50*	
number of drinks	0.358	***	0.342	***	0.306	***	0.306	***
last week in 1980	*0.02*		*0.03*		*0.04*		*0.04*	
age started to	-0.110		-0.173		0.060		0.059	
drink	*0.09*		*0.17*		*0.28*		*0.28*	
birth weight	1.278	***	-0.440		0.563		0.567	
	0.43		*1.09*		*2.13*		*2.13*	
N	1428		714		542		542	
WOMEN								
years of schooling	-0.110		-0.228		-0.048		0.030	
	0.10		*0.16*		*0.22*		*0.32*	
high school	0.329		-0.284		-0.491		-0.561	
graduate	*0.50*		*0.67*		*0.80*		*0.83*	
number of drinks	0.464	***	0.421	***	0.323	***	0.322	***
last week in 1980	*0.02*		*0.02*		*0.04*		*0.04*	
age started to	-0.068	**	-0.130	*	0.005		0.006	
drink	*0.03*		*0.07*		*0.08*		*0.08*	
birth weight	0.084		0.166		-0.866	*	-0.873	*
	0.23		*0.50*		*0.52*		*0.52*	
N	2406		1203		643		643	

Other controls (where applicable): parents' education and drinking habits,
twins' age category, month when questionaire was filled in.
Standard deviations are in italics.
Estimates significant 1%, 5% and 10% significance level are marked ***, ** and *, respectively.

schooling and excessive drinking. The significance of lagged dependent variable in the regressions in table 5.2 suggests that drinking behaviour persists over time, reinforcing our earlier conclusions for regular drinking.

5.2 Schooling and alcohol-related problems

While there appears to be no significant relationship between schooling and drinking, regular or excessive, schooling might still help prevent alcohol-related problems. Our regression specifications for alcohol-related problems include past drinking, age when started to drink and birth weight as control variables. Past drinking, which comprises the total number of drinks last week in 1980, the maximum daily intake on a typical week in 1980, and the maximum daily

Table 5.2 The impact of schooling on excessive drinking

	MEN				WOMEN			
	All twins		Monozygotic twins		All twins		Monozygotic twins	
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Maximum drinks per day on a typical week								
years of schooling	-0.193***	-0.043	-0.029	0.112	-0.062*	-0.133**	-0.071	-0.045
	0.06	*0.10*	*0.21*	*0.31*	*0.04*	*0.06*	*0.09*	*0.13*
max drinks	0.318***	0.305***	0.274***	0.275***	0.322***	0.260***	0.091**	0.090**
per day in 1980	*0.02*	*0.03*	*0.05*	*0.05*	*0.01*	*0.02*	*0.04*	*0.04*
N	1426	713	272	272	2386	1193	637	637
Maximum drinks per day ever in lifetime								
years of schooling	-0.617***	-0.126	-0.431	-0.733	-0.055	0.187	0.037	0.030
	0.15	*0.23*	*0.41*	*0.59*	*0.08*	*0.12*	*0.14*	*0.21*
N	1506	753	284	284	2596	1298	693	693
Standard deviation of weekly alcohol intake								
years of schooling	-0.034*	0.005	-0.024	-0.032	-0.010	-0.028	-0.009	-0.010
	0.02	*0.03*	*0.07*	*0.10*	*0.01*	*0.02*	*0.02*	*0.04*
std deviation	0.315***	0.297***	0.247***	0.247***	0.348***	0.277***	0.156***	0.156***
in 1980	*0.02*	*0.03*	*0.05*	*0.05*	*0.01*	*0.02*	*0.03*	*0.03*
N	1622	811	297	297	2882	1441	764	764

Specifications: (1) - pooled OLS, (2) - all twin fixed effects, (3) - MZ twin effects, (4) - same, corrected for meas. error.
Other controls (where applicable): parents' education and drinking habits, high school graduation dummy
age when started to drink, birth weight.

intake ever in a lifetime, is used to control for the individual history of both regular as well as excessive drinking. We do not include present drinking, because the alcohol problem events may have preceeded it. To save space, we focus only on the fixed-effects regression results for monozygotic twins, skipping the previous steps shown in tables 5.1 and 5.2 (these results are available on request).

Looking at the results for the standard alcohol problem indicators (table 5.3), we find a negative impact of schooling on the alcohol dependence (AD) indicator and CAGE test score for men. Instrumenting own-reported schooling to correct for measurement error predictably increases its estimates; in fact, the only significant estimates for men are from the IV regressions. According to our regressions, an extra year of schooling reduces the chances of developing AD for men by about 5%. Recalling that an average man in our sample has 12.5 years of schooling and is 27% likely to have AD symptoms, men attaining the highest level of scholing (17 years) run a very small risk of having an AD, while for those with only 5 years of schooling the risk is as high as 60%. Similarly for the CAGE test score, given the sample average of 1.4, completing

Table 5.3 The impact of schooling on the incidence of alcohol-related problems, standard indicators

	MZ twin fixed effects				MZ twin fixed effects			
	uninstrumented		instrumented		uninstrumented		instrumented	
	Alcohol dependence (AD) indicator							
	MEN				WOMEN			
years of schooling	-0.021		-0.049	*	0.029	**	0.009	
	0.02		*0.03*		*0.01*		*0.02*	
number of drinks	0.002		0.002		0.000		0.000	
last week in 1980	*0.00*		*0.00*		*0.00*		*0.00*	
max drinks p/day on	0.009		0.009		0.004		0.004	
typical week in 1980	*0.01*		*0.01*		*0.01*		*0.01*	
max drinks	0.004		0.004		-0.001		-0.001	
per day ever	*0.00*		*0.00*		*0.00*		*0.00*	
N	211				369			
	Five-shot test							
years of schooling	0.022		-0.138		-0.114		-0.078	
	0.10		*0.14*		*0.07*		*0.11*	
number of drinks	0.068	***	0.070	***	0.046	***	0.045	***
last week in 1980	*0.01*		*0.01*		*0.02*		*0.02*	
max drinks p/day on	-0.003		-0.009		0.010		0.008	
typical week in 1980	*0.04*		*0.04*		*0.04*		*0.04*	
max drinks	0.047	***	0.046	***	0.138	***	0.138	***
per day ever	*0.01*		*0.01*		*0.02*		*0.02*	
N	236				470			
	CAGE test							
years of schooling	-0.034		-0.150	**	-0.053	*	-0.036	
	0.05		*0.07*		*0.03*		*0.05*	
number of drinks	0.018	***	0.019	***	0.006		0.006	
last week in 1980	*0.01*		*0.01*		*0.01*		*0.01*	
max drinks p/day on	0.016		0.011		0.003		0.002	
typical week in 1980	*0.02*		*0.02*		*0.02*		*0.02*	
max drinks	0.015	**	0.014	*	0.040	***	0.040	***
per day ever	*0.01*		*0.01*		*0.01*		*0.01*	
N	270				632			

Other controls: high school graduation dummy, age when started to drink, birth weight.

secondary education can make a difference between acceptable (<2) and potentially problematic ($\geq$2) CAGE scores.

For women, on the other hand, the evidence on the effect of schooling on problem drinking is less convincing. Here schooling appears to affect positively the chances of developing an AD, though the magnitude of this effect goes down as we instument schooling to correct for measurement error, which puzzling effect we observe for all the three variables in table 5.3.

There is some evidence that schooling helps reduce CAGE test score, but again this effect is weak.

Table 5.4 The impact of schooling on the incidence of alcohol-related problems, own indicators

	MZ twin fixed effects				MZ twin fixed effects			
	uninstrumented		instrumented		uninstrumented		instrumented	
	Inability to stop/reduce drinking							
	MEN				WOMEN			
years of schooling	-0.037		-0.113	*	0.048	*	0.024	
	0.04		*0.06*		*0.03*		*0.04*	
number of drinks	0.009		0.010	*	0.003		0.003	
last week in 1980	*0.01*		*0.01*		*0.01*		*0.01*	
max drinks p/day on	0.017		0.014		-0.013		-0.012	
typical week in 1980	*0.02*		*0.02*		*0.02*		*0.02*	
max drinks	0.019	***	0.019	***	0.021	***	0.021	***
per day ever	*0.01*		*0.01*		*0.01*		*0.01*	
N	219				357			
	Behavioural problems caused by drinking							
years of schooling	0.041		-0.062		-0.039		-0.086	*
	0.06		*0.08*		*0.04*		*0.05*	
number of drinks	0.012		0.013	*	0.002		0.003	
last week in 1980	*0.01*		*0.01*		*0.01*		*0.01*	
max drinks p/day on	0.051	**	0.047	**	0.006		0.008	
typical week in 1980	*0.02*		*0.02*		*0.02*		*0.02*	
max drinks	0.022	***	0.022	***	0.031	***	0.031	***
per day ever	*0.01*		*0.01*		*0.01*		*0.01*	
N	214				360			
	Health problems caused by drinking							
years of schooling	-0.028		-0.108	***	0.020		-0.002	
	0.03		*0.04*		*0.02*		*0.02*	
number of drinks	0.003		0.004		-0.008	**	-0.007	**
last week in 1980	*0.00*		*0.00*		*0.00*		*0.00*	
max drinks p/day on	0.007		0.003		0.009		0.010	
typical week in 1980	*0.01*		*0.01*		*0.01*		*0.01*	
max drinks	0.014	***	0.013	***	0.010	***	0.010	***
per day ever	*0.00*		*0.00*		*0.00*		*0.00*	
N	208				350			

Other controls: high school graduation dummy, age when started to drink, birth weight.

Turning to the results for our own indicators (table 5.4), we see that, for men, schooling is a significant and important determinant of inability to stop/reduce drinking and health problems (perceived to be) cased by drinking. For women, schooling reduces the incidence of behavioural problems. There are also signs for a *positive* effect of schooling on inability to reduce drinking for women, but the evidence is not conclusive. Our estimates suggest that 10-12 years of schooling can make a difference for an average individual between having an alcohol problem and not having it. Past drinking remains an important determinant of alcohol-related problems for both sexes. It is not surprising that people who drink more are also more likely to have developed an alcohol problem of some sort. What is remarkable is that schooling helps reduce the incidence of drinking problems *given* the amount of alcohol consumed. [6] It looks like schooling reduces the incidence of alcohol-related problems by helping individuals control their behaviour so that their drinking is less likely to result in a problem.

[6] We have also tried running the alcohol problems regressions without controls for past drinking, but the results were similar to the full specification, so we skipped them.

6 Extensions and robustness checks

6.1 Count data regressions

Note that many of the dependent variables in our study, such as the number of drinks and and alcohol problems, are non-negative integers or “counts”. The observations concentrate on relatively few discrete values and are skewed to the right. Thus the normal distribution, upon which the linear estimator we have used so far is based, renders only a poor approximation of the dependent variables’ empirically observed distributions. Although many distributions do converge to normal in the asymptote, estimation efficiency losses may occur on a limited sample such as ours. Here we specify and run Poisson regressions, which the basic estimation technique in the count data econometrics (Cameron and Triverdi, 1998), for the number of drinks and alcohol problems[7].

This model assumes that the number of occurrences of a particular event (in our case, having a drink or developing an alcohol-related problem) for a given individual within a defined time period (e.g., a week) follows a Poisson distribution. The probability of y_{ij} realisations of this event is defined as

$$p(y_{ij}) = \frac{e^{-E(y_{ij})} \cdot E(y_{ij})^{y_{ij}}}{y_{ij}!}, \tag{6.1}$$

where, to limit the domain of $E(y)$ to positive values only, it is specified as

$$E(y_{ij}) = \alpha_i \cdot e^{\mathbf{x}'_{ij}\beta} \tag{6.2}$$

Note that for this specification β measures the percentage change in the dependent variable in response to a small change in the respective x. To convert the implied percentage changes in the absolute changes, we use sample average of the dependent variable as benchmark, which is common practice.

The fixed-effects estimator for β can be derived from maximising the likelihood function based on the assumed Poisson distribution (equation 6.1). Hausman et al. (1984) showed that the estimation procedure can be simplified considerably by maximising the likelihood of the

[7] For one important dependent variable in our study – the AD score – a fixed-effects logit estimator is applied. We skip the detalied discussion of this estimator here, noting only that a very similar approach based on conditioning the likelihood function on the within twin pair sums of the dependent variable needs to be applied in order to obtain consistent estimates.

distribution of y conditional on the distribution of the within twin pair sum of y: [8]

$$\prod_i \prod_j p(y_{ij}|y_{i1}+y_{i2}) = \prod_i \left[\frac{e^{-\alpha_i \sum_j e^{\mathbf{x}'_{ij}\beta}} \cdot \alpha_i^{y_{i1}+y_{i2}} \cdot e^{\mathbf{x}'_{i1}\beta \cdot y_{i1} + \mathbf{x}'_{i2}\beta \cdot y_{i2}}}{y_{i1}! \cdot y_{i2}!} \cdot \frac{(y_{i1}+y_{i2})!}{e^{-\alpha_i \sum_j e^{\mathbf{x}'_{ij}\beta}} \cdot \left(\alpha_i \sum_j e^{\mathbf{x}'_{ij}\beta}\right)^{y_{i1}+y_{i2}}} \right]$$

$$= \prod_i \left[\frac{e^{\mathbf{x}'_{i1}\beta \cdot y_{i1} + \mathbf{x}'_{i2}\beta \cdot y_{i2}}}{y_{i1}! \cdot y_{i2}!} \cdot \frac{(y_{i1}+y_{i2})!}{\left(\sum_j e^{\mathbf{x}'_{ij}\beta}\right)^{y_{i1}+y_{i2}}} \right] \tag{6.3}$$

The estimator in equation 6.3 is consistent for a wide class of discrete distributions including Poisson as long as the regressors are exogenous and the conditional expectation of y (equation 6.2) is correctly specified (Wooldridge, 1999, pp. 80-81; Cameron and Triverdi, 2005, pp. 805-806). It is therefore rather immune to the 'overdispersion' problem, when the dependent variable's conditional mean is not equal to its conditional variance (as is the case for Poisson distribution). Using panel-robust standard errors (Wooldridge, 1999, p. 83) will further strengthen statistical properties of the Poisson fixed-effects estimator.

The price to pay for the advantages of the Poisson fixed-effects estimator is the loss of twin groups for whom the dependent variable is zero for both twins. As can be seen from equation 6.3, the likelihood function for such groups is 1 regardless of the values of β, so they do not contribute to the total likelihood. On the other hand, the reduction of the effective sample to the twin pairs where at least one twin has an alcohol problem will permit a closer look at the effect of schooling on problem drinking by twins with a family and/or genetic predisposition to it. This effect should arguably be larger than the population-averaged one which we estimated earlier.

Table 6.1 compares count data regression estimates of the impact of one year of schooling on a variety of indicators, with their linear counterparts (corrected for measurement error in schooling) applied on the same sub-samples of monozygotc twins. Because no significant results are found for women, we focus here on the male sample only. The two sets of estimates support our earlier findings on the negative impact of schooling on problem drinking. IV estimates often exceed those from count regressions, which is expected since the Poisson estimator is not immune to measurement error in schooling. More importantly, the IV estimates from table 6.1, obtained for twins of whom at least one had an alcohol problem at least once in life, exceed those reported earlier for the entire sample (tables 5.2-5.4). The differences can be small (as with the CAGE test score) or large (as with the AD score), depending on how different the relevant samples are. A general tendency, however, is there, suggesting that schooling matters more for people with a family history of problem drinking, than for the whole population.

[8] This result stems from the Rao-Blackwell theorem, which implies that conditioning the likelihood function for a given parameter on the sufficient statistic for this parameter produces a no worse estimation result than the original, unconditional, likelihood. For a random variable x with the probability distribution depending on a parameter θ, $P(x,\theta)$, the sufficient statistic for θ is a function of x, $t(x)$, such the probability distribution of x conditional on $t(x)$ does not depend on θ. That is, $P(x|\theta,t) = P(x|t)$. If y follows a Poisson distribution (equation 6.1), the sufficient statistic for each α_i is $\sum_j y_{ij}$. Indeed, as can be seen from equation 6.3, the likelihood of y conditional on $y_{i1}+y_{i2}$ does not depend on α.

Table 6.1 Estimates of the impact of schooling on alcohol intake and problem drinking

	MEN				WOMEN	
Dependent variables	count regression		linear IV		count regression	linear IV
no. drinks last week	-0.685		-0.305		-0.007	0.049
	0.47		*0.65*		*0.25*	*0.34*
max drinks p/day on a typical week	-0.148		0.117		-0.014	-0.042
	0.19		*0.31*		*0.08*	*0.13*
max drinks p/day ever	-0.611	*	-0.724		-0.102	0.034
	0.37		*0.59*		*0.14*	*0.21*
AD	-0.292	**	-0.443	**	0.077	0.054
	0.14		*0.19*		*0.08*	*0.10*
five-shot	-0.084		-0.138		-0.058	-0.079
	0.11		*0.14*		*0.08*	*0.11*
CAGE	-0.101	*	-0.161	**	-0.050	-0.045
	0.05		*0.07*		*0.04*	*0.06*
inability to stop	-0.075		-0.117		0.041	0.045
	0.06		*0.08*		*0.05*	*0.07*
behavioural problems	-0.125		-0.099		-0.077	-0.199
	0.10		*0.11*		*0.08*	*0.13*
health problems	-0.215	**	-0.300	**	0.101	-0.002
	0.09		*0.12*		*0.10*	*-0.40*

6.2 Personality

Part of the link between schooling and (problem) drinking may be explained through personality traits. Personality has been found to affect significantly the risk of alcohol dependence by Slutske et al. (2002), who used the same twins data as ours, but the extent to which personality mediates the relationship between schooling and problem drinking has not as yet been established. In this subsection we probe into this issue by including measures of personality in the regressions for alcohol-related problem scores, leaving other controls the same. Following the approach in Slutske et al. (2002), we combine questions from the Tridimensional and Eysenck Personality Questionnaires and calculate the novelty seeking, harm avoidance, reward dependence, extraversion, neuroticism, psychoticism and lie scores.

A principal component analysis on these seven scores yields three dimensions of personality, together explaining 70% of the total variance in the scores. The first dimension, labelled "negative emotionality" (NE), explains 32% of the total variance in the scores and is most significantly associated with harm avoidance and neuroticism. High scores on this dimension point to high levels of anxiety, insecurity, mood changes and general unhappiness. The second dimension, “behavioural undercontrol” (BU), encapsulating novelty seeking, psychoticism and lie and explaining 21%, is associated with implusiveness, thrill seeking, rebelliousness and lack of responsibility. The third dimension, “positive emotionality” (PE), is mostly made up of reward dependence and extraversion and explains the remaining 17%. High scores on this dimension are earned by lively, friendly and outgoing people.

Empirically, personality is a potentially powerful environmental determinant of problem drinking. Genes account for about a third in the variation in personality dimensions among individuals, less than, for example, in schooling. Positive correlations between BU score and problem drinking are consistently observed for both sexes. Also, higher scores on NE and BU are associated with significantly less schooling for men (one-standard deviation change in NE and BU accounts for just under 0.2 years of schooling), even controlling for monozygotic twin fixed effects. Curiously, the same scores are positively associated with schooling for women, though the correlations are not significant. Can personality be one of the mechanisms through which schooling affects problem drinking by men?

Table 6.2 reports further regression results for all problem drinking measures with and without personality dimensions included. (The inclusion of new variables reduces the sample size on which the results in table 6.2 are based, hence slight differences in the estimates with tables 5.3 and 5.4.) The most important personality dimension for problem drinking by both sexes is behavioural undercontrol; Slutske et al. (2002) reach a similar conclusion for their alcohol addiction measure. For women, the link between schooling and problem drinking is still insignificant, even though the estimates for schooling often go up once personality is included (because of positive correlations of BU with schooling, and negative with problem drinking, for women). For men, on the contrary, the inclusion of BU reduces the estimates for schooling most of the times [9]. Personality thus accounts for part of the previously observed effect of schooling on problem drinking. However, whether education teaches one to be more wise and responsible, or more wise and responsible people achieve more schooling, remains an open question. At any rate, that schooling remains a significant determinant of behavioural and health problems caused by drinking suggests mechanisms other than BU through which the impact of schooling is realised.

[9] Health problems perceived to be caused by drinking is a notable exception, but the underlying sample is quite small (94 observations), so the results may be susceptible to outliers.

Table 6.2 The impact of schooling on problem drinking with and without personality controls

	MEN				WOMEN			
	Max. drinks on a single day ever							
years of schooling	-0.655	*	-0.581		-0.120		-0.153	
NE			-0.195				-0.412	**
PE			0.614				0.362	*
BU			0.634				0.607	***
	Five-shot test							
years of schooling	-0.115		-0.092		-0.071		-0.078	
NE			-0.158				-0.150	*
PE			-0.143				0.010	
BU			0.277	**			0.167	
	CAGE test							
years of schooling	-0.110	**	-0.078		-0.050		-0.048	
NE			-0.004				-0.037	
PE			-0.096				-0.004	
BU			0.217	***			0.164	***
	Inability to stop/reduce drinking							
years of schooling	-0.081		-0.054		0.026		0.024	
NE			0.048				0.001	
PE			0.074				-0.042	
BU			0.201	**			0.116	
	Behavioural problems caused by drinking							
years of schooling	-0.163	*	-0.138	*	-0.083		-0.124	
NE			0.289	***			-0.039	
PE			-0.019				-0.090	
BU			0.137	*			0.525	***
	Health problems caused by drinking							
years of schooling	-0.279	***	-0.317	***	-0.031		-0.089	
NE			0.014				0.140	
PE			0.054				-0.010	
BU			0.484	***			0.527	**

Standard deviations are skipped for brevity. All other controls remain.

7 Conclusions

TO BE COMPLETED

References

Anderson, P. and B. Baumberg, 2006, *Alcohol in Europe*, London: Institute of Alcohol Studies.

APA, 2000, *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR, 4th edition*, American Psychiatric Association.

Ashenfelter, O. and A. Krueger, 1994, Estimates of the economic return to schooling from a new sample of twins, *American Economic Review*, vol. 84, no. 5, pp. 1157–1173.

Babor, T., J. Higgins-Biddle, J. Saunders and M. Monteiro, 2001, *The Alcohol Use Disorders Identification Test Guidelines for Use in Primary Care*, World Health Organization.

Baltagi, B. and J. Griffin, 1995, A dynamic model for liquor: the case for pooling, *Review of Economics and Statistics*, vol. 77, no. 3, pp. 545–554.

Behrman, J. and P. Taubman, 1989, Is schooling “mostly in the genes”? Nature-nurture decomposition using data on relatives, *Journal of Political Economy*, vol. 97, no. 6, pp. 1425–1446.

Behrman, J. and M. Rosenzweig, 1999, ‘Ability’ biases in schooling returns and twins: A test and new estimates, *Economics of Education Review*, vol. 18, no. 2, pp. 159–167.

Black, S., P. Devereux and K. Salvanes, 2007, From the cradle to the labor market? the effect of birth weight on adult outcomes, *Quarterly Journal of Economics*, vol. 122, no. 1, pp. 409–439.

Bonjour, D., L. Cherkas, J. Haskel, D. Hawkes and T. Spector, 2003, Returns to education: Evidence from U.K. twins, *American Economic Review*, vol. 93, no. 5, pp. 1799–1812.

Cameron, A. and P. Triverdi, 1998, *Regression analysis for count data*, Cambridge University Press.

Cameron, A. and P. Triverdi, 2005, *Microeconometrics: Methods and applications*, Cambridge University Press.

Cutler, D.M. and E. Glaeser, 2005, What explains differences in smoking, drinking, and other health-related behaviors?, *American Economic Review*, vol. 95, no. 2, pp. 238–242.

Cutler, D.M. and A. Lleras-Muney, 2006, Education and health: Evaluating theories and evidence, NBER working paper 12352.

Droomers, M., C. Schrijvers and J. Mackenbach, 2004, Educational differences in starting excessive alcohol consumption: explanations from the longitudinal GLOBE study, *Social Science and Medicine*, vol. 58, no. 10, pp. 2023–2033.

Foxcroft, D., D. Ireland, D.J. Lister-Sharp, G. Lowe and R. Breen, 2003, Longer-term primary prevention for alcohol misuse in young people: a systematic review, *Addiction*, vol. 98, pp. 397–411.

Furnée, C.A., W. Groot and H. Maassen van den Brink, 2008, The health effects of education: a meta-analysis, *European Journal of Public Health*, vol. 18, no. 4, pp. 417–421.

Griliches, Z., 1979, Sibling models and data in economics: Beginnings of a survey, *Journal of Political Economy*, vol. 87, no. 5, pp. S37–S64.

Hausman, J.A., B. Hall and Z. Griliches, 1984, Econometric models for count data with an application to the patents – R&D relationship, *Econometrica*, vol. 52, pp. 909–938.

Heath, A., K. Bucholz, P. Madden, S. Dinwiddie, W. Slutske, L. Bierut, D. Statham, M. Dunne, J. Whitfield and N. Martin, 1997, Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency in findings for women and men, *Psychological Medicine*, vol. 27, pp. 1381–1396.

Häkkinen, U., M.R. Järvelin, G. Rosenqvist and J. Laitinen, 2006, Health, schooling and lifestyle among young adults in Finland, *Health Economics*, vol. 15, pp. 1201–1216.

Isacsson, G., 1999, Estimates of the return to schooling in Sweden from a large sample of twins, *Labour Economics*, vol. 6, pp. 471–489.

Kenkel, D., 1991, Health behavior, health knowledge, and schooling, *Journal of Political Economy*, vol. 99, no. 2, pp. 287–305.

Knopik, V., A. Heath, P. Madden, K. Bucholz, W. Slutske, E. Nelson, D. Statham, J. Whitfield and N. Martin, 2004, Genetic effects on alcohol dependence risk : re-evaluating the importance of psychiatric and other heritable risk factors, *Psychological Medicine*, vol. 34, pp. 1519–1530.

Lopez, A., C. Mathers, M. Ezzati, D. Jamison and C. Murray, eds., 2006, *Global Burden of Disease and Risk Factors*, Oxford University Press and The World Bank.

Mayfield, D., G. McLeod and P. Hall, 1974, The cage questionnaire: Validation of a new alcoholism screening instrument, *American Journal of Psychiatry*, vol. 131, pp. 1121–1123.

Miller, P., C. Mulvey and N. Martin, 1995, What do twins studies reveal about the economic returns to education? A comparison of Australian and U.S. findings, *American Economic Review*, vol. 85, no. 3, pp. 586–599.

Miller, P., C. Mulvey and N. Martin, 2001, Genetic and environmental contributions to educational attainment in Australia, *Economics of Education Review*, vol. 20, p. 211Ű224.

Oers, J. van, I. Bongers, L. van de Goor and H. Garretsen, 1999, Alcohol consumption, alcohol-related problems, problem drinking, and socioeconomic status, *Alcohol and Alcoholism*, vol. 34, no. 1, pp. 78–88.

Park, C. and C. Kang, 2008, Does education induce healthy lifestyle?, *Journal of Health Economics*, vol. 27, pp. 1516–1531.

Seppä, K., J. Lepistö and P. Sillanaukee, 1998, Five-shot questionnaire on heavy drinking, *Alcoholism: Clinical and Experimental Research*, vol. 22, no. 8, pp. 1788–1791.

Sharpe, D., M. Abdel-Ghany, H.Y. Kim and G.S. Hong, 2001, Alcohol consumption decisions in Korea, *Journal of Family and Economic Issues*, vol. 22, no. 1, pp. 7–24.

Slutske, W., A. Heath, P. Madden, K. Bucholz, D.J. Statham and N. Martin, 2002, Personality and the genetic risk for alcohol dependence, *Journal of Abnormal Psychology*, vol. 111, no. 1, pp. 124–133.

Su, S.J.B. and S. Yen, 2000, A censored system of cigarette and alcohol consumption, *Applied Economics*, vol. 32, no. 6, pp. 729–737.

Webbink, H., N. Martin and P. Visscher, 2008a, Does education reduce the probability of being overweight?, CPB Discussion Paper 102.

Webbink, H., N. Martin and P. Visscher, 2008b, Does teenage childbearing increase smoking, drinking and body size?, *Journal of Health Economics*, vol. 27, no. 4, pp. 888–903.

WHO, 2007, Expert committee on problems related to alcohol consumption, WHO Technical Report Series 944, World Health Organization.

Wooldridge, J.M., 1999, Distribution-free estimation of some nonlinear panel data models, *Journal of Econometrics*, vol. 90, pp. 77–97.

Yen, S., 1995, Alternative transformations in a class of limited dependent variable models: alcohol consumption by us women, *Applied Economics Letters*, vol. 2, pp. 258–262.