

# Peer Effects in UK Adolescent Substance Use: Never Mind the Classmates?

Duncan McVicar\*  
Melbourne Institute of Applied Economic and Social Research  
University of Melbourne  
Vic 3010, Australia  
[dmcvicar@unimelb.edu.au](mailto:dmcvicar@unimelb.edu.au)  
Tel: +61(0)3 90354509

&

Arnold Polanski  
Queen's University Management School  
Queen's University Belfast  
University Road  
Belfast BT7 1NN  
United Kingdom  
Tel: +44(0)2890975024  
[a.polanski@qub.ac.uk](mailto:a.polanski@qub.ac.uk)

This version: November 2010

JEL: I12, Z13

Keywords: peer effects, alcohol, school class, friends

Acknowledgements: Thanks to the ESPAD team for the data. The United Kingdom ESPAD study was conducted by Professor Martin Plant and Dr Patrick Miller of the University of the West of England, Bristol (UWE). It was mainly funded by the Wates Foundation and UWE. Additional support was provided by the Joseph Rowntree Foundation, the Oakdale Trust, Butcombe Brewery Ltd, Dr George Carey, the Jack Goldhill Charitable Trust, R& J Lass Charities Ltd and the North British Distillery Company Ltd. Thanks also to seminar participants at Queen's University Belfast, the Melbourne Institute of Applied Economic and Social Research, the Irish Economic Association Conference 2009, the European Meeting of the Econometric Society 2009, the European Association of Labour Economists Conference 2009 and the Australian Conference of Economists 2010 for useful comments on earlier drafts.

\*Corresponding author

### **Abstract**

This paper estimates peer influences on the alcohol, tobacco and cannabis use of a school based sample of UK 15 year olds. We present evidence of large, positive and statistically significant peer effects in all three behaviours when classmates are taken as the reference group. When friends are taken as the reference group, using self-reports of perceived friends' substance use, we also find large, positive and statistically significant associations with own substance use, although we stop short of inferring causality in this case. When both reference groups are considered simultaneously, the influence of classmates on own behaviour either disappears or is much reduced in magnitude, whereas the association between own and friends' behaviours doesn't change. The suggestion is that classmate behaviour is primarily relevant only inasmuch as it proxies for friends' behaviour, with classmates that are not also friends having relatively little influence on adolescent substance use.

## 1. Introduction

Peer effects (or social interactions effects) refer to externalities in which the actions or characteristics of a reference group affect an individual's behaviour. Such effects are possible in a huge variety of social contexts, with recent studies exploring areas including educational performance (e.g. Hoxby, 2000; Angrist and Lang, 2004); criminal behaviour (e.g. Glaeser et al., 1996; Kling et al., 2005; Pattachini and Zenou, 2008); smoking and drinking (e.g. Kremer and Levy, 2008; DeSimone, 2007; Cutler and Glaeser, 2008); benefits claiming (e.g. Bertrand et al., 2000); and neighbourhood influences on various other socio-economic outcomes (e.g. Evans et al., 1992; Topa, 2001; Ioannides and Zabel, 2003). But the literature is yet to arrive at a consensus about the existence of such effects, let alone their magnitudes. Brock and Durlauf (2001), Soetevent (2006) and Ioannides (2008) provide recent reviews.

This lack of evidence-based consensus is also true for the specific literature concerning peer effects in adolescent substance use – an issue of acute policy interest – despite the conventional wisdom that they are a key factor influencing individual behaviour. Eisenberg (2004, p1) puts it as follows: “For adolescents in particular, it is almost accepted as fact that peers exert strong effects on each other's substance use behaviour. Yet in practice it is extremely difficult to perform an empirical test that verifies these effects.” The difficulty arises from a combination of tough data requirements – at the very least information is required on own behaviour and behaviour of a relevant reference group – and the potential endogeneity of peer behaviour (see Manski, 1993; Brock and Durlauf, 2001). Many studies have addressed these problems by exploiting readily available school-based survey data together with an instrumental variables (IV) approach. This reliance on school-based survey data together with concerns over non-random selection of friends, however, means that much of the evidence on peer effects in adolescent substance use treats the school or school class as the reference group. But for behaviours that mostly take place outside of school hours, such as cannabis use, it is questionable whether these are the most relevant reference groups.<sup>1</sup>

---

<sup>1</sup> Questions about which peers have the most influence for which behaviours, although often overlooked in the empirical economics literature on peer effects, are not new to the social sciences (see, for example, Granovetter (1973) and Oetting and Beauvais (1987)).

In this paper we exploit UK data on substance use among classmates and among friends to provide new estimates of peer effects in the UK, and further to show that friends are the more relevant reference group for these behaviours. Because most friends are drawn from the class but not all classmates are friends, classmates' behaviour is relevant to own behaviour primarily as a proxy for friends' behaviour, at least for smoking and cannabis use. Estimates of peer effects between classmates are therefore likely to underestimate peer effects between friends, and zero peer effects between classmates does not imply zero peer effects between friends.

The remainder of the paper is set out as follows. The following section briefly discusses the different kinds of data available and the associated econometric problems in estimating peer effects. Section 3 summarises existing estimates of peer effects in adolescent substance use for different reference groups. Section 4 provides an initial description of our data. Section 5 sets out our estimation approach. Section 6 presents and discusses our own estimates of peer effects where both the school class and friends are treated as reference groups. Section 7 concludes.

## **2. Identifying Peer Effects with Different Kinds of Data**

As is well known, the existence of a statistical association between peer behaviour and own behaviour does not necessarily imply a *causal* relationship from peer behaviour to own behaviour (see Manski, 1993; Brock and Durlauf, 2001). There are a number of potential identification issues, including unobserved confounding factors, non-random selection of peers, measurement error in peer behaviour, systematic misreporting of peer behaviour, and, at least in linear models, the difficulty in separating the impacts of peer characteristics and of peer behaviour on own behaviour. These problems are variously important, and suggest different estimation strategies, in different kinds of data covering different reference groups. On top of these identification issues is the issue of which reference groups are relevant for which kinds of behaviours. For studies of how a child's educational performance is influenced by the performance of classmates at school, for example, there is no issue here since by definition the relevant peer group is the school class. But for studies of peer effects in adolescent substance use, are social influences likely to come from friends or classmates or both?

Three different kinds of survey data have generally been used in previous econometric studies of peer influences in adolescent substance use. Common to all is that they are based on samples of young people either across local neighbourhoods or schools with information collected on own substance use and own characteristics. But data sources differ in the extent to which friendship links are observed, in the definition of the relevant reference group, and in the amount of information available about the reference group.

The ‘gold standard’ in the context of adolescent substance use is network data where friendship links between different individuals in the sample are explicitly observed, i.e. we know person  $i$  is connected to person  $j$  and so on. The US National Longitudinal Survey of Adolescent Health (Add Health) study, which identifies up to ten friendship links between individual pupils within schools, is an example. An immediate advantage is that such data allow both classmates and friends to be examined as potential reference groups. Eisenberg (2004) and Clark and Loheac (2007) do so, although both studies explore impacts of each reference group separately. Another advantage is that friendship reference groups are likely to vary across individuals in the same school or school class in which case Manski’s (1993) ‘reflection problem’ doesn’t arise. On the other hand, such data are just as susceptible to shared confounding factors as other school-based survey data. Further, endogenous selection of peers is likely to be more of a problem where friends are taken as the reference group than where schoolmates or classmates are taken as the reference group (Clark and Loheac, 2007). Friendship links are also only observed for friends within schools which may lead to measurement error in peer behaviour for those with friends outside school. Fortunately a variety of methods to arrive at unbiased estimates of peer effects are potentially available with such data, including school fixed effects, IV, and using lags of peer behaviour in place of contemporaneous values where data are collected longitudinally as in the case of the Add Health study. Unfortunately such data are still relatively uncommon.

More commonly used in this context are survey based data where no friendship links are observed but where the school class or the neighbourhood is taken as the relevant reference group, depending on the sampling frame. Case and Katz (1991) is an example of a study using neighbourhood survey data, with city blocks taken as the relevant reference group. Gaviria and Raphael (2001), Powel et al. (2005) and Soetevent and Kooreman (2007) are examples of studies with the school or school class taken as the reference group. An advantage of neighbourhood

data over school data is that substance use is more likely to take place in the neighbourhood context rather than the school context. On the other hand city blocks are not as well defined as schools or school classes (Gaviria and Raphael, 2001) and school-based surveys stand a better chance of obtaining information on the reference group population (the whole school or school class) as opposed to a sample of the relevant reference group population.<sup>2</sup> Although selection bias is likely to be less of a problem with these data than for estimates of peer effects between friends, it is possible that families self-select into neighbourhoods (and therefore self-select their children into neighbourhood schools) on the basis of some characteristics that are correlated with adolescent substance use, and other confounding factors are just as likely, so peer behaviour is again potentially endogenous. Linear models may also be susceptible to the reflection problem so that the impact of peer behaviour cannot be separately identified from the impact of peer characteristics. Most existing studies have made progress by assuming no contextual effects – that peer characteristics do not directly influence an individual’s behaviour – and adopting an IV strategy where peer characteristics instrument for peer behaviour.

A third kind of survey data is where each individual reports their perception of reference group substance use, whether among friends (e.g. Norton et al., 2003) or among classmates (Norton et al., 2003 again; Kawaguchi, 2004). On the one hand, it may be perceived behaviour rather than actual behaviour that matters for peer effects. Data on perceived friend’s behaviour can also, at least in principal, capture behaviour of friends both inside and outside of any given school or neighbourhood, unlike school-based network data such as provided by the Add Health study. On the other hand, individuals might systematically misreport peer behaviour (Norton et al., 2003), on top of the usual potential biases due to selection of peers and other confounding effects. There may also be a lack of information on peer characteristics other than the reported behaviours – at least this is the case in the perceptions data used in the current paper – which makes an IV approach difficult.

In this paper we use a rich data set – the UK part of the 2003 European Schools Survey Project on Alcohol and Other Drugs (ESPAD) – that combines elements of the second and third types of data. As a school-based survey we have information on actual behaviour of classmates and on

---

<sup>2</sup> Micklewright et al. (2010) show that measurement error in peer variables can lead to estimates of peer effects that are biased towards zero where only a random sample of peers is observed for each individual.

classmates' characteristics, allowing a standard IV approach to the estimation of peer effects in substance use where the school class is taken as the relevant reference group. But because the data also include information on the perceived behaviour of each individual's friends, we can also estimate peer effects, or at least associations between own and friends' substance use, from this alternative reference group, both separately and simultaneously with classmates.

### 3. Existing Estimates of Peer Effects in Adolescent Substance Use

Several recent studies provide estimates of peer effects in adolescent substance, using school-based survey data, where the school, the school year or the school class are treated as the reference group. These studies usually focus on the prevalence of a particular behaviour in the reference group and how it impacts on the decision of an individual to *participate* in the behaviour, often using data on whether the individual has, say, smoked at least one cigarette in the last 30 days. Most provide estimates of peer effects where peer behaviour is treated as exogenous together with estimates where it is treated as endogenous. Despite their similarities, however, the studies draw mixed conclusions on the magnitude and even the existence of peer effects in adolescent substance use.

For the US, Eisenberg (2004) finds zero peer effects in either smoking, drinking or cannabis use where the school is treated as the reference group and school grade structure is used to instrument for peer behaviour. Clark and Loheac (2007) find positive peer effects where the school year is treated as the relevant reference group and where lagged peer behaviours are included in place of current peer behaviours. These are larger for boys than for girls, with elasticities of around one quarter (boys drinking), one sixth (girls drinking), one fifth (boys and girls smoking), one quarter (boys cannabis use) and one sixth (girls cannabis use). But when school fixed effects are included the only peer effect that remains statistically significant at 95% is for boys drinking (elasticity around one fifth). In contrast, Gaviria and Raphael (2001) find positive and significant peer effects in drinking, smoking and drug use (cannabis and cocaine treated together), where the school year is taken as the reference group, whether they instrument for peer behaviour (with peer family background) or not. Their single equation estimates suggest elasticities of around one sixth (smoking), one fifth (drinking) and one quarter (drug use), and their IV estimates suggest elasticities of around one sixth (smoking) and one third (drinking and

drug use). Powell et al. (2005) use a similar approach to look at peer effects in smoking with the school class treated as the reference group, finding an elasticity of own smoking to peer smoking of around one half whether peer behaviour is instrumented or not. Kawaguchi (2004) finds positive and significant peer effects in drinking, smoking and cannabis use, where the school year is taken as the reference group, but using perceived peer behaviour in place of actual peer behaviour. Elasticities are around one third (drinking) and one half (smoking and cannabis use) whether school fixed effects are included or not. Norton et al. (2003) provides estimates of peer effects where schoolmates in the same year are treated as the reference group and where both actual and perceived peer behaviours are observed. Elasticities where actual peer behaviours are used are around two fifths (smoking) and one third (drinking). Corresponding elasticities where perceived peer behaviours are used are around one fifth (smoking) and one third (drinking). Outside of the US, Soetevent and Kooremen (2007) find evidence of peer effects in adolescent smoking for the Netherlands where the school class is taken as the reference group, but they disappear when school dummies are included to capture unobservables.<sup>3</sup> We didn't find any estimates of school or class based peer effects in adolescent substance use for the UK.

Now consider studies that define friends, or some subset of friends, as the relevant reference group. Once again, despite their similarities, these studies draw mixed conclusions on the magnitude of peer effects in adolescent substance use, although unlike the schoolmate/classmate reference group studies, all find evidence of non-zero peer effects. For the US, Clark and Loheac (2007) use information on who is a friend with who in the Add Health data to estimate peer effects between school-based friends in drinking, smoking and cannabis use, using lagged peer behaviour and school fixed effects to aid identification. They find positive and significant peer effects in each case for both male and female peer groups, with elasticities of around one eighth (boys and girls drinking), one fifth (boys and girls smoking), one eighth (boys cannabis use) and one tenth (girls cannabis use). Eisenberg (2004) also uses Add Health data to estimate peer effects amongst school-based friends but using something more akin to a natural experiment approach, where he explores the impact on own substance use of have a 'using' friend leave the

---

<sup>3</sup> Soetevent and Kooreman (2007) argue that this school-based reference group may be less appropriate for some behaviours than for others. For example, in the absence of school fixed effects, they find stronger peer effects for more 'school-based' behaviours such as truancy and weaker peer effects for less 'school-based' behaviours such as smoking. Soetevent (2006) also includes a discussion of reference group definition.

school. He obtains elasticities of around one twentieth (drinking and smoking) and one eighth (cannabis use). Norton et al. (2003) use data on actual and perceived smoking for three ‘best friends’ to provide two alternative sets of estimates of peer effects, finding elasticities of around one half (actual smoking) and three quarters (perceived smoking). Krauth (2007) estimates peer effects in adolescent smoking using perceived (same sex) friends’ smoking and providing both ‘naive’ estimates (treating perceived peer smoking as exogenous) and estimates that account for selection assuming peer group selection on unobservables is of equal magnitude to peer group selection on observables. Estimated elasticities are around two thirds and one twentieth respectively. Krauth (2005) provides similar estimates using Canadian data on perceived smoking of friends, but finds larger peer effects with elasticities of around one (naive model) and one half (model accounting for selection). Finally, Miller and Plant (2003) study peer effects in adolescent alcohol, tobacco and drug use using an earlier sweep of the same UK survey data we use here, treating friends as the reference group and again using data on perceived substance use. They provide only single equation estimates (treating peer behaviours as exogenous) and their results are not reported in the form of impacts on participation probabilities and so are not readily comparable with the other estimates discussed here. Nevertheless, they show strong, positive associations between own substance use and perceived substance use among friends for all three behaviours.

Of these studies, only Eisenberg (2004) and Clark and Loheac (2007) provide estimates of peer effects for different reference groups – friends and the school year – using the same data set. In both studies, at least for smoking and cannabis use, it is only friends that appear relevant. Clark and Loheac (2007) put this down to a combination of larger peer effects for friends than for classmates and positive bias due to endogenous selection of friends. For alcohol use, Clark and Loheac’s (2007) estimates suggest both reference groups are relevant, although Eisenberg’s (2004) estimates again suggest friends are relevant where classmates are not. Because peer effects are estimated separately for each reference group, however, we cannot rule out the possibility that classmates are merely proxying for friends in this case. If we want to test whether friends and classmates are both relevant reference groups then we need to estimate a model where substance use of both reference groups is included simultaneously.

#### 4. The ESPAD Data

The data used here are drawn from the UK part of the 2003 sweep of ESPAD survey which randomly selected a representative sample of schools across the country from which single classes of pupils in their final year of compulsory education, i.e. aged 15 or 16 years, were selected at random. All (present) members of each class were given the ESPAD questionnaire to complete, anonymously, during school hours, and under exam conditions. The completed forms were then sealed in envelopes, collected by staff, and posted to the survey team. After cleaning a small number of spoiled forms, the resulting database contains self-reported information on over 2000 individuals across 77 schools.

The questionnaires ask a number of questions relating to own smoking, drinking and drug taking behaviour. We focus on the following questions:

- How frequently have you smoked cigarettes during the last 30 days?
- On how many occasions in the last 30 days (if any) have you had any of the following to drink (beer, wine, spirits, also include sprits mixed with soft drinks)?
- On how many occasions (if any) have you used cannabis (marijuana, dope, grass, pot) or hashish (hash or hash oil) during the last 30 days?

Each of these questions has a seven-point scales for responses, ranging from ‘none/not at all’ through to ‘more than 20 cigarettes/40 or more drinks/40 or more occasions using cannabis’. For comparison with much of the existing literature, we dichotomise the resulting data into binary indicators where 0 denotes ‘none/not at all’ and 1 denotes at least some smoking tobacco, consumption of alcohol, or use of cannabis. By summing across the school class (other than the individual concerned) we construct variables for the prevalence of these behaviours in the class defined as the proportions of the class that have smoked, drank alcohol or used cannabis in the last 30 days. Table 1 shows the prevalence of these behaviours across the sample. Note that alcohol use is very much a mainstream behaviour for this age group, whereas tobacco and cannabis use are minority behaviours, albeit substantial minorities in each case.

Crucially for our purposes the questionnaire also asks each individual to report information relating to the behaviour of their friends as well as themselves. In other words we can separately

identify the behaviour of each individual's friends from that of each individual's class. Note that because the survey is entirely anonymous we cannot identify who in the class is regarded by who as a friend, or even whether an individual's friends are in or outside his or her class at school. This also means we have no information on the *characteristics* of friends. The friends' behaviour questions take the following form:

- How many of your friends would you estimate...
  - Smoke cigarettes?
  - Drink alcoholic beverages (beer, wine, spirits)?
  - Smoke cannabis (dope, grass, pot, marijuana) or hashish?

Responses are given on a five point scale labelled 'none', 'a few', 'some', 'most', and 'all'. In what follows we use both the raw five point scale to measure friends' behaviours as well as simple dichotomous indicators for having any friends that smoke or use cannabis or having at least some friends that drink.<sup>4</sup> Table 1 shows the sample means for these five point scales.

Note that the questions about own substance use refer to the last 30 days whereas those on perceived friends' substance use do not specify a time period. We proceed on the assumption that this distinction is not critical – both questions can be interpreted as referring to 'current' use. A more important distinction, however, is that our substance use measure for classmates is expressed as a (quasi-)continuous proportion whereas our perceived substance use measure for friends is expressed as a proportion on a five point scale. Although we estimate models including behaviours for both reference groups in these different forms, to facilitate comparisons between them we construct similar five point scale measures for the proportions of classmates engaged in each behaviour. To do so we take the distributions of individuals across the whole sample in each of the five 'friend-use' categories and apply these shares to split the quasi-continuous classmate proportions into five similarly-distributed categories. Table 1 shows means and standard deviations for these constructed scales are very close to those for friends. As we might expect, given that many friends are likely to come from within the school class, these variables are positively correlated, with  $\rho=.19$  (drinking),  $\rho=.14$  (smoking) and  $\rho=.13$  (cannabis use).

---

<sup>4</sup> Only two percent of the sample report having no friends that drink.

Individuals also report on a wide range of questions about their own characteristics, attitudes and family backgrounds. From these we select a small number of individual characteristics and family background controls which will plausibly impact on own substance abuse behaviour: a binary dummy for whether the individual has one or more older siblings engaging in the relevant behaviour; a binary dummy for gender (male=1, 0 otherwise); a binary dummy for father with higher education; a binary dummy for describing yourself as ‘academically able’ derived from a seven point scale asking ‘how good you think you are at schoolwork compared to other people your age’ (1= excellent, well above average, above average, 0 otherwise); a binary dummy for living in a two-parent household (including step-parents); and a variable for ‘money usually spent in a week for personal needs without parents’ control’ denoted in £s. Our priors for these control variables are uncontroversial: we expect those with higher educated fathers, those living in two-parent households, those with better academic ability, and those with less disposable income to engage in less substance use. Table 1 presents sample means.

## 5. Approach to Estimation

This paper extends the approach of Gaviria and Raphael (2001), Powell et al. (2005) and others by including both reference groups in a simple reduced form regression for substance use behaviour as given below:

$$y_i^* = \alpha + \delta_1 \bar{y}_{1i} + \delta_2 \bar{y}_{2i} + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad (1)$$

where  $y_i^*$  is a latent variable denoting the extent of substance use in the last 30 days for individual  $i$ ,  $\bar{y}_{1i}$  is the prevalence of substance use among  $i$ 's classmates,  $\bar{y}_{2i}$  is the prevalence of (perceived) substance use among  $i$ 's friends,  $X_{1i}$  to  $X_{ki}$  are the set of control variables discussed in the previous section and  $u_i$  is a normally distributed error term. Following these earlier studies, (1) makes the assumption that peer characteristics (e.g. their family background) have no *direct* impact on own behaviour, although they may have an *indirect* impact via peer substance use. We do not actually observe  $y_i^*$  but the following dichotomous indicator for substance use:

$$y_i = \begin{cases} 0 & \text{if } y_i^* \leq 0 \\ 1 & \text{if } y_i^* \geq 0 \end{cases} \quad (2)$$

We therefore estimate (1) and (2) by maximum likelihood as a series of probit models, initially treating peer behaviour as exogenous (although allowing for correlation of errors within classes).

We explore a number of different specifications of the reference group behaviour variables. First we assume  $\delta_2=0$ , i.e. that only classmates influence own behaviour, using the continuous measures of classmate substance use prevalence. Results are presented in Table 2 columns 1, 3 and 5, and essentially they provide UK estimates of peer effects along the lines of the (single equation) estimates of Gaviria and Raphael (2001) and Powell et al. (2005) for the US. Second, we assume  $\delta_1=0$ , i.e. that only friends influence own behaviour, using simple dichotomous measures of perceived prevalence of substance use amongst friends. Results are presented in Table 3 columns 1, 3 and 5, and essentially provide UK estimates of peer effects along the lines of Norton et al. (2003) and (the naive estimators of) Krauth (2005, 2007). Results where school fixed effects are included in (1) to capture common unobserved influences on substance use are presented in Table 3 columns 2, 4 and 6. Third, we replace the continuous measure of classmate substance use with the constructed five point scale measure, again assuming  $\delta_2=0$ . Results are presented in column 1 of Tables 4-6. Fourth we similarly replace the dichotomous indicator of friends perceived substance use with the five point scale measure, again assuming  $\delta_1=0$ . Results are presented in column 2 of Tables 4-6. Fifth, we include peer effects from both classmates and friends simultaneously, where substance use behaviours for both reference groups are measured on the five point scales. Results are presented in column 3 of Tables 4-6. These are the key results for the purposes of distinguishing classmate impacts from friend impacts.

Given the identification concerns discussed earlier, we then allow peer behaviour to be endogenous, and re-estimate (1) and (2) as IV probit models (see StataCorp, 2009), with classmate family background measures instrumenting for classmate substance use prevalence. This follows the approach of earlier studies such as Gaviria and Raphael (2001) and Powell et al. (2005) and results are presented in Table 2 columns 2, 4 and 6. Unfortunately, the same IV strategy is not possible where friends are taken as the reference group because we do not observe any background information for friends. The estimates of peer effects from friends will therefore combine actual peer effects with a mix of biases due to confounding effects, selection, random measurement error and systematic misreporting of peer behaviour, for which it is not possible to

assign an overall direction. We are therefore reluctant to place a causal interpretation on the friends estimates.

Finally, to test robustness, we re-estimate the models with both reference groups included simultaneously using the continuous prevalence measure for classmate behaviour, and using binary dummies for the categories on the five point scales of classmate and friends behaviours. We do not report these results separately given their similarity with the results we do report, although we do refer to them in the discussion where appropriate. We also re-estimate all models as linear probability models and using two stage least squares, although again we do not report the results given their similarity with the probit results.

## **6. Results and Discussion**

Before discussing our peer effects estimates it is worth briefly considering the estimated impacts on substance use of the control variables. These are robust across the various versions of the model and take expected signs. For alcohol use, having an older sibling that drinks increases own drinking probability by between .11 and .14; an extra ten pounds of own spending money per week increases own drinking probability by between .04 and .06; and own drinking probability is unrelated to gender, academic ability, living in a two parent household and father's education level. For smoking, having an older sibling that smokes increases own smoking probability by between .17 and .23; being a boy and being 'academically able' reduces smoking probability by up to .08 and .17 respectively; living in a two parent household and having a higher educated father reduces smoking probability by between .09 and .13, and between .05 and .09 respectively; an extra ten pounds of own spending money per week increases own smoking probability by up to .04. For cannabis use, having an older sibling that uses cannabis increases own cannabis use probability by between .09 and .23; being a boy increases cannabis use probability by between .06 and .09; being 'academically able' reduces cannabis use by up to .05 but is at best marginally significant; living in a two parent household and having a higher educated father reduces smoking cannabis use probability by between up to .07 and .05 respectively; an extra ten pounds of own spending money per week increases own smoking probability by up to .04.

## 6.1. School Class As Reference Group

First consider the single equation probit estimates for alcohol use in Table 2 column 1. There is a large, positive and highly statistically significant association between own drinking and prevalence of alcohol use among classmates, with a one percentage point increase in the proportion of the class that drink associated with a .6 percentage point increase in the probability that the individual drinks.<sup>5</sup> The point estimate falls slightly to .55 when we instrument for peer drinking (column 2 of Table 2). Weak instrument caveats aside, the similarity of the single equation probit and IV probit estimates, together with the Wald test reported in Table 2, suggests we cannot rule out exogeneity of classmate drinking prevalence, as suggested by Gaviria and Raphael (2001) for the US (in which case the single equation probit estimates are to be preferred). Notice that our peer effect estimate for the UK is larger than the range suggested for the US in Section 3.

The single equation probit estimates for tobacco use are presented in column 3 of Table 2. Again there is a large, positive and highly statistically significant association between own smoking and prevalence of tobacco use among classmates, with a one percentage point increase in the proportion of the class that smoke associated with a .37 percentage point increase in the probability that the individual smokes. The point estimate increases to .47 when we instrument for peer drinking (column 4 of Table 2) but again we cannot rule out exogeneity of classmate smoking prevalence in line with Gaviria and Raphael (2001) and Powell et al. (2005) for the US. The magnitude of this peer effect is close to that estimated for the US by Powell et al. (2005), Kawaguchi (2004) and Norton et al. (2003).

Next consider the single equation probit estimates for cannabis use presented in column 5 of Table 2. Once again there is a large, positive and highly statistically significant association between own cannabis use and prevalence of cannabis use among classmates, with a one percentage point increase in the proportion of the class that use cannabis associated with a .28 percentage point increase in the probability that the individual uses cannabis. The point estimate more than doubles, to .61, when we instrument for peer cannabis use (column 6 of Table 2), and in this case the evidence suggests we cannot treat peer cannabis use as exogenous. Gaviria and

---

<sup>5</sup> Marginal effects can be interpreted as elasticities where the peer behaviour variable is expressed as a class proportion.

Raphael (2001) also obtain an IV estimate that is larger than their single equation estimate – they suggest this may be driven by negative simultaneity bias – although in both cases our point estimates for the UK are larger in magnitude than their point estimates for the US.

So for all three behaviours we find evidence of strong, positive and highly significant peer effects between classmates in the UK, whether we instrument for peer behaviour or not. In terms of magnitude the peer effect appears larger for alcohol and cannabis use than for tobacco use, which reflects rankings found by Gaviria and Raphael (2001), for example, but not Norton et al. (2003) or Kawaguchi (2004). The estimated magnitudes are in line with earlier US estimates for tobacco smoking but larger than earlier US estimates for alcohol and cannabis use.

## **6.2. Friends as Reference Group**

Now consider the estimates where friends are taken as the reference group, but where we only have information on the perceived substance use of friends. Estimates are presented in Table 3, for models with and without school fixed effects, and where we capture reference group behaviour with binary dummies for whether some friends drink, a few friends smoke or a few friends use cannabis. By including the fixed effects we control for time-invariant unobserved factors that might influence own and friends' behaviours, e.g. the existence of a shop close to the school that sells alcohol or cigarettes to adolescents, which will likely reduce any positive bias on the estimated peer effect coefficients. On the other hand we might introduce additional biases of uncertain sign because of the incidental parameters problem (see Greene, 2001). The similarity between our probit and our (unreported) linear model fixed effects estimates, however, suggests any such biases are not large in this case. Even so, both the simple probit and fixed effects probit estimates may still be subject to a mix of biases of uncertain overall sign.

First consider alcohol use. Adolescents that report at least some friends that use alcohol are 44 percentage points more likely to drink than those that report none or only a few friends that drink. Including school fixed effects reduces the size of this association from .44 to .36, suggesting that the fixed effects do control for some school-level confounding factors that otherwise positively bias the peer effect estimate. These results are consistent in sign and statistical significance with the existing estimates of Miller and Plant (2003) for the UK and

Eisenberg (2004) and Clark and Loheac (2007) for the US, although the magnitudes are not directly comparable. They are also internally consistent in sign and statistical significance with our own Table 2 estimates where the school class is taken as the reference group, but again the two sets of estimates are not directly comparable in terms of magnitude.

Adolescents that report at least a few friends that smoke are 51 percentage points more likely to smoke than those that report no friends that smoke. Including school fixed effects slightly reduces the size of this association from .51 to .47, again suggesting some school-level confounding factors that otherwise positively bias the peer effect estimate. These results are also consistent in sign and statistical significance with our own Table 2 estimates where the school class is taken as the reference group, with the existing estimates of Miller and Plant (2003) for the UK and with the longer list of studies for the US that examine peer effects in smoking, although again magnitudes are not directly comparable. What we don't know is the extent to which we can interpret this statistical association, even with the fixed effects, as a causal peer effect, although we can be guided to some extent by earlier smoking studies. Krauth (2005), for example, suggests selection accounts for around half of the overall association between own smoking and reports of friends' smoking, and if this carried over to our own estimates then we would still see a large and positive association between own and friends' smoking. Norton et al. (2003) suggests the association between own smoking and perceived friends' smoking is around 1.5 times larger than that between own smoking and actual friends' smoking, and again if this carried over to our own case, and we were interested in actual peer smoking rather than perceived peer smoking, we would still be left with a large and positive peer effect in smoking between friends.

Adolescents that report at least a few friends that use cannabis are 35 percentage points more likely to use cannabis than those that report no friends that use cannabis, and this association barely changes when school fixed effects are included. Once again these results are consistent in sign and statistical significance, but not directly comparable in magnitude, with our own estimates where the class is taken as the reference group (Table 2), with the existing estimates of Miller and Plant (2003) for the UK and with those of Eisenberg (2004) and Clark and Loheac (2007) for the US. But again we don't know the extent to which we can interpret this statistical

association, even with the fixed effects, as a peer effect, and as for alcohol use, the existing literature provides less guidance in the case of cannabis use than the case of tobacco smoking.

So for all three behaviours we find evidence of strong, positive and highly significant associations between own substance use and substance use of friends as perceived and reported by the individual. These associations are robust to inclusion of school fixed effects, but may still be subject to numerous biases of uncertain overall sign (and perhaps different signs and/or magnitudes for the different behaviours). Just because we can treat classmate drinking and smoking as exogenous does not imply that we can treat self-reports of perceived friends' substance use as exogenous. On the other hand, if friends are largely drawn from within school classes, the existence of peer effects between classmates does suggest that peer effects also exist between friends, in which case these associations will be picking up causal effects at least to some extent.

### **6.3. Both Reference Groups**

Now consider the estimates presented in Tables 4-6 where, to facilitate comparison between the reference groups, we use five point scales for the proportion of classmates and friends that drink, smoke or use cannabis. Because the classmate substance use variables refer to actual classmate behaviour, we are implicitly treating perceptions of friends' behaviours as proxies for actual friends' behaviours here (for a discussion of the implications of this see Norton et al., 2003). For simplicity we treat the five point scales as if they were cardinal, which assumes the impact on own substance use of moving from 'a few' to 'some' classmates or friends using the substance is the same as the impact of moving from 'some' to 'most'. All the conclusions hold, however, where we replace the five point scales with dummies for individual points on each scale, and our focus is on the sign, significance, stability and relative magnitudes of the estimates rather than their absolute magnitudes. Because we cannot instrument for friends' behaviours following the usual method, and because we cannot include school fixed effects where the class is included as a reference group, we only report single equation probit estimates. As a consequence, although we can proceed on the basis of classmates' drinking and smoking being exogenous, we cannot rule out a downwards bias on the coefficient for classmates' cannabis use and we cannot rule out biases of uncertain sign on the friends' behaviours coefficients. Further, the presence of one or

more endogenous right hand side variables may also bias the coefficients of the exogenous variables, again in an uncertain direction. But even with these caveats, a clear picture emerges of the relative influences of the two reference groups that is unlikely to be driven by biases alone.

First consider alcohol use and Table 4 which presents the estimated marginal effects first for the model where only classmates are included as a reference group, then for the model where only friends are included as the reference group, and finally where both classmates and friends are included as reference groups. In the first case we again see a large, positive and highly statistically significant impact of classmates' drinking on own probability of drinking, with a 'one category' move up the scale increasing probability of own drinking by 6.4 percentage points. In the second case we see a positive and highly significant impact (or at least association) of friends' drinking on own probability of drinking, but in magnitude around three times as large as that from the classmates. Of course this could reflect either a larger causal effect from friends, positive bias, or some combination of the two. In the third case, where we include both reference groups, the association between classmates' drinking and own drinking disappears whereas the association with friends' drinking barely changes. The implication is that once you control for friends' drinking, classmates' drinking has no additional influence on own drinking. In other words, friends are the relevant reference group for adolescent alcohol use, and classmates that are not friends do not appear to be relevant for own alcohol use.

Next consider tobacco use and Table 5. Where only classmates are included as a reference group we again see a positive impact of classmates' smoking on own probability of smoking, with a 'one category' move up the scale increasing probability of own smoking by 4.1 percentage points, but this association is only statistically significant at the 90% level. Where only friends are included as a reference group the marginal effect is much larger and is statistically significant at the 99% level. Where both reference groups are included the association between classmates' smoking and own smoking disappears whereas the association with friends' smoking barely changes. So, as for drinking, the implication is that once you control for friends' smoking, classmates' smoking has no additional influence on own smoking.

Finally consider cannabis use and Table 6. Where only classmates are included as a reference group once again there is a marginally significant positive impact of classmates' cannabis use on

own probability of cannabis use, with a ‘one category’ move up the scale increasing probability of own cannabis use by 2.4 percentage points. Bear in mind, however, that the equivalent estimate in Table 2 was shown to be downwards biased, with an IV estimate double in magnitude. Where only friends are included as a reference group the marginal effect is once again much larger and is statistically significant at the 99% level. Where both reference groups are included the association between classmates’ cannabis use and own cannabis use disappears whereas the association with friends’ cannabis use barely changes. So, as was the case for drinking and smoking, the implication is that once you control for friends’ cannabis use, classmates’ cannabis use has no additional influence on own cannabis use.

Are there plausible alternative explanations – other than that friends are the relevant reference group for these behaviours and classmates are not – for the pattern of results presented in Tables 4-6? Is it possible, for example, that there are no peer effects here at all and that all we are picking up is a positive association between own and friends’ reported behaviours due to, say, selection? Given the evidence in Table 2 suggesting that classmates’ drinking and smoking can be treated as exogenous, and therefore that the large, positive association with own drinking and smoking can be interpreted as a casual effect, this seems unlikely. So although part of the marginal effect on friends’ substance use in each case may be picking up positive biases, it is unlikely that *all* of the marginal effect in each case is driven by positive biases.<sup>6</sup> It also seems unlikely that the inclusion of potentially endogenous friends’ behaviour variables biases all classmates’ behaviour coefficients to zero, especially when control variables generally retain statistical significance.

It does seem possible, however, that we introduce measurement error when using the constructed five point scales for classmates’ substance use, which could bias the marginal effects towards zero. But this bias would be present whether friends’ substance use behaviour was included in the model or not, so is unlikely to explain the contrast between estimated peer effects from

---

<sup>6</sup> Arguably, if we are prepared to explicitly treat classmates’ behaviour as proxying for friends’ behaviour, then we may be able to recover the causal impact of friends’ behaviour on own behaviour by using classmates’ behaviours to instrument for friends’ behaviours. This method appears feasible with our data, at least in the case of smoking and cannabis use, with the resulting IV estimates suggesting peer effects between friends that are around 1.5 and 1.25 times as large as the single equation estimates in Table 3, respectively.

classmates when friends' behaviours are included in the models and when they are omitted. Nevertheless, to check robustness to this we re-estimate the models where both friends and classmates are included simultaneously as reference groups, but with prevalence of substance use among classmates captured by the original continuous proportion measures. For all three behaviours, the marginal effects on friends' use are again unchanged from those presented in Tables 4-6. There is some muddying of the waters, however, in terms of classmates' impacts, which leads us to tone down our earlier conclusion to some extent. For cannabis use, we once again obtain a zero marginal effect, but recall that we cannot rule out downwards bias in this estimate, regardless of how classmates' cannabis use is measured. For tobacco use, we obtain a positive marginal effect that is statistically significant only at the 90% level, and around half the magnitude of the equivalent marginal effect in Table 2. For alcohol use, we obtain a positive and clearly statistically significant impact of classmates' drinking on own drinking, although also around half the magnitude of the equivalent marginal effect in Table 2. So for alcohol and tobacco use, and perhaps for cannabis use, it seems we cannot entirely rule out that classmates have an impact over and above that of friends, although we can show that this impact is much reduced in magnitude by the inclusion of friends in the model.

How do these results fit with the existing empirical literature? Recall that both Eisenberg (2004) and Clark and Loheac (2007) provide estimates of peer effects for both reference groups using Add Health data. Both studies suggest that friends are more relevant for smoking and cannabis use, as is the case here. For alcohol use, Clark and Loheac's (2007) estimates suggest both reference groups may be relevant, whereas Eisenberg's (2004) estimates suggest only friends are relevant. Our results show that alcohol use is the behaviour for which it is hardest to rule out an influence from classmates over and above that from friends.

## **7. Conclusions**

This paper estimates peer influences on the alcohol, tobacco and cannabis use of a sample of UK 15 year olds, contributing new estimates to a literature predominantly using US data. As is the case for many existing studies of peer effects in adolescent substance use, we provide estimates of peer effects where the school class is taken as the relevant reference group. Our findings suggest the existence of large, positive and highly statistically significant peer effects in all three

behaviours, whether peer behaviour is instrumented or not. Further, although we don't observe friendship links explicitly in the data, we are able to use information on friends' perceived behaviours as reported by each individual to estimate peer effects between friends. Again the resulting estimates suggest large, positive and highly statistically significant associations between own substance use and friends' perceived substance use in all three behaviours, although it is difficult to place a purely causal interpretation on these associations.

The main contribution of the paper, however, is to show that the influence of classmates is much diminished, and may even disappear altogether, where both reference groups (friends and classmates) are included in a multivariate regression model for own behaviour. In other words friends are the more relevant reference group for peer effects in substance use, and classmates that are not friends may have little or no influence on adolescent substance use.<sup>7</sup> The implication is that a finding of zero peer effects between classmates, where only classmates are included as a reference group, does not imply zero peer effects between friends or other possible reference groups, in which case we may be closer to an empirical consensus on the existence of peer effects in adolescent substance use than a naive reading of the literature might otherwise suggest.

### References

- Angrist, J. And Lang, K. (2004). 'Does school integration generate peer effects? Evidence from Boston's METCO Program.' *American Economic Review*, 94, 1613-34.
- Bertrand, M., Luttmer, E. And Mullainathan, S. (2000). 'Network effects and welfare cultures.' *The Quarterly Journal of Economics*, 115, 1019-55.
- Brock, W. And Durlauf, S. (2001). Interaction-based models. In *Handbook of Econometrics*, vol. 5, J. Heckman and E. Leamer (eds.). Amsterdam: North-Holland.
- Case, A.C. and Katz, L.F. (1991). 'The company you keep: the effects of family and neighbourhood on disadvantaged youths.' NBER working Paper No. 3705, NBER, Cambridge MA.

---

<sup>7</sup> School-based survey data that include explicit information on friendships, such as the Add Health data, would allow this to be tested more directly by estimating peer effects from classmates that are not identified as friends.

- Clark, A.E. and Loheac, Y. (2007). 'It wasn't me, it was them! Social influence in risky behaviour of adolescents.' *Journal of Health Economics*, 26, 4, 763-84.
- Cutler, D. and Glaeser, E.L. (2008). 'Social interactions and smoking.' NBER Working Paper 13477, NBER, Cambridge MA.
- DeSimone, J.A. (2007). 'Fraternity membership and drinking behaviour.' NBER Working Paper No. 13262, NBER, Cambridge MA.
- Eisenberg, D. (2004). 'Peer effects for adolescent substance use: do they really exist?' Working Paper UC-Berkeley School of Public Health.
- Evans, W.N., Oates, W.E. and Schwab, R.M. (1992). 'Measuring peer group effects: a study of teenage behaviour.' *Journal of Political Economy*, 100, 5, 966-91.
- Gaviria, A. And Raphael, S. (2001). 'School-based peer effects and juvenile behaviour.' *The Review of Economics and Statistics*, 83, 257-68.
- Glaeser, E.L., Sacerdote, B. And Scheinkman, J.A. (1996). 'Crime and social interactions.' *The Quarterly Journal of Economics*, 111, 2, 507-48.
- Granovetter, M. (1973). 'The strength of weak ties.' *American Journal of Sociology*, 78, 1360-80.
- Greene, W.H. (2001). 'Fixed and random effects in nonlinear models.' New York University Working Paper EC -01-01.
- Hoxby, C. (2000). 'Peer effects in the classroom: learning from gender and race variation.' NBER Working Paper No. 7867, NBER, Cambridge MA.
- Ioannides, Y. And Zabel, J. (2003). 'Neighbourhood effects and housing demand.' *Journal of Applied Econometrics*, 18, 563-84.
- Ioannides, Y. (2008). 'Social interactions (empirics),' in *New Palgrave Dictionary of Economics 2<sup>nd</sup> Edition*, S. Durlauf and L. Blume (eds.), Palgrave Macmillan.

- Kawaguchi, D. (2004). 'Peer effects on substance use among American teenagers.' *Journal of Population Economics*, 17, 351-67.
- Kling, J., L.R., Ludwig, J. And Katz, L.F. (2005). 'Neighbourhood effects on *crime for female and male youth: evidence from a randomized housing voucher experiment.*' *The Quarterly Journal of Economics*, 120, 87-130.
- Krauth, B.V. (2005). 'Peer effects and selection effects on smoking among Canadian youth.' *Canadian Journal of Economics* 38, 735-756.
- Krauth, B.V. (2007). 'Peer and selection effects on youth smoking in California.' *Journal of Business and Economic Statistics*, 25, 288-98.
- Kremer, M. and Levy, D. (2008). 'Peer effects and alcohol use among college students.' *Journal of Economic Perspectives*, 22, 189-206.
- Manski, C. (1993). 'Identification of endogenous social effects: the reflection problem.' *Review of Economic Studies*, 60, 531-42.
- Micklewright, J., Schnepf, S.V. and Silva, P.N. (2010). 'Peer effects and measurement error: the impact of sampling variation in school survey data.' Mimeo, Institute of Education, University of London.
- Miller, P. and Plant, M. (2003). 'The family, peer influences and substance use: findings from a study of UK teenagers.' *Journal of Substance Use*, 8, 19-26.
- Norton, E.C., Lindrooth, R.C. and Ennett, S.T. (2003). How measures of perception from survey data lead to inconsistent regression results: evidence from adolescent and peer substance use.' *Health Economics*, 12, 139-48.
- Oetting, E. R., & Beauvais, F. (1987). Peer cluster theory, socialization characteristics, and adolescent drug use: A path analysis. *Journal of Counseling Psychology*, 34(2), 205-213.
- Patacchini, E. And Zenou, Y. (2008). 'The strength of weak ties in crime.' *European Economic Review*, 52, 209-36.

Powell, L.M., John, A. And Hana, R. (2005). 'The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behaviour.' *Journal of Health Economics*, 24, 950-68.

Soetevent, A.R. (2006). 'Empirics of the identification of social interactions: an evaluation of the approaches and their results.' *Journal of Economic Surveys*, 20, 193-228.

Soetevent, A.R. and Kooreman, P. (2007). 'A discrete choice model with social interactions: an application to high school teen behaviour.' *Journal of Applied Econometrics* 22: 599-624.

StataCorp (2009). *Stata Release 11*. Texas: Stata Corporation.

Topa, G. (2001). 'Social interactions, local spillovers and unemployment.' *Review of Economic Studies*, 68, 261-95.

**Table 1: Adolescent Substance Use Sample Means/Proportions (Standard Deviations)**

Individual has used alcohol in last 30 days	.750
Own drinking alcohol scale 1-7	2.79 (1.55)
Proportion of class drinking alcohol in last 30 days: constructed 5 point scale	3.92 (.889)
Proportion of friends drinking alcohol: 5 point scale	3.93 (.885)
Individual has smoked tobacco in last 30 days	.297
Own smoking scale 1-7	1.88 (1.55)
Proportion of class smoking in last 30 days: constructed 5 point scale	2.94 (.959)
Proportion of friends smoking: 5 point scale	2.94 (.954)
Individual has used cannabis in last 30 days	.199
Own cannabis scale 1-7	1.48 (1.22)
Proportion of class using cannabis in last 30 days: constructed 5 point scale	2.38 (1.07)
Proportion of friends using cannabis: 5 point scale	2.40 (1.07)
Older sibling(s) drinking dummy	.551
Older sibling(s) smoking dummy	.268
Older sibling(s) cannabis dummy	.147
Boy	.511
Father HE	.475
Academically able	.573
Two parent household	.679
Own money per week	15.79 (14.42)
Number of schools	76
Nobs	1791

**Table 2: Peer Effects in Alcohol, Tobacco and Cannabis Use, School Class as Reference Group, Probits & IV Probits, Marginal Effects (Standard Errors)**

	Alcohol Probit	Alcohol IV Probit	Tobacco Probit	Tobacco IV Probit	Cannabis Probit	Cannabis IV Probit
Proportion of class using alcohol in last 30 days (%)	.593*** (.072)	.554* (.291)				
Proportion of class using tobacco in last 30 days (%)			.365*** (.121)	.466*** (.133)		
Proportion of class using cannabis in last 30 days (%)					.279** (.116)	.606*** (.136)
Older sibling that drinks	.142*** (.021)	.142*** (.022)				
Older sibling that smokes			.231*** (.029)	.228*** (.029)		
Older sibling that uses cannabis					.231*** (.027)	.221*** (.027)
Boy	-.007 (.020)	-.007 (.021)	-.059* (.032)	-.059* (.031)	.087*** (.027)	.080*** (.025)
Father HE	-.019 (.022)	-.019 (.022)	-.075*** (.025)	-.071*** (.026)	-.048** (.023)	-.040* (.023)
Academically able	-.033 (.024)	-.033 (.025)	-.165*** (.029)	-.162*** (.030)	-.041 (.026)	-.039 (.026)
Two parent household	.019 (.025)	.019 (.025)	-.115*** (.030)	-.110*** (.031)	-.069*** (.026)	-.060** (.026)
Own money per week	.006*** (.001)	.006*** (.001)	.004*** (.001)	.004*** (.001)	.004*** (.001)	.004*** (.001)
Pseudo R <sup>2</sup>	.068		.112		.086	
Nobs	1791	1791	1791	1791	1791	1791
Instrument(s)		a, b		a, c, d, e		a, c
Wald test: exogeneity of peer behaviour (p-value)		.886		.362		.018
F-stat (instrument relevance)		3.65		20.58		14.50
Chi-square p-value (instrument exclusion)		.970		.437		.166

Notes: \*\*\*, \*\* and \* denote statistical significance at 99%, 95% and 90% levels. The dependent variable is whether you drank alcohol, smoked tobacco or used cannabis in the last 30 days. Errors are clustered at classroom level. Marginal effects are calculated at means (peer behaviour, money) or at zero (all binary dummies). For the peer behaviour variables they indicate the difference in the probability of own substance use between an individual in class where no classmates drink/smoke/use cannabis and a class where all classmates do so. The IV probits use the following instruments for peer behaviour: a=class proportion with older sibling drinking/smoking/using cannabis; b=class mean money per week; c=class share with higher educated father; d=class share describing themselves as academically able; e=class share living in two parent households. As a rough rule of thumb for instrument relevance an F-stat below 10 may indicate a weak instrument. Higher p-values for the chi-square test indicate excludable (valid) instruments. Higher p-values for the Wald test suggest peer behaviour can be treated as exogenous.

**Table 3: Peer Effects in Alcohol, Tobacco and Cannabis Use, Friends as Reference Group, Self-reports, Probits & Fixed Effects Probits, Marginal Effects (Standard Errors)**

	Alcohol Probit	Alcohol FE Probit	Tobacco Probit	Tobacco FE Probit	Cannabis Probit	Cannabis FE Probit
At least <i>some</i> friends used alcohol in last 30 days (%)	.437*** (.044)	.355*** (.046)				
At least a few friends smoked in last 30 days (%)			.506*** (.098)	.474*** (.091)		
At least a few friends used cannabis in last 30 days (%)					.350*** (.043)	.332*** (.044)
Older sibling that drinks	.139*** (.022)	.134*** (.022)				
Older sibling that smokes			.228*** (.027)	.217*** (.029)		
Older sibling that uses cannabis					.187*** (.027)	.169*** (.026)
Boy	<.001 (.022)	-.024 (.021)	-.051 (.034)	-.076** (.033)	.070*** (.021)	.057*** (.021)
Father HE	-.022 (.023)	-.023 (.024)	-.087*** (.025)	-.064** (.026)	-.044** (.022)	-.025 (.022)
Academically able	-.027 (.024)	-.033 (.025)	-.162*** (.029)	-.140*** (.030)	-.047* (.026)	-.030 (.026)
Two parent household	.018 (.026)	.022 (.025)	-.125*** (.031)	-.094*** (.031)	-.074*** (.026)	-.051* (.026)
Own money per week	.005*** (.001)	.005*** (.001)	.004*** (.001)	.004*** (.001)	.003*** (.001)	.003*** (.001)
Class fixed effects	No	Yes***	No	Yes***	No	Yes***
Pseudo R <sup>2</sup>	.110	.164	.124	.175	.143	.188
Nobs	1791	1749	1791	1776	1791	1742

Notes: \*\*\*, \*\* and \* denote statistical significance at 99%, 95% and 90% levels. The dependent variable is whether you drank alcohol, smoked tobacco or used cannabis in the last 30 days. Errors are clustered at classroom level. Marginal effects are calculated at means (money) or at zero (all binary dummies). For the peer behaviour variables they indicate the difference in the probability of own alcohol use between an individual with at least some friends that drink relative to an individual with no or only a few friends that drink, and an individual with at least a few friends that smoke/use cannabis relative to an individual with no friends that smoke/use cannabis. The fixed effects probits include school dummies and schools where all or no individuals drink/smoke/use cannabis are excluded.

**Table 4: Alcohol Use Probits, Reference Group Five Point Scales, Marginal Effects (Robust Standard Errors)**

	Model 1 (class as ref group)	Model 2 (friends as ref group)	Model 3 (both ref groups)
Proportion in class that drink (constructed 5 point scale)	.064*** (.021)		.026 (.020)
Proportion of friends that drink (5 point scale)		.209*** (.016)	.204*** (.016)
Older sibling drinks	.143*** (.021)	.110*** (.021)	.109*** (.021)
Boy	-.003 (.021)	<.001 (.020)	-.002 (.020)
Father HE	-.019 (.023)	-.016 (.023)	-.015 (.023)
Academically able	-.030 (.025)	-.023 (.024)	-.023 (.024)
Two parent house	.020 (.027)	.022 (.026)	.024 (.026)
Money to spend	.006*** (.001)	.004*** (.001)	.005*** (.001)
Pseudo R <sup>2</sup>	.056	.177	.179
Nobs	1791	1791	1791

Notes: \*\*\*, \*\* and \* denote statistical significance at 99%, 95% and 90% levels. The dependent variable is whether you drank alcohol in the last 30 days. The peer behaviour variables are 5 point scales in each case: as recorded in the data for friends and as constructed from class proportions for classmates (1=none, 2=a few, 3=some, 4=most, 5=all). Errors are clustered at classroom level. Marginal effects are calculated at means (peer behaviour, money) or at zero (all binary dummies). For the peer behaviour variables they indicate the impact on the probability of own alcohol use of a one unit increase in peer alcohol use (a one point move up the five point scale).

**Table 5: Smoking Probits, Reference Group Five Point Scales, Marginal Effects (Robust Standard Errors)**

	Model 1 (class as ref group)	Model 2 (friends as ref group)	Model 3 (both ref groups)
Proportion in class that smoke (constructed 5 point scale)	.041* (.023)		.020 (.021)
Proportion of friends that smoke (5 point scale)		.257*** (.018)	.253*** (.018)
Older sibling smokes	.233*** (.029)	.170*** (.027)	.169*** (.027)
Boy	-.059* (.032)	-.030 (.031)	-.030 (.030)
Father HE	-.077*** (.025)	-.057** (.026)	-.053** (.026)
Academically able	-.167*** (.030)	-.112*** (.030)	-.109*** (.030)
Two parent house	-.118*** (.030)	-.094*** (.030)	-.088*** (.030)
Money to spend	.004*** (.001)	.002 (.001)	.002 (.001)
Pseudo R <sup>2</sup>	.110	.240	.241
Nobs	1791	1791	1791

Notes: \*\*\*, \*\* and \* denote statistical significance at 99%, 95% and 90% levels. The dependent variable is whether you smoked tobacco in the last 30 days. The peer behaviour variables are 5 point scales in each case: as recorded in the data for friends and as constructed from class proportions for classmates (1=none, 2=a few, 3=some, 4=most, 5=all). Errors are clustered at classroom level. Marginal effects are calculated at means (peer behaviour, money) or at zero (all binary dummies). For the peer behaviour variables they indicate the impact on the probability of own smoking of a one unit increase in peer smoking (a one point move up the five point scale).

**Table 6: Cannabis Use Probits, Reference Group Five Point Scales, Marginal Effects (Robust Standard Errors)**

	Model 1 (class as ref group)	Model 2 (friends as ref group)	Model 3 (both ref groups)
Proportion in class that use cannabis (constructed 5 point scale)	.024* (.014)		.006 (.013)
Proportion of friends that use cannabis (5 point scale)		.170*** (.015)	.169*** (.016)
Older sibling uses cannabis	.233*** (.027)	.105*** (.025)	.104*** (.024)
Boy	.089*** (.027)	.068*** (.022)	.067*** (.023)
Father HE	-.049** (.023)	-.031 (.021)	-.030 (.021)
Academically able	-.042 (.026)	-.012 (.022)	-.012 (.022)
Two parent house	-.071*** (.026)	-.036* (.020)	-.035* (.020)
Money to spend	.004*** (.001)	.002*** (.001)	.002*** (.001)
Pseudo R <sup>2</sup>	.085	.273	.274
Nobs	1791	1791	1791

Notes: \*\*\*, \*\* and \* denote statistical significance at 99%, 95% and 90% levels. The dependent variable is whether you used cannabis in the last 30 days. The peer behaviour variables are 5 point scales in each case: as recorded in the data for friends and as constructed from class proportions for classmates (1=none, 2=a few, 3=some, 4=most, 5=all). Errors are clustered at classroom level. Marginal effects are calculated at means (peer behaviour, money) or at zero (all binary dummies). For the peer behaviour variables they indicate the impact on the probability of own cannabis use of a one unit increase in peer cannabis use (a one point move up the five point scale).