Cross-section and panel estimates of peer effects in early adolescent cannabis use: With a little help from my 'friends once removed'

Accepted Manuscript

Cross-section and panel estimates of peer effects in early adolescent cannabis use: With a little help from my ‘friends once removed’

John Moriarty, Duncan McVicar, Kathryn Higgins

PII: S0277-9536(16)30302-1

DOI: 10.1016/j.socscimed.2016.06.032

Reference: SSM 10705

To appear in: *Social Science & Medicine*

Received Date: 6 November 2015

Revised Date: 9 June 2016

Accepted Date: 11 June 2016

"Cross-section and Panel Estimates of Peer Effects in Early Adolescent Cannabis Use: With a Little Help from my 'Friends Once Removed' "

Abstract:

Peer effects in adolescent cannabis are difficult to estimate, due in part to the lack of appropriate data on behaviour and social ties. This paper exploits survey data that have many desirable properties and have not previously been used for this purpose. The data set, collected from teenagers in three annual waves from 2002-2004 contains longitudinal information about friendship networks within schools (N = 5,020). We exploit these data on network structure to estimate peer effects on adolescents from their nominated friends within school using two alternative approaches to identification. First, we present a cross-sectional instrumental variable (IV) estimate of peer effects that exploits network structure at the *second degree*, i.e. using information on friends of friends who are not themselves ego's friends to instrument for the cannabis use of friends. Second, we present an individual fixed effects estimate of peer effects using the full longitudinal structure of the data. Both innovations allow a greater degree of control for correlated effects than is commonly the case in the substance-use peer effects literature, improving our chances of obtaining estimates of peer effects than can be plausibly interpreted as causal. Both estimates suggest positive peer effects of non-trivial magnitude, although the IV estimate is imprecise. Furthermore, when we specify identical models with behaviour and characteristics of randomly selected school peers in place of friends', we find effectively zero effect from these 'placebo' peers, lending credence to our main estimates. We conclude that cross-sectional data can be used to estimate plausible positive peer effects on cannabis use where network structure information is available and appropriately exploited.

Keywords (2-5)

Peer Effects; Adolescence; Cannabis Use; Instrumental Variables; Network Structure

The authors hereby state we have no conflicts of interest of which we are aware.

1. Introduction

Early adolescent cannabis use has been linked to a variety of negative outcomes for the user (e.g. Patton et al., 2002; van Ours and Williams, 2009; Kuepper et al., 2011). Parents and policy makers may therefore want to discourage adolescents from using cannabis. When considering how to prevent widespread early initiation into cannabis use, it is desirable to know the extent to which one person's decision to use cannabis causes the choices of those around her to change (and vice-versa), i.e. peer effects. This can alter the cost-effectiveness of a given prevention strategy, as well as best advice for parents as to how involved they should be in their children's peer relationships. Young people themselves may also benefit from awareness both of how the behaviour of others affects their decision-making autonomy and how their own behaviour may steer others towards potentially harmful outcomes.

Estimating peer effects in cannabis use or any other behaviour, however, has been shown to be difficult. Despite considerable effort in the literature to overcome the identification problems raised by Manski (1993) and others, two resources have been in scant supply, namely panel data and data on network structure within schools. Because different individuals tend to nominate different groups of friends in such data, with nominations not always reciprocated, this achieves at least a partial separation of persons who are the *subjects* of peer effects and those peers who are the *source* of the effect (whereas such separation is difficult where the peer group simply consists of the reference group minus the individual; see Angrist, 2014). This separation, however, comes at the cost of the likely exacerbation of problems associated with other sources of correlation in behaviour within groups, e.g. young people's "homophilous" tendency to befriend those most similar to themselves.

Studies in this strand of the literature, which almost invariably use the friendship nominations data from the National Longitudinal Study of Adolescent Health (Add Health), have tried to overcome this by using lagged peer behaviour in place of current peer behaviour (e.g. Clark and Lohéac, 2007), by using various peer characteristics – such as percentage of peers whose parents smoke, who live with both biological parents, or who report having easy access to cigarettes or alcohol at home – to instrument for peer behaviour (e.g. Ali and Dwyer, 2009; Ali et al., 2011; Card and Giuliano, 2013; Halliday and Kwak, 2012), or by exploiting the mobility of school friends, either from switching schools or from graduating (Eisenberg, 2004). Where they look at cannabis use, these studies tend to report positive and significant, and in the case of Ali et al. (2011) quite large, peer effects. Concerns about potential correlated effects remain with each of these approaches, however, and causality is difficult to infer.

This paper also exploits network structure to estimate peer effects on adolescents from their nominated friends within school. But by adopting an instrumental variables (IV) strategy that exploits network structure at the *second degree* (following the pioneering suggestion of Bramoullé et al., 2009), we make a novel contribution to the cross-sectional peer effects literature. Specifically, we use information on friends of friends who are not themselves ego's friends – we call these 'friends once removed' – to derive an instrument for the cannabis use of friends which is more plausibly exogenous than instruments based directly on friends' characteristics. (This particular estimate was first reported in an earlier working paper version of parts of this paper; see Moriarty et al., 2012). Building on this innovation, we re-test the same model using behaviour and characteristics of randomly selected school peers in place of friends. This provides a falsification test, as any effect from these 'placebo' friends would suggest that effects in the main model are not the result of interactions and interpersonal influence.

Furthermore, by exploiting a data set that features *longitudinal* data on network structure within schools, we make a significant additional contribution. Specifically, we are also able to estimate a model including individual fixed effects. Both innovations allow a greater degree of control for correlated effects than is commonly the case in this literature, improving our chances of obtaining estimates of peer effects than can be plausibly interpreted as causal.

The remainder of this paper is set out as follows. Section 2 describes the data, drawn from the Belfast Youth Development Study (BYDS). Section 3 describes our approach to estimation. Section 4 presents and discusses results and Section 5 concludes.

2. The Belfast Youth Development Study

The BYDS is a longitudinal study of youth behaviour that tracks pupils in a single cohort across a sample of 42 schools in Northern Ireland from age 11/12 years (the first year of secondary school) through to the final year of compulsory secondary education at age 15/16 years. The first wave of data was collected in 2001 and the final (in-school) wave was collected in 2005. In this paper we focus on three consecutive waves of the BYDS – from wave 2, collected in 2002 when participants were aged 12/13 to wave 4, collected in 2004 when participants were aged 14/15. (Waves 1 and 5 are much smaller. In wave 1 fewer schools had been recruited than was the case for later waves. We omit wave 5 because of high individual-level attrition between waves 4 and 5.) 5,020 pupils answered at least one questionnaire between wave 2 and wave 4. In order to start where much of the literature leaves off, we also examine wave 3 as a cross-section. Wave 3 rendered the largest response from mainstream school pupils of any single wave (N = 4456), followed by wave 2 (N = 4,293) and wave 4 (N = 3900). Questionnaires were completed in school under exam conditions, placed by the students in sealed envelopes and collected by staff. Participation was contingent in the first instance on the agreement of school principals. Pupils who did not wish to participate could leave the questionnaire blank. Parents were informed of the school's participation and sent a consent withdrawal slip. If this was not returned, parents' consent was assumed. Ethical approval for this secondary analysis was granted by the Research Ethics Committee at the School of Sociology, Social Policy and Social Work. For further information on the BYDS, and on cannabis use within the BYDS sample, see McCrystal et al. (2007).

In the current study, cross-sectional analysis draws from Wave 3 responses of 4,289 participants. These are the participants who gave a valid response to questions relating to cannabis use and who nominated at least one friend, allowing for reference group cannabis use to be calculated. Longitudinal analysis draws on responses from the 2,839 individuals who make up the (valid) balanced panel. These are the participants who provided the above information in all three waves.

Our analysis uses information on an individual's cannabis use provided by the answers to two survey questions. First, we construct a binary dummy for cannabis use in the last year – taking the value 1 if the individual reports having used cannabis at least once, and 0 otherwise – using answers to the following question: "Have you tried cannabis in the last year?" This binary approach is the most common approach in the literature, although most studies report on substance use in the last 30 days or month rather than the last 12 months. BYDS does not ask about cannabis use over the shorter time period, but there is a follow-up question on frequency of use, as follows: "Thinking about cannabis, which of these statements best describes you: I have only used it once; I have used it between 2 and 5 times; I use it about once every month; I use it every week; I used to take it but don't anymore?" 17.72% of the wave 2 sample report having used cannabis at least once in the last 12 months, and 5.25% report using cannabis at least monthly. This rises to 28.37% and 11.5% in wave 3 and 36.65% and 18.16% in wave 4. We use this information to generate an additional dummy for at least monthly use. Very few students in our sample desist from cannabis use during the study period; but 8%, 9% and 13% respectively report uptake of cannabis use in waves 2-4, i.e. use among those who did not previously report use (see Table 1).

Respondents were also given space to nominate up to 10 friends within the school grade. The average number of reported friends in the sample is 7.35, and few list the maximum of 10 friends. In what follows we treat all those *not* nominated by an individual as 'non-friends', but potential *friends once removed*, should they be nominated by one of ego's friends. (Where individuals are constrained by the maximum number of nominations, our conjecture is that friends are listed roughly in order of the perceived 'closeness' of the friendship, so if there are non-listed friends they are not close friends.) Reference group behaviour, which closely resembles own reported behaviour, is captured through derived variables denoting the proportion of nominated friends that report using cannabis in the last 12 months / monthly. Note that we also have information on whether friendship nominations are reciprocated. Over the observed period of this study participants receive a fairly stable proportion of reciprocal nominations from nominated friends, at just below two thirds on average.

In addition to information on use of cannabis and other substances, the BYDS contains data on a host of individual and family background characteristics, as well as limited information at the school level. Greater likelihood of drug use has been shown to be associated with being from more deprived socioeconomic groups (Buu et al., 2009); with being male (Amaro et al., 2001); with being older (McVicar & Polanski, 2014); with having older siblings (Latimer et al., 2004); with living in single-parent or reconstituted households (Griesback et al., 2003); and with lower educational ambitions (van Ours & Williams, 2009). We include controls for all of these factors (for more detail on the derivation of controls using the BYDS data see Moriarty & Higgins, 2015). At the school level, we also use dummy controls for single-sex schools, academically selective and Catholic-maintained schools (which drop out in our models including school fixed-effects). We generate binary dummies for missing values for the subset of variables where they occur and include these as additional controls.

Additionally, we control for the following characteristics of the individual's social network: number of friends named; whether the individual reports their friends in the school but outside the grade being older (older school friends), whether the individual reports their friends outside the school being older (older non-school friends). Table 1 reports sample means for all controls and for the cannabis use variables by wave. Note the similarity of sample characteristics between the balanced panel and the wave 3 cross-section.

<Table 1 around here>

Note that possession of cannabis is illegal in Northern Ireland, as it is in the rest of the UK, and this was the case throughout the period we look at here. However, cannabis was downgraded from a class B to a class C drug throughout the UK in January 2004, i.e. after wave 3 but prior to wave 4 of the BYDS survey. In effect this meant those possessing small amounts of cannabis would not be arrested but would receive a police warning. So our cross-sectional estimates refer to a period where cannabis was classified as a class B drug. Our fixed effects model – which spans the downgrading – controls for this and any other common shocks by including time fixed effects alongside individual fixed effects. Estimates of the prevalence of cannabis use among adolescents in Northern Ireland at this time tend to be higher than those for most other European countries, but broadly in line with or slightly below those for other parts of the UK (see Higgins et al., 2004).

3. Estimation Approach

We present both cross-section estimates and panel estimates of peer effects in cannabis use, initially using a single wave of BYDS data, then using three consecutive waves. Our primary motivation for first presenting cross-section estimates is that, with the exception of BYDS, survey data on adolescent substance use containing information on network structure tends to be cross-sectional in nature (e.g. Add Health). We show that by more fully exploiting network structure along the lines of Bramoullé et al. (2009), such data can potentially generate estimates of peer effects that are more plausibly interpreted as causal and are also qualitatively in line with those obtained from panel models estimated on longitudinal network structure data.

3.1. Cross-sectional IV estimates

We begin by using data drawn from wave 3 to explore the extent to which nominated friends influence own behaviour. Along the lines of Clark and Lohéac (2007), Ali and Dwyer (2009), Ali et al. (2011) and Halliday and Kwak (2012), we first specify a linear-in-means model, estimated by OLS, as follows:

(1) $$y_{is} = \delta \bar{Y}_{if} + X'_{is}\beta + \alpha_s + \varepsilon_{is}$$

where y_{is} is a binary dummy equal to 1 if individual *i* reports using cannabis in the reference period and 0 otherwise. Reference group behaviour is denoted by $\bar{Y}_{if}$, (the proportion of *i*'s nominated friends using cannabis), where the subscript *f* denotes the set of nominated friends. X_{is} and α_s denote observable individual controls (for ego) and school dummies, respectively, and ε_{is} is an error term which we allow to be clustered at the school level. The parameter of interest is δ, which captures the association between own cannabis use and the prevalence of cannabis use among nominated friends.

Even within schools, however, $\bar{Y}_{if}$ is unlikely to be uncorrelated with the error term ε_{is}. We therefore follow several earlier studies by instrumenting for $\bar{Y}_{if}$ and adopting a two stage least squares approach (e.g. Ali and Dwyer, 2009, who instrument for peer behaviour using the proportion of peers whose parents smoke, who live with both biological parents, and who report having easy access to cigarettes at home; Ali et al., 2011, who instrument for peer behaviour using the latter two variables plus the proportion of peers who report being satisfied with their relationships with their parents plus the proportion of peers who report easy access to alcohol at home; Halliday and Kwak, 2012, who instrument for peer behaviour using the proportion of peers living with biological parents and whose mothers/fathers have college degrees).

Our point of departure from these earlier IV studies, however, is that instead of using observed information about friends themselves to instrument for their cannabis use, we follow Bramoullé et al. (2009) – who apply the method to study peer effects in the sports activities of adolescents in the Add Health data – by using information on *friends once removed* to instrument for friends' behaviour. These are individuals within the school year who are nominated as friends by ego's friends, but not by ego herself. This is illustrated in Figure 1, where individual ID46 (ego) has nominated six friends (ID21, ID5, ID71, ID30, ID17 and ID8). They in turn nominate their own friends. Of these, only those nominated by ego's friends who are not also nominated by ego, i.e. those with no direct link to ego, are counted as *friends once removed* (ID3, ID51, ID61, ID12, ID41, ID9 and ID11).

< Figure 1 around here >

The exploitation of information about these friends once removed to instrument for friends' cannabis use is formalised as follows:

(2) $$y_{is} = \delta_1 \bar{Y}_{if} + X'_{is}\beta_1 + \bar{X}'_{if}\beta_2 + a_s + e_{is}$$

(3) $$\bar{Y}_{if} = \pi \bar{Z}_{for} + X'_{is}b_1 + \bar{X}'_{if}b_2 + \theta_s + \varphi_{is}$$

Here, the 'friends once removed' instrument, which is assumed to be excludable from (2), is represented by $\bar{Z}_{for}$. The parameter of interest is now δ_1 which, under the assumption of excludability, can be interpreted as the impact of friends' cannabis use on own cannabis use, i.e. the peer effect. Note that De Giorgi et al. (2009) also exploit network structure in a similar way – in their case tutorial colleagues of tutorial colleagues rather than friends of friends – to study peer effects in academic performance. Ours is the first study to use this IV approach to estimate peer effects in adolescent substance use.

Homophilic selection of friends once removed – which would lead one to question the excludability assumption – is considerably less likely than in the case of nominated friends. For one thing, most individuals nominate fewer than ten friends, i.e. they do not exhaust the space on the questionnaire for nominations. For another, the average size of a school grade in the BYDS sample is just over 100 individuals, so significant proactive *non*-selection is unlikely. Further, in contrast to the IV approaches of Ali and Dwyer (2009), Ali et al. (2011) and Halliday and Kwak (2012), our IV approach enables *all* observed characteristics of friends to be controlled for in (2) as contextual effects $\overline{X}_{if}'$. In other words our case for the excludability of the instrument rests not on an assumption of no contextual effects (i.e. of no impact of friends' characteristics on own behaviour), but on an assumption of orthogonality of own behaviour with the characteristics of *friends once removed*, drawn from among the *non*-friends in the class. The trade-off for this more plausible case for instrument excludability is a potentially weaker instrument than would be the case were we to use the standard approach, i.e. an instrument which is likely to be less strongly correlated with friends' cannabis use.

Specifically, we use the proportion of friends once removed who report having older non-school friends to instrument for the proportion of nominated friends using cannabis. We know from a controls-only regression that having older non-school friends is a strong predictor of own cannabis use (with a coefficient of 0.11, statistically significant at the 99.9% level – see Table 3), and there is a positive relationship in the reduced form that is statistically significant at 95% (but not 99%), suggesting a reasonably strong first stage, i.e. that the instrument is correlated with the behaviour of ego's nominated friends. . We also include school fixed effects here to wash out any remaining correlated effects at the school level. Also note that in the IV model we are able to include in $\overline{X}_{if}'$ a control for the proportion of *friends* reporting having older non-school friends. This ensures the reduced form correlation between the instrument and own cannabis use does not work through ego having access to cannabis through contact with the older non-school friends of friends.

Table 2 shows that our instrument is orthogonal to all but two of ego's observed characteristics, including whether ego reports having older non-school friends (also included as a control). The first exception is an association with having more cars (a proxy for wealth) which may be a chance association, as there is no intuitive reason why the friends of friends of wealthier pupils might prefer older non-school friends. (We would expect one from 17 correlations to be significant at the 95% level purely at random.) The second is an association between the instrument and ego having friends from older grades, which appears to be picking up characteristics of immediate friends, including friends' network connections, given that the coefficient is non-significant when these contextual effects are controlled for, as they are in our IV model.

<Table 2 around here>

3.2. Individual fixed effects estimates

Because BYDS includes information on network structure not just at one wave but at multiple waves, it also enables an individual fixed effects approach to estimation. We use data across waves 2-4 to estimate the following model on the balanced panel:

$$y_{ist} = d\bar{Y}_{ift} + X'_{ist}\psi + \lambda_i + \lambda_s + \lambda_t + u_{ist} \qquad (4)$$

where *f* denotes nominated friends at time t, λ_i and λ_t are individual and wave fixed effects, and X_{ist} now omits time-invariant controls.

In estimating (4) we are investigating whether friends' cannabis use is associated with own cannabis use controlling for all time-invariant differences between individuals, whether observed or unobserved, and for observable time-varying differences. (Only unobserved time-varying heterogeneity remains as a possible source of correlated effects.) Conditional on these controls, and with two additional assumptions, we interpret a statistically significant relationship between friends' cannabis use and own use as indicating a potentially causal relationship, i.e. a peer effect. The two additional assumptions are first that any peer effect from ego to friends is quantitatively trivial given the size of nominated friendship groups and the fact that many nominations are unreciprocated, and second, in contrast to our earlier IV approach, that there are no contextual effects. Neither assumption is strictly necessary if we are prepared to loosen our interpretation of the estimated δ as capturing peer effects writ large rather than peer effects from a particular reference group behaviour. Ours is the first study in this literature to present individual fixed estimates where nominated friends are the reference group. (Sen (2009) uses multiple waves of the Waterloo Smoking Prevention Programme data to estimate a fixed effects model of peer effects in smoking behaviour, but relies on class-level average behaviour, rather than having nominated friends as a reference group.)

To test sensitivity (to non-random attrition among other things) we also estimate the model on the unbalanced panel (with and without an additional dummy variable for attrition in next wave). We also estimate a version of the model on the balanced panel with peer cannabis use lagged one year (along the lines of Clark and Lohéac, 2007), and on the balanced panel with lagged rather than current friendship nominations, i.e. friendship measured at time t-1.

Note that the phrasing and structure of items regarding ego's older school and non-school friends varied substantially across the three waves. Thus we cannot extend our IV approach in the cross-section to the panel case using these data. Having said that, even if an identical question had been asked, it would be unlikely to display sufficient genuine - i.e. not attributable to measurement error – variation over three waves to make an IV with individual fixed effects approach feasible here, given the sample size.

4. Results and Discussion

Table 3 presents results from estimating both (1) and (2) and (3) on the wave 3 cross-section and (4) on the balanced panel. As a point of reference, estimating (1) by OLS suggests a strong, highly significant and positive association between own cannabis use and friends' cannabis use, with a one percentage point increase in the proportion of friends that report using cannabis in the last 12 months associated with a .55 percentage point increase in own probability of reporting cannabis use in the last 12 months.

<Table 3 around here>

Turning to our IV estimate – column 4 of Table 3 - the estimated coefficient on friends' cannabis use is smaller than the OLS equivalent as we would expect, although it remains large and positive. Note that this estimate is smaller than the closest comparable IV estimate of Ali et al. (2011), despite being for use in the last 12 months. However, the imprecision of the estimate renders it statistically insignificant. The high standard error reflects the small sample size coupled with a 'weak-ish' instrument – note the first stage F statistic of 10.8 – i.e. that our instrument is only detecting a small amount of exogenous variation in friends' cannabis use prevalence. The bottom line is that the positive IV estimate provides tentative evidence for a peer effect in early adolescent cannabis use, without the need for an unsupported 'no contextual effects' assumption, but that the instrument is insufficiently strong to detect a clear peer effect signal in a sample the size of BYDS.

Consider now the results from our fixed effects model as reported in column 5 of Table 3. Again we see evidence of a large, positive and in this case statistically significant association between own cannabis use and friends' cannabis use. Given the fixed effects and other controls and our assumption regarding the direction of causality, this too provides tentative evidence for peer effects in early adolescent cannabis use.

Note that in both the IV and OLS models estimated on the cross-section, associations with control variables are generally in line with those found in previous studies cited earlier in the paper (e.g. cannabis use is increasing with age, is positively associated with living in single-parent or reconstituted households, and is positively associated with having lower educational ambitions). The latter effect survives in the individual fixed effects model, along with a negative impact from paternal employment, but with the individual fixed effects themselves explaining 78% of the variation in the dependent variable, none of the other controls are significantly associated with own cannabis use.

<Table 4 around here>

These estimates are robust to a variety of sensitivity analyses (see Table 4). For example, the IV estimates are very similar when estimated on the wave 3 sample restricted to the balanced panel (albeit with larger standard errors), as are the fixed effects estimates when estimated on the unbalanced panel (with or without the inclusion of a dummy for attriting in the next wave, which is itself statistically insignificant). Both IV and fixed effects point estimates are similar when using the monthly cannabis use dummy in place of the dummy for cannabis use in the last 12 months (although again with larger standard errors). The main exception to the robustness of our cross-sectional results is the IV model excluding observed contextual effects, which produces a coefficient closer to the original OLS association, suggesting that controlling for contextual effects is important and providing further suggestive evidence that identification strategies that rest on the assumption of no contextual effects are questionable. When we ourselves estimate an IV model where we replace our friends once removed instrument with the same variable at the friends level, i.e. where we instrument with the proportion of friends who report having older non-school friends rather than the proportion of friends once removed who report having older non-school friends, we obtain an estimate that is actually larger in magnitude than the OLS estimate, but that is much more precisely estimated than our preferred IV estimate.

Using lagged friends' cannabis use (defined either as the cannabis use at *t* of friends nominated at *t*-1, or like Clark and Lohéac (2007) as the cannabis use at *t*-1 of friends nominated at *t*) in the fixed effects model gives a small positive and statistically significant coefficient, though in each case we see a dilution of the effect compared to that estimated for current friends' current use. Clark and Lohéac (2007) argue that using the lag of peer behaviour in place of current peer behaviour avoids Manski's reflection problem, so obviating the need for the no contextual effects assumption. On the other hand peer outcomes are likely to be correlated over time so that lagged peer behaviours proxy for current behaviours and may therefore still be subject to the reflection problem (Manski, 1993). Lagged peer outcomes will also overlook recent changes in peer behaviour which might impact on own outcomes and are therefore at best interpretable as lower bounds estimates of the peer effect of interest (Hanushek et al., 2003). The motivation behind also estimating a version of the model using current use of lagged friends is similar, i.e. to exploit the churn over time in friendship networks to try to wash out reverse causality. Using the behaviour at *t*-1 of nominated friends at *t*-1, we find a zero. Our interpretation of this is that it reflects the extent of change taking place in young adolescents' social milieu and in cannabis use over this one-year interval. Whereas the positive effects in the two preceding lag model specifications reflect continuation between *t*-1 and *t*, either via continued friendships or continued behaviour, in the final model we stack the cards against finding a peer effect too far.

Finally, to provide further support to our argument that the IV and fixed effects estimates presented above can be tentatively interpreted as indicating causal relationships, we re-estimate both models using randomly-generated friendship nominations from among the non-friends, i.e. for each individual we randomly replace each of their friendship nominations with a 'placebo friend' from among the non-friends within the school grade. These falsification tests are reported in Table 5. Friends once removed are defined in the usual way for these 'placebo friends'. In each case the estimated association between own cannabis use and that of 'placebo friends' is zero. This is also the case for the OLS estimate on the wave 3 cross-section and fixed effects using 'placebo friends'' current use.

<Table 5 around here>

5. Conclusions

This paper presents estimates of peer influences in early adolescent cannabis use using school-based survey data not previously exploited for this purpose. The observed increase in cannabis use over the three years focused on (age 12-age 15) highlights that this is a period where many are first exposed to cannabis use. However, users remain in the minority and there is clear value in understanding the influential role peer behaviour appears to have in determining which young people join that minority. Despite its relatively small sample size, the BYDS data set has a number of attractive properties allowing us to make a significant contribution to the substance use peer effects literature.

First, by exploiting information on nominated friends, the reference group can be limited to those peers with whom social interactions regularly occur. Ours is one of only a handful of studies that provide such estimates and we do so using new data (not Add Health) not previously used for this purpose. Furthermore, such network data affords us an estimation approach using information for friends of friends who are not nominated friends of the individual to instrument for nominated friends' behaviour. Combined with school fixed effects, this approach offers a plausible method for identifying peer effects in cannabis use where only cross-sectional data are available, without relying on the usual assumption of no contextual effects. Ours is the first paper to adopt this approach – first suggested by Bramoullé et al. (2009) – in this particular literature. Given the scarcity of data containing network structure over multiple waves, this IV approach demonstrates an effective utilisation of cross-sectional data containing network structure which could be usefully replicated in other contexts.

Second, by capitalising on the longitudinal structure of the BYDS data, we can compare estimates from this cross-sectional IV approach with those from an alternative individual fixed effects approach resting on different assumptions. Given the uniqueness of our data in that it features longitudinal information on network structure within schools, ours is the first study in this literature to present such estimates where the reference group is nominated friends.

The estimates from both approaches suggest large peer effects in cannabis use – though smaller than those suggested by Ali et al. (2011) – between early adolescent friends, with a 10% increase in friends' cannabis use associated with an increase in the probability of ego's use by between 2.5% and 4%, although our IV estimate is too imprecise to be statistically significant at standard levels. (The imprecision of the main IV estimate, reflecting the relatively small sample size of the BYDS study, is the main weakness of the paper.) But even if in isolation each estimate can only be interpreted as tentatively suggesting peer effects in adolescent cannabis use, taken together they provide reasonably strong evidence – using novel approaches and new data – that peer effects exist and are non-trivial in magnitude. Policy makers can therefore potentially exploit social multipliers – get more bang for their buck – in interventions to reduce adolescent cannabis use. Parents can also potentially reduce the probability of their adolescent children taking up cannabis use by influencing the friendship network formation of their children. Finally, communicating the results of this research to young people themselves may raise awareness both of how the behaviour of others affects their decision-making autonomy and how their own behaviour may steer others towards potentially harmful outcomes.

References

Ali, MM., Amialchuk, A. and Dwyer, DS. (2011). 'The social contagion effect of marijuana use among adolescents.' *PloS One,* 6(1), e16183-e16188.

Ali, MM. and Dwyer, DS. (2009). 'Estimating peer effects in adolescent smoking behavior: a longitudinal analysis.' *Journal of Adolescent Health*, 45(4), 402-408.

Amoro, H., Blake, S., Schwarz, P. and Flinchbaugh, F. (2001). ‘Developing theory-based substance abuse prevention programs for young adolescent girls.’ *Journal of Early Adolescence,* 21(3), 256-293.

Angrist, JD. (2014). 'The perils of peer effects.' *Labour Economics,* 30, 98-108.

Bramoullé, Y., Djebbari, H. and Fortin, B. (2009). 'Identification of peer effects through social networks.' *Journal of Econometrics*, 150, 41-55.

Buu, A., DiPiazza, C., Wang, J., Puttler, LI., Fitzgerald, HE., Zucker, RA. (2009). ‘Parent, family, and neighborhood effects on the development of child substance use and other psychopathology from preschool to the start of adulthood.’ *Journal of Studies on Alcohol and Drugs,* 70(4), 489–498.

Card, D. and Giuliano, L. (2013). 'Peer effects and multiple equilibria in the risky behavior of friends.' *Review of Economics and Statistics,* 95(4), 1130-1149.

Clark, AE. and Lohéac, Y. (2007). '"It wasn't me, it was them!" Social influence in risky behaviour by adolescents.' *Journal of Health Economics,* 26, 763-764-784.

De Giorgi, G., Pellizzari, M. and Redaelli, S. (2009). 'Be as careful of the company you keep as of the books you read: peer effects in education and on the labour market.' Working Paper No. 14948, National Bureau of Economic Research, Cambridge, MA.

Eisenberg, D. (2004). 'Peer effects for adolescent substance use: Do they really exist?' *Working Paper UC Berkeley School of Public Health.*

Griesbach, D., Amos, A., and Currie, C. (2003). ‘Adolescent smoking and family structure in Europe.’ *Social Science and Medicine*, 56, 41-52.

Halliday, TJ. and Kwak, S. (2012). 'What is a peer? the role of network definitions in estimation of endogenous peer effects.' *Applied Economics,* 44(3), 289-302.

Hanushek, EA., Kain, JF., Markman, JM. and Rivkin, SG. (2003). 'Does peer ability affect student achievement?' *Journal of Applied Econometrics,* 18(5), 527-544.

Higgins, K., Percy, A. and McCrystal, P. (2004). ‘Secular trends in substance use: the conflict and young people in Northern Ireland.’ *Journal of Social Issues*, 60(3), 485-506.

Kuepper, R., van Os, J., Lieb, R., Wittchen, HU., Höfler, M. and Henquet, C. (2011). 'Continued cannabis use and risk of incidence and persistence of psychotic symptoms: 10 year follow-up cohort study.' *British Medical Journal,* 342, 745.

Latimer, W., Floyd, LJ., Kariis, T., Novotna, G., Exnerova, P. and O’Brien, M. (2004). ‘Peer and sibling substance use: predictors of substance use among adolescents in Mexico.’ Revista Panamericana de Salud Publica, 15(4), 225-232.

Manski, CF. (1993). 'Identification of endogenous social effects: The reflection problem.' *The Review of Economic Studies*, 60(3), 531-542.

McCrystal, P., Percy, A. and Higgins, K. (2007). 'Frequent cannabis use among 14/15 years old in Northern Ireland.' *Drug and Alcohol Dependence,* 88(1), 19-27.

McVicar, D. and Polanski, A. (2014). 'Peer effects in UK adolescent substance use: Never mind the classmates?' *Oxford Bulletin of Economics and Statistics*, 76(4), 589-604.

Moriarty, J. and Higgins, K. (2015). 'Effects of peer network interactions on adolescent cannabis use,' Journal of Criminal Psychology, 5(2), 75-91.

Moriarty, J., McVicar, D. and Higgins, K. (2012). 'Peer effects in Adolescent cannabis use: It's the friends, stupid.' *Melbourne Institute of Applied Economic and Social Research Working Paper Series: 27/12.*

Patton, G. C., Coffey, C., Carlin, J. B., Degenhardt, L., Lynskey, M. and Hall, W. (2002). 'Cannabis use and mental health in young people: Cohort study.' *British Medical Journal*, 325(7374), 1195-1198.

Sen, A. (2009). 'Estimating the impacts of household behavior on youth smoking: Evidence from Ontario, Canada.' *Review of Economics of the Household*, 7(2), 189-218.

van Ours, JC. and Williams, J. (2009). ‘Why parents worry: Initiation into cannabis use by youth and their educational attainment.’ *Journal of Health Economics,* 28(1), 132-142.

Table 1: Observed Variables and Sample Means

Variable	Balanced Panel Wave 2	Balanced Panel Wave 3	Balanced Panel Wave 4	Cross-Section Wave 3
Own use of cannabis in the past 12 months	.14	.26	.36	.28
Proportion of nominated friends using in past 12 months	.16	.27	.37	.30
Cannabis uptakers	.08	.09	.13	-
Cannabis desisters	<.01	.01	.06	-
Uses cannabis at least monthly	.04	.09	.18	.11
Sex-male	.45	.45	.45	.47
Age at which began school, years	11.68	11.68	11.68	11.7
Number of friends also valid	6.44	6.81	6.76	6.52
Proportion of nominations that are reciprocated	.65	.66	.65	.65
Number of brothers	1.02	1.00	.98	1.03
Number of sisters	.96	.95	.92	.98
Number of cars in household	1.5	1.57	1.59	1.46
Lives with single parent	.13	.14	.14	.15
Lives with parent and step-parent or parent's partner	.07	.07	.08	.08
Lives in alternative family structure	< .01	.02	.02	.02
Family structure missing	.01	.01	.01	.01
Wants to leave school at 16	.20	.18	.19	.20
Wants to leave school at 16 missing	.02	.02	-	.03
Has older friends in the school	-	.21	.20	.23
Has older friends outside of school	-	.21	-	.22
Friends once removed have older friends outside of school		.22		.23
Mother works full-time	.35	.35	.36	.33
Mother works part-time	.30	.29	.29	.29
Mother work missing	.03	.09	.19	.10
Father works full-time	.66	.65	.66	.63
Father works part-time	.08	.08	.07	.08
Father work missing	.06	.18	.09	.20
N. Obs	2,839	2,839	2,839	4,416

Note: The wave 3 cross-section includes all wave 3 respondents other than those with missing cannabis use information, age information, pupils in alternative education.

Figure 1: Network diagram of "friends once removed"

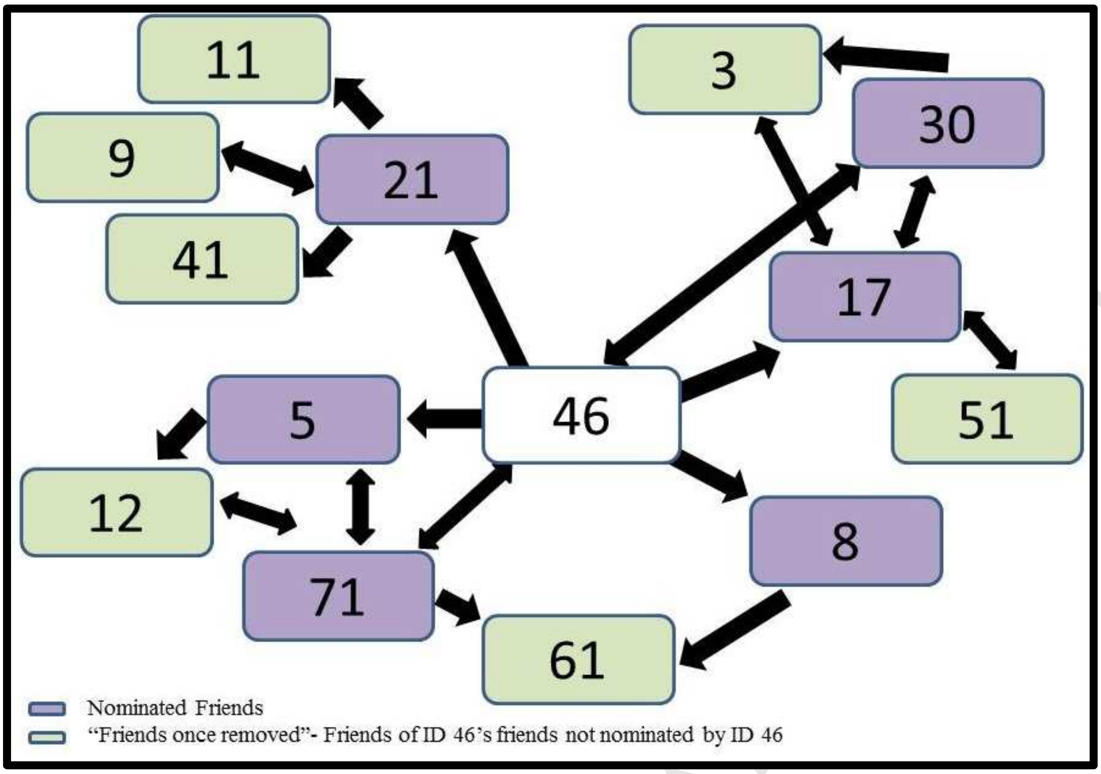

Table 2: Balancing Tests: OLS associations with instrument ("friends once removed" having older non-school friends)

	Association with instrument, no contextual effects	Association with instrument, controlling for contextual effects
Sex-male	.006 (.085)	.005 (.012)
Age	.055 (.036)	.041 (.036)
Number of friends	.187 (.466)	.286 (.455)
Number of brothers	.151 (.122)	.134 (.121)
Number of sisters	-.027 (.120)	-.052 (.121)
Number of cars	.359* (.148)	.385* (.131)
Single parent family	.044 (.034)	.023 (.041)
Step family	.006 (.032)	-.001 (.034)
Alternative family	-.022 (.018)	-.028 (.018)
Want to leave school at 16	.079 (.044)	.041 (.037)
No response on educational ambition	.007 (.019)	.000 (.022)
Older School Friends	.143* (.067)	.097 (.050)
Older Non-School Friends	.080 (.069)	.041 (.062)
Mother works full-time	-.053 (.056)	-.049 (.056)
Mother works part-time	.001 (.045)	.011 (.049)
Father works full-time	-.023 (.047)	.010 (.042)
Father works part-time	-.036 (.037)	-.039 (.039)
School fixed effects	Yes	Yes
Time (wave) fixed effects	No	No
Individual fixed effects	No	No
Nobs	4276	4276

Notes: * significant at 95% level. Standard errors are clustered at the school level. The first column gives coefficients for separate OLS regressions of the ego characteristic on the IV controlling for school fixed effects only. The second column also controls for the full set of contextual effects, i.e. the characteristics of nominated friends, including the proportion of friends having older non-school friends.

Table 3: Estimated Peer Effects in Cannabis Use in Last 12 Months

	Wave 3 Cross Section			Balanced Panel
	OLS (controls only)	OLS (w/school dummies)	IV	Fixed Effects
Nominated friend's cannabis use	-	.553 * (.029)	.366 (.396)	.295* (.024)
Sex-male	.045 (.029)	.028 (.019)	.104 (.077)	-
Age	.065* (.018)	.044 * (.02)	.043* (.017)	-
Number of friends	.004 (.003)	.005 (.003)	.003 (.003)	-.003 (.002)
Number of brothers	.009 (.007)	.003 (.006)	.004 (.007)	-.001 (.012)
Number of sisters	.008 (.007)	.005 (.007)	.004 (.007)	-.01 (.011)
Number of cars	.012* (.006)	.009 (.006)	.01 (.006)	-.003 (.007)
Single parent family	.103* (.028)	.086 * (.025)	.086* (.027)	-.016 (.028)
Step family	.124* (.028)	.092 * (.025)	.093* (.03)	.009 (.036)
Alternative family	.060 (.041)	.037 (.048)	.038 (.037)	.043 (.046)
Want to leave school at 16	.141* (.023)	.101 * (.018)	.102* (.033)	.036* (.015)
No response on educational ambition	.206* (.043)	.186 * (.045)	.186* (.049)	.028 (.042)
Older School Friends	.111* (.017)	.078 * (.016)	.076* (.021)	-
Older Non-School Friends	.109* (.017)	.086 * (.017)	.088* (.018)	-
Mother works full-time	.036* (.018)	.034 * (.017)	.035* (.017)	.030 (.017)
Mother works part-time	.024 (.020)	.026 (.017)	.025 (.017)	.005 (.016)
Father works full-time	-.019 (.025)	-.007 (.023)	-.005 (.025)	-.039* (.018)
Father works part-time	-.001 (.028)	-.007 (.030)	-.005 (.025)	-.050* (.024)
School fixed effects	Yes	Yes	Yes	Yes
Contextual effects (mean friends' characteristics)	No	No	Yes	No
Time (wave) fixed effects	No	No	No	Yes
Individual fixed effects	No	No	No	Yes
N. Observations	4416	4289	4276	2839
F (Excluded instrument, first stage)	-	-	10.78*	-
First stage association between instrument and friends' cannabis use	-	-	.160* (.048)	-
R^2 / within R^2	.16	.24	.24	.15/.11

Notes: * significant at 95% level. Standard errors are clustered at the school level in the cross-section and at the individual level for the fixed effects model. Additional controls (not listed) for missing values for the following variables: number of brothers/sisters, family structure, number of cars, mother & father employment. Panel outcome is current cannabis use in wave 2/3/4.

Table 4: Sensitivity Analysis, Estimated Peer Effect Coefficient (Standard Error)

		Coefficient	Standard Error	First stage F stat	No. obs
Wave 3 IV (Balanced panel only)		.257	.593	4.29	2833
Wave 3 cross-section IV no contextual effects		.591*	.263	17.4	4276
Wave 3 IV, monthly use		.285	.962	2.93	4243
Wave 3 IV, friends IV in place of friends once removed IV		.768*	.200	20.8	4289
Fixed effects:					
Wave 2-4, Unbalanced	Friends at *t*, use at *t*	.266*	.022		11708
Wave 2-4, Balanced	Friends at *t*, use at *t-1*	.076*	.036		5678
	Friends at *t-1*, use at *t*	.176*	.031		5678
	Friends at *t-1*, use at *t-1*	-.020	.035		5678
	Friends' at *t* monthly use at *t*	.344*	.032		8443

Notes: * Significant at 95% probability level. Standard errors clustered at the school level in the cross-section and at the individual level for the fixed effects models. Models include the controls as listed in Table 2. IV models use the proportion of "friends once removed" who prefer older non-school friends as a single instrument.

Table 5: Falsification Test: Random Peer Groups

	Coefficient	Standard Error	No. obs.
OLS, wave 3 cross-section	-.043	.042	4247
IV, wave 3 cross-section	-.186	.492	4235
FE, balanced panel	.037	.021	7296

Notes: All controls as for the models reported in Table 3. Standard errors clustered at the school level in the cross-section and at the individual level for the fixed effects models. 'Nominated friends' generated randomly from among the non-friends within grade. The first stage F-statistic for the instrument in the IV model is 7.03.

- Credibly-identified peer effects estimates for adolescent cannabis use are shown
- Uses rich new survey data with information on friendship networks within schools across multiple consecutive waves
- Cross-sectional identification achieved by instrumenting using friends of friends characteristics
- Longitudinal dimension of data exploited to produce individual fixed effects
- Friends’ cannabis use appears to impact on own cannabis use

SSM-D-15-03093R1

"Cross-section and Panel Estimates of Peer Effects in Early Adolescent Cannabis Use: With a Little Help from my 'Friends Once Removed' "

John Moriarty, Administrative Data Research Centre for Northern Ireland,

Duncan McVicar, Queen's University Belfast School of Management

Kathryn Higgins, Institute for Child Care Research, Queen's University Belfast

Corresponding author:
John Moriarty,
Administrative Data Research Centre for Northern Ireland,
Room 02.013 2nd Floor,
Institute for Clinical Science Block B,
Royal Victoria Hospital,
Grosvenor Road,
Belfast, BT126BJ

j.moriarty@qub.ac.uk

Acknowledgements

The authors gratefully acknowledge the work of the many researchers, fieldworkers and data managers involved in collating and assembling the BYDS data over the past 17 years. We also acknowledge the input of seminar participants at multiple venues.

The Belfast Youth Development Study was funded by the Northern Ireland Health and Social Care Research and Development Division. Moriarty's PhD studentship was funded by Atlantic Philanthropies via the Improving Children's Lives Initiative.