Published in final edited form as:

Soc Sci Med. 2016 March ; 152: 61–69. doi:10.1016/j.socscimed.2016.01.036.

The Effect of Peers on HIV Infection Expectations among Malawian Adolescents: Using an Instrumental Variables/School Fixed Effect Approach

Jinho Kim*

Department of Sociology, University of Wisconsin–Madison, 1180 Observatory Drive, Madison, WI 53706

Abstract

Malawian adolescents overestimate their HIV infection risk. Understanding why they do so is important since such overestimation is likely to be linked to later-life outcomes. This study focuses on the influence peers have on HIV infection expectations. I use novel school-based survey data collected in Malawi between October 2011 and March 2012 ($n = 7,910$), which has more reliable measures of peers' HIV infection expectations than other studies. I employ a combined instrumental variables/fixed effects methodology designed to address several methodological challenges in estimating peer effects, including self-selection of friends, the issue of unobserved environmental confounders, and the bi-directionality of peer effects. Several tests are conducted in order to assess the robustness of the specifications. Results suggest that a one-percentage-point increase in the mean probabilistic expectation of HIV infection among peers increases an adolescent's own subjective expectation of infection by an average of 0.65 percentage points. This paper shows that peer influence is greater for males than for females. Results also suggest that the peer effects on HIV infection expectations are only statistically significant among those lacking more complete knowledge of HIV/AIDS.

Keywords

HIV/AIDS; risk perceptions; social interactions; peer effects; sub-Saharan Africa

1. Introduction

Adolescents' risk perceptions are found to be less accurate than adults, especially when it comes to mortality expectations (Fischhoff et al., 2000; Millstein and Halpern-Felsher, 2002). For example, adolescents in the U.S. tend to have unrealistically high expectations of early death (Jamieson and Romer, 2008). Similarly, Malawian adolescents clearly overestimate their risk of HIV infection (Anglewicz and Kohler, 2009). In the sample used for this study (7,910 secondary school students with an average age of 16), the reported mean subjective probability of having HIV is 16%, while in reality, the HIV prevalence rate among Malawian adolescents aged 15-19 is only 2.7% (National Statistical Office, 2011). These high HIV infection expectations are especially surprising considering that many

*Corresponding author. jinhokim@ssc.wisc.edu, Telephone: (608) 571-9315, Fax: (608) 262-8400.

adolescents are sexually inexperienced: among those aged 19-24 with more than secondary education, 12.6% of women and 35.2% of men initiated sexual activity before age 18 (National Statistical Office, 2011).

Social scientists and medical researchers have long been interested in discovering what determines individual perceptions of health risk, particularly because those perceptions are important prerequisites for behavior change in health behavior models (e.g., Ajzen and Fishbein, 1980). In the area of health and HIV/AIDS research, a number of factors have been shown to be associated with an individual's HIV risk perception, including some demographic characteristics, schooling, past sexual practices, self-efficacy, living with chronically ill family members, and so on (Barden-O'Fallon et al., 2004; Macintyre et al., 2004; Prohaska et al., 1990). Another line of recent research has suggested that communication and interactions with network partners such as relatives, friends, and acquaintances are important determinants of individuals' perceptions of HIV infection risk. For example, studies find that individuals who have network partners with high perceived risk of contracting HIV are themselves more likely to worry about contracting HIV (Bühler and Kohler, 2003; Smith and Watkins, 2005; Smith, 2003). However, these studies contain an important caveat that the findings may not be interpreted as causal. More specifically, these studies fail to account for the selection of individual's network partners, by making the assumption that one's network is randomly formed, which does not reflect reality. However, two studies account for social network endogeneity (i.e., the non-random selection of individuals into their networks) by using individual fixed effects and instrumental variables strategies. Similar to the evidence from other studies, they find that in Malawi and Kenya, network partners' risk perceptions of contracting HIV have an impact on individuals' own perceptions, even after accounting for selection into individual's social networks (Helleringer and Kohler, 2005; Kohler et al., 2007).

However, the findings from these previous studies are prone to measurement biases. They use the respondent's *perceptions* of network partners' HIV infection expectations, not network partners' *actual* reports of such expectations. The relevant social psychological bias includes projection bias, also known as the false consensus effect (Ross et al., 1977). Projection bias occurs when an individual "projects" his or her own beliefs and behaviors onto others, unconsciously assuming that most people think and behave like him or her. So, individuals may report their own HIV infection expectations twice: once for individual expectations and once for network partners' expectations. This would clearly inflate estimates of the effect that peers' expectations have on individual expectations.

In addition, to my knowledge, there exists no research that has focused on the role of social interactions and networks in the transmission of HIV infection expectations among adolescents. As described above, given that many adolescents in sub-Saharan Africa remain sexually inexperienced during adolescence and HIV prevalence among adolescents are substantially lower than adults, the way in which adolescents construct their own HIV risk perceptions may be different from adults. For example, having little or no past sexual experience to base their HIV infection expectations on, adolescents may rely more heavily on external sources of information. In particular, peers often serve as the most critical source of information and influence, as adolescents grow independent from parents and spend an

increasing amount of time with peers (Giordano, 2003). Moreover, since uncertainty and anxiety about the future (e.g., unrealistic fatalism and uncertainty about HIV status) among adolescents are shown to have long-term detrimental effects on educational attainment, health, and fertility decisions (Duke et al., 2002; Nguyen et al., 2012; Trinitapoli and Yeatman, 2011), understanding critical determinants of HIV infection risk overestimation among adolescents is crucial for guiding policies to foster human capital accumulation and improve health in sub-Saharan Africa.

In this study, I aim to establish credible evidence of peer influence on HIV risk expectations by separating selection and confounding effects from peer influence effects using a quasi-experimental framework. Moreover, in order to avoid potential biases resulting from using respondents' *perceptions* of network partners' HIV infection expectations that previous studies use, I use expectations reported directly by peers. Although no existing research reveals gender heterogeneity in the effects of network partners' HIV risk perceptions, given substantial gender differences in friendship formation and social interactions among adolescents (Crosnoe, 2000; Maccoby, 1990; Moore, 1990), I investigate whether peer effects in HIV risk perceptions vary by gender. Finally, adolescents who have appropriate knowledge of HIV/AIDS may be less reliant on peers' HIV infection expectations than those with less knowledge, so I stratify the results based on level of knowledge about HIV/AIDS.

Conceptualizing and Estimating Peer Effects

Manski (1993) distinguishes among the following three types of social effects: endogenous effects, contextual effects, and correlated effects. Endogenous effects occur when an individual's propensity to behave in some way varies with the behavior of the reference group. In the case of adolescents' HIV infection expectations, endogenous effects can occur if an individual is likely to have a higher expectation of HIV infection if his peers also have high expectations—that is, if their expectations are interdependent. Contextual effects (or exogenous effects) occur when an individual's propensity to behave in some way varies according to the exogenous characteristics of the reference group. Lastly, correlated effects occur when individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments. For example, a similar level of HIV infection expectation among students attending the same school may be due to their exposure to a shared school environment (e.g., HIV/AIDS education).

Among these categories of peer effects, this study focuses on identifying endogenous peer effects because they entail potential social multiplier effects (or spillover effects), which occur when policy and other interventions are multiplied through social interactions. While empirically isolating endogenous peer effects from contextual and correlated effects is crucial for obtaining more credible estimates of peer influence, there are several critical challenges. First, peer groups (assumed to be classmates in this paper) can be endogenously determined by parental choices (e.g., residential choices, private school enrollment), such that estimated peer effects may be attributable to the similarity among individuals rather than peer influences. Second, potential unobserved environmental characteristics that influence individual and peer group choices simultaneously (i.e., correlated effects) may confound the peer effect, leading to biased estimates of the peer effects. A third empirical

challenge is separately examining bi-directional effects in social relationships. Since peer behavior affects individual-level behavior and vice versa, a standard linear empirical model of social interactions using an average measure of behavior by a reference group suffers from identification problems (termed the “reflection problem” by Manski (1993)).

In order to address these methodological challenges and accurately quantify endogenous peer effects, this paper uses an instrumental variables/fixed effects methodology. First, I employ an instrumental variable strategy to overcome the identification problem caused by the simultaneity of group- and individual-level decisions. The validity of the instrumental variable strategy in this study critically depends on school fixed effects. School-level fixed effects will control for the main source of selection into schools, such that the instrument variable would be uncorrelated with school-level characteristics. Additionally, controlling for school-level fixed effects eliminates unobserved school-level factors that are shared by all individuals within the same school, thus minimizing correlated effects. In order to assess the validity of the specifications, I conduct balancing tests to show that the instrument of the study is plausibly quasi-random across cohorts, within schools. In addition, following Lavy and Schlosser (2011), I use falsification tests using placebo treatments to confirm that estimated peer influence is not driven by any short-term effects of school-level unobserved confounders.

2. Methods

2.1. Data

The present study uses school-based survey data collected between October 2011 and March 2012 as part of an HIV/AIDS prevention program for Malawian adolescents, supported by Daeyang Luke Hospital in Malawi, the Korea International Cooperation Agency (KOICA) and Africa Future Foundation in Korea. The target population of the baseline survey was all students in Grade 1 to Grade 3 (equivalent to grades 9 and 11 in the U.S. education system) who were initially enrolled in 33 public secondary schools in major traditional authorities (TA) within Lilongwe District: Chimutu, Chitukula, Kalumba, and Tsabango in 2011. The survey consists of two sets of questionnaires: 1) an in-class self-report survey and 2) a private interview with HIV Testing and Counseling (HTC). In the survey, students answered questions on basic demographic characteristics, household assets, health, labor market participation, school performance, time/risk preference, HIV-related knowledge and behaviors, and attitudes towards HIV/AIDS. After the in-class survey, each student was confidentially interviewed about sexual experiences by a same-sex enumerator. The questionnaires and interviews were conducted in Chichewa, the dominant language in the study site and Malawi's national language. Approximately 80% of initially enrolled secondary school students ($n = 7,971$) participated the survey. Those who did not answer key questions on individual- and family-level characteristics were dropped, leaving the analytic sample of 7,910.

2.2. Variables

The dependent variable of this study is an individual's probabilistic expectation of being infected with HIV. Since the endogenous peer effects are of particular interest in this study,

the independent variable is the mean probabilistic expectation of HIV infection among classmates.

Collecting data on expectations has been known to be difficult in developing countries because respondents often have limited formal education and are unfamiliar with the basic properties of probabilities (Attanasio, 2009). However, recent studies have found that people in developing countries can generally understand and answer probabilistic questions, suggesting that measuring probabilistic expectations in such contexts is indeed feasible and valuable (Delavande et al., 2011; Paula et al., 2013; Trinitapoli and Yeatman, 2011). To increase the reliability of responses to probabilistic expectations, researchers can adapt to local contexts and ask the questions differently. Using some practical examples (e.g., rainfall expectations) or visual aids (e.g., a pile of beans or stones) may be helpful in explaining the concept of probability (Delavande and Kohler, 2009). In the school-based survey for this study, a similar context-sensitive effort was made to collect reliable data on HIV infection expectations. During a thirty-minute pre-interview training session, a ruler graded from 0 to 100 was used to help students conceptualize their expectations in probability terms. Then, students were asked about their perceived chances of contracting HIV.

In the empirical models, I include a set of individual- and peer-level control variables. To control for individual-level characteristics, I include gender, age, ethnicity, grade level, family acceptance of respondent's hypothetical HIV infection, parental education, asset-based wealth index, parental death, co-residence of both parents, and family members' death due to HIV/AIDS. Peer-level characteristics include percent female, percent Yao people, percent other ethnicity, percent of parents who completed secondary school, mean wealth index, percent with any parent dead, percent living with both parents, number of sampled classmates, and mean family acceptance of respondent's hypothetical HIV infection. Table 1 presents the summary statistics for the sample. Gender differences in male and female samples are shown in Table A1 in the supplementary files.

2.3. Analytic strategy

The primary empirical specification for this study is the linear-in-means model of social interactions (Manski, 1993):

$$Y_{igs} = X_{igs}B + \bar{X}_{-igs}\delta + W_s\theta + a\bar{Y}_{-igs} + \varepsilon_{igs} \quad (1)$$

where Y_{igs} is the subjective probability of HIV infection of individual i in grade g in school s , individual and family characteristics are contained in a vector X , and classmate characteristics are measured as grade-level averages of the X vector excluding the individual, labeled $\bar{X}_{-igs}$. In all results, the error terms, ε_{igs} , are clustered at the school level. A vector of school dummies, W_s , eliminates unobserved school-level characteristics or confounding factors shared by all individuals within the same school. These school-level fixed effects should control for some important environmental factors that may influence both peers' and individuals' HIV infection expectations (e.g., local HIV prevalence). Importantly, I find that approximately 50% of variation in the peer measure of interest and the instrumental variable is within school (Table A2 in the supplementary files), which validates the strategy of using within-school variation for these grademate variables. By

contrast, a substantial portion of variation in sociodemographic composition variables is between schools. Finally, $\bar{Y}_{igs}$ is the grade-level average outcome excluding the individual (i.e., grade-level average of probabilistic expectations of being infected by HIV). The main coefficient of interest is the endogenous effect, α , which demonstrates the extent to which individuals are influenced by their peers' HIV infection expectations.

As noted above, due to the “reflection problem,” α in Eq. (1) is not identified without using instrumental variables or other methods (Manski, 1993). To identify endogenous peer effects, following Fletcher (2011, 2010), I use a combined instrumental variables/fixed effects approach. The instrument used in this study is the proportion of classmates with family members who died of HIV/AIDS. This is based on responses to the question “Do you have [immediate or extended] family members (e.g., parents, children, grandchildren, uncles, cousins, etc.) who have already died of HIV/AIDS?” Eighteen percent of respondents reported having family members who died of HIV/AIDS (Table 1).

I estimate the instrumental variable estimator through a two-stage least square framework. In the first stage, classmates' mean subjective probability of HIV infection are predicted as a function of instruments and the exogenous variables of the main equation. In the second stage, the predicted value from the first stage replaces the peer-group variable (i.e., $\bar{Y}_{igs}$ in Eq. (1)) in the main least square regression. In order to be compelling instruments, these variables must meet two criteria: (1) they must be strongly correlated with classmates' mean subjective probability of HIV infection, and (2) they must be validly excluded from a respondent's own subjective probability of HIV infection (exclusion restriction). The F-statistic in the first-stage regression (reported in the result tables) demonstrates that the instrument is strongly linked with classmates' mean subjective probability of HIV infection, so the first criterion is met. The second criterion is untestable, though the intuition for the exclusion restriction is that the instrumental variable is assumed to affect an individual's subjective probability of HIV infection only through its effect on the classmates' mean subjective probability of HIV infection.

Importantly, in this study the validity of the instrument depends on the use of school fixed effects. To argue for the excludability of the instrument, I assume that the distribution of classmates with family members who died from HIV/AIDS is quasi-random across grades *within* schools. This means that the proportion of classmates with a family member who succumbed to HIV may be highest in Grade 1 cohort in one school while it is highest in Grade 3 cohort in another school. While it is likely that students with family members who died of HIV/AIDS attend different schools than students without such an experience, using school fixed effects relies on variation in the proportion of classmates whose family members died of HIV/AIDS within the same school. Unlike students in more developed countries, classes and grades in less developed countries like Malawi may not be as clearly differentiated by age. Classes could be more endogenous because students are more likely to repeat a grade and have more flexibility in leaving school and then returning later (when they are older) to the same grade. Therefore, balancing tests are necessary to establish evidence that the instruments are quasi-exogenous.

In Panel A of Table 2, balancing tests provide evidence against the validity of the instrument without school fixed effects. For example, the instruments are shown to be strongly related to parental socioeconomic status (i.e., education, wealth, and occupational status) (Panel A), suggesting that the instrument might be related to unmeasured family characteristics that might also be related to the outcome. On the other hand, Panel B shows that by controlling for school fixed effects, the instrument is not correlated with any individual- or family-level characteristics. This suggests that the proportion of family members' death due to HIV can be treated as plausibly exogenous (or quasi-random) within schools and parents are not systematically changing student's schools based on these cohort characteristics (Bifulco et al., 2011; Lavy and Schlosser, 2011). Therefore, the instrumental-variable model can be validated only when it is used with school fixed effects, which eliminate selection effects.

One may still be concerned about the violation of the exclusion restriction assumption: students could feasibly learn about a death in their classmate's family, which may in turn affect their own HIV expectations directly, rather than through peers' HIV expectations. However, students report the deaths of immediate and extended family members (73% of those who reported family members' HIV death have both parents alive (not shown)) and regressions control for peers' parental death at the school-grade level, so the instrument is likely to capture distant family members' HIV death. It is reasonable to assume that classmates would not necessarily know about these deaths and will only be influenced via the effect of these deaths on their peers' HIV infection expectations. It is, however, still possible that there are other uncontrolled channels through which the instrument may affect an individual's HIV infection expectations. In order to test the sensitivity of the instrumental variable estimates to deviations from the exclusion restriction, I apply the procedure suggested by Conley and colleagues (2012) that produces valid inferential statements without imposing the assumption that the instrument exactly satisfies the exclusion restriction requirement.

The IV estimates in this study are interpreted as local average treatment effects (LATE) (Imbens and Angrist, 1994). For the LATE interpretation to be valid, a monotonicity condition must be satisfied, namely, that everyone's HIV infection expectations move in the same direction—no student actually decreases (rather than increases) his or her HIV infection expectations in the presence of peers having higher expectations. This monotonicity assumption would be incorrect if people heavily rely on non-medical beliefs about HIV (e.g., cultural or religious beliefs). For example, some people may believe that God decides who gets HIV regardless of own protection strategies (Bunnell et al., 2005). Or if people believe that there would be a fixed number of HIV-positive people (the number is capped), they would think that increased number of HIV infection around them actually reduces their own risk of having HIV. Although the monotonicity assumption is not empirically testable, evidence from the Malawi Demographic and Health Survey (DHS) shows that secondary school students have a relatively high level of knowledge about HIV prevention and low level of misconceptions about HIV transmission. This suggests that they are likely to rely on medically based beliefs about HIV (National Statistical Office, 2011).

3. Results

3.1. Main regression results

Table 3 presents the estimates of peer endogenous effects from the baseline ordinary least squares (OLS) regression, the instrumental variable (IV) approach, and the combined instrumental variables/fixed effects approach (IV/FE) for the pooled sample of males and females. The full results are available in the supplementary files. In Column 1, although the result from OLS regression shows that the coefficient for endogenous peer effects is positive and highly significant, it is hard to interpret since the coefficient is likely biased due to the reflection problem, selection, and correlated effects. The two-stage least square results are presented in Column 2. In this model, the coefficient for endogenous peer effects is not statistically significant, though its magnitude remains similar to the OLS coefficient. However, one may be concerned that the instruments are correlated with characteristics that predict selection into the chosen school. In Column 3, to account for potential selection effects and to eliminate school-level unobservables that simultaneously influence individual and grade-level HIV infection expectations, school fixed effects are controlled.

In the IV/FE model (Column 3), the reported first-stage F-statistic is nearly 11, which exceeds the recommended cutoff of an F-statistic of 10 (Staiger and Stock, 1997). The results show that after controlling for school fixed effects, the coefficient of interest increases from 0.44 in the OLS model and 0.42 in the IV model to 0.65, suggesting that the estimates obtained from OLS and IV without school fixed effects were relatively underestimated. Although the Hausman tests provide no statistical evidence on the differences between the estimates of endogenous peer effects across specifications (results not shown), based on the aforementioned critical disadvantages of OLS and IV-only models and strong empirical support from the balancing tests, I favor the IV/FE specification over the other two. To interpret the coefficient of endogenous peer effects in the IV/FE model, increasing the mean probabilistic HIV expectation among classmates by one percentage point will increase an individual's own HIV infection expectation by an average of 0.65 percentage points. (Note that within-school variation in classmates' mean subjective probability of HIV infection is approximately two percentage points)

In order to assess the validity of the results, I conduct falsification tests where grade g 's peer measures of subjective probability of HIV infection are replaced with grade $g-1$ or $g+1$ measures within the same school (this is a "placebo treatment"). Table 4 reports the results of a falsification check on the preferred specification (i.e., IV/FE model in Table 3). Column 1 presents baseline results for comparison. Columns 2 and 3 estimate the same specification while replacing the actual measures of peer-level subjective probability of HIV infection with measures from the wrong grade in the same school. If unobserved contextual factors are driving the relationships reported in Column 1, then results using the false peer measures should be similar to the actual peer measures. Instead, I find that using the false peer measures leads to null findings (Columns 2 and 3), suggesting that the preferred estimates in Column 1 are unlikely to be driven by unobserved contextual factors.

I also present the results from alternative specifications in Table A3 in the supplementary files, where I use dichotomized measures of the perceived probability of being infected by

HIV (15%, 25%, 35%, and 45%). These results also support evidence for the peer effects in HIV infection expectations. As the cut-off increases, the peer effect on an individual's HIV expectation increases.

Lastly, in order to assess the robustness of the IV estimates, I adopt the bounds approach proposed by Conley and colleagues (2012) to draw inferences that are valid if the instrumental variable does not perfectly satisfy the exclusion restriction assumption. For the sake of brevity, the results are reported for 95% confidence intervals and detailed descriptions of the procedure are discussed in the supplementary files. The results from this sensitivity check suggest that even if peers' family members' dying of HIV/AIDS had a relatively large direct impact on an individual's HIV infection expectations, the IV estimates of endogenous peer effects remain positive and statistically significant for both the pooled and male samples. In sum, the IV estimates of endogenous peer effects are very robust, even to substantial departures from a perfect IV.

3.2. Gender analysis

Next, I stratify the results by gender to consider whether the peer effect on HIV infection expectation differs for males and females. Panel A of Table 5 shows interesting divergent patterns by gender. The OLS estimates for males and females do not seem to be very different (0.47 vs. 0.40) (Columns 1 and 4), but, in the preferred model where an IV/FE model is used, the estimated coefficient for peer influence dramatically increases to 1.18 among males (Column 3), whereas the coefficient for females decreases to 0.14 (Column 6). Statistical tests (using seemingly unrelated estimation) confirm the differences in the IV/FE estimates of endogenous peer effects between male and female students ($p = 0.077$, not shown in the table). These differential gender patterns suggest that compared with the preferred specification of IV/FE, OLS underestimates the peer effect among males but overestimates the peer effect among females. The increased peer estimates among male adolescents in IV and IV/FE models imply that endogenous peer effects become more important when accounting for the "reflection problem," which is consistent with the peer effects literature using similar methodologies (Lundborg, 2006).

In this initial analysis of gender differences, only the outcome was stratified by gender; the independent variable, peers' mean subjective probability of HIV infection, still relied on pooling male and female responses. This assumes that unlike adults, adolescents (particularly in school settings) could have more social interactions with the opposite sex through casual relationships and school activities (e.g., lunch, soccer, etc.) (Tavory, 2014). In other words, the assumption is that male and female students are equally likely to respond to male and female peers' expectations. However, evidence from qualitative interviews shows that social networks tend to be gendered among young adults in Malawi (Kohler et al., 2007; Moore, 1990; Watkins and Warriner, 2003). This suggests the possibility that adolescents respond only to the influences of same-sex peers. To test this, I separate classmates by gender to create separate average peer HIV expectations for males and females, and then estimate the effect on individual expectations separately by gender.

Basically, the findings of the analysis in Panel B are qualitatively similar to Panel A: peers' HIV infection expectations affect an individual's expectations among males, but not

females. The results in Panel B of Table 5 suggest that endogenous peer effects are positive and statistically significant among male students (Column 3), though the magnitude is smaller than in Panel A. Although the coefficient of endogenous peer effects among female students is negative and statistically insignificant, the weak first-stage result makes it hard to draw any firm conclusions of the female results (Column 6). Lastly, I also examine whether there are cross-gender peer effects in HIV infection expectations (Table A4 in the supplementary files). I find no statistically significant endogenous peer effects for both male and female students.

Eligible possible sources of gender differences in endogenous peer effects may include gender-based differential selection into schools, differential causal peer effect, and maybe something else. In an attempt to examine the possibility that gender differences may be attributable to gender-based differential selection into schools, I predict the probability of adolescents' being in schools with high mean subjective probability of HIV infection. In short, I find that boys are more highly selected into better schools in terms of school-level HIV infection expectations as well as other socioeconomic characteristics (main effect of gender), whereas parental resources protect girls from selecting into bad schools (interaction effect between mother's education and gender) (Table A5 in the supplementary files). These results suggest that gender-based differential selection seems to be a reasonable partial explanation of gender differences in endogenous peer effects, though I cannot quantify how much of the differences are attributable to selection.

3.3. Results stratified by levels of HIV/AIDS knowledge

In order to gain a sense of whether adolescents with more complete knowledge of HIV/AIDS are less reliant on peers' HIV infection expectations, I stratify the results based on level of knowledge about HIV/AIDS. The level of knowledge about HIV/AIDS was measured based on the respondent's total number of correct answers to 20 HIV/AIDS-related questions (Table A6 in the supplementary files). The average number of correct answers is approximately 17 out of 20 questions. The sample is divided into two sub-groups: a high-level group (>17) and a low- or medium-level group (≤ 17). Table 6 demonstrates that those with a low- or medium-level of HIV/AIDS knowledge are influenced by peers' HIV risk perceptions, whereas the peer effects are smaller and statistically insignificant among those with a higher level of HIV/AIDS knowledge (Columns 1 and 2). This contrast is even larger among male adolescents (Columns 3 and 4).

4. Discussion

In this study, I use a social interactions framework and a unique sample of Malawian adolescents to examine whether adolescent HIV infection expectations are influenced by peers' HIV infection expectations. The major findings of this study are that peers have significant and substantial causal effects on an individual's HIV risk perceptions. The results indicate that a one-percentage-point increase in mean probabilistic expectation of HIV infection among classmates increases an adolescent's own subjective probability by an average of 0.65 percentage points. This finding suggests that Malawian adolescents' overestimation of their own probability of HIV infection appears to be partly due to peer

expectations. Falsification tests support the validity of the results, suggesting that the estimates are not short-term effects of unobservable confounders at the school level.

Gender-stratified models show that endogenous peer effects on HIV infection expectations are larger among male adolescents, suggesting that males are more sensitive to peers' HIV infection expectations. These results are inconsistent with those of Kohler and colleagues (2007), which show that among Malawian adults the influences of social networks on HIV risk perception do not differ by gender. By contrast, similar to the current study, Godlonton and Thornton (2012) find that having neighbors who learn HIV results influences males' probability of learning HIV results more than women. Gender differences in peer effects may be corroborated by evidence that men are more likely to talk about their sexual experiences and the implications for HIV infection (Kaler, 2004; Watkins, 2004). In addition, gender differences may be also attributable to the distinct ways in which male and female adolescents form friendships and interact with peers. For example, male adolescents tend to interact loosely with a broad group of peers, whereas female adolescents prefer more intense relationships with a few close friends (Crosnoe, 2000; Kim, 2015; Maccoby, 1990).

The key finding of this study, that an individual's expectation of HIV infection is influenced by peers' expectations, could be relevant to policy. Policies that help individual adolescents may have spillover effects on untreated adolescents through social interactions. This is especially true for policies that target adolescents' knowledge and attitudes about HIV/AIDS. For example, providing more detailed and complete information on HIV/AIDS, especially with respect to transmission and prevention of HIV (e.g., providing information on differential distributions of HIV infection by age and gender (Dupas, 2011)), may improve risk assessment and encourage behavior change among adolescents. As a result of interventions like this, adolescents (particularly those who are not sexually experienced) may lower their HIV risk perceptions by internalizing the HIV/AIDS information. To the extent that adolescents revise their HIV risk perceptions based on their peers' expectations, the policy effect will be multiplied through social interactions.

While this paper is able to address many of the relevant econometric issues in estimating peer effects, there are several limitations. First, it is well known that school-based samples are not representative of the general population in less developed countries, including Malawi, where the net enrollment rate (NER) for secondary school was 31.2 percent in 2013 (World Bank, 2015). Therefore, the sample used in this study is comprised of a relatively advantaged group of adolescents. Despite the limited external validity of this study, a positive educational gradient in HIV prevalence in most sub-Saharan African countries, including Malawi (Fortson, 2008), suggests that the sample used in this study is potentially of considerable academic and policy interest. Nonetheless, future research should consider how peers might influence HIV risk perceptions among out-of-school adolescents. In fact, peer effects in HIV infection expectations among non-students may differ from peer effects among students since early transition to sexual activity and marriage is more common among adolescents who do not attend schools (National Statistical Office, 2011), and their peers are more likely to be diverse in terms of demographic and socioeconomic characteristics.

Second, causal inference can be threatened when using the instrumental variables strategy if the instrument fails to satisfy the exclusion restriction assumption. Admittedly, there is no way of knowing for sure whether this assumption is met. There are several ways that the HIV-related death of a peer's family member can affect an individual's HIV risk perceptions. For example, if a classmate's immediate family member dies of HIV/AIDS, an individual may attend the funeral, and it could be the salience of this death that affects his or her own expectations of HIV infection. Given the high HIV prevalence in the country, even if one does not know for certain that the death was caused by HIV/AIDS, it might affect subjective perceptions of the prevalence in the community. Although controlling for parental death at the school-grade level suggests that the instrument mostly reflects distant family members' death due to HIV/AIDS (and classmates are less likely to know the victim directly), this only reduces, and does not eliminate, the concern that the exclusion restriction assumption may be violated. However, the procedure proposed by Conley and colleagues (2012) to check for the sensitivity of IV estimation results when this restriction is relaxed yields highly stable estimates for a wide range of relaxations, which suggests that the instrument is valid.

Despite these limitations, this study makes several contributions to the literature. Overall, this paper provides rigorous evidence of causal social network effects among adolescents in a developing country where such evidence is scarce. The results suggest that peers' HIV risk perceptions significantly affect an individual's own expectation of having HIV, providing a possible explanation for Malawian adolescents' tendency to overestimate their chances of HIV infection. The finding that peers influence HIV infection expectations among males more than they do for females is new and intriguing, and could be relevant for policy. This paper adds significantly to previous research by using novel data from a large school-based survey in Malawi and actual HIV risk perceptions reported by peers. The results are valuable both for increasing our understanding of how peers influence HIV risk perceptions and for guiding policies that focus on lowering uncertainty about HIV status and the future among adolescents.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Ajzen, I.; Fishbein, M. Understanding attitudes and predicting social behaviour. Prentice-Hall, Inc; Englewood Cliffs, New Jersey: 1980.
- Anglewicz P, Kohler H-P. Overestimating HIV infection: The construction and accuracy of subjective probabilities of HIV infection in rural Malawi. *Demogr Res.* 2009; 20:65–96.10.4054/DemRes.2009.20.6 [PubMed: 19672478]
- Attanasio OP. Expectations and Perceptions in Developing Countries: Their Measurement and Their Use. *Am Econ Rev.* 2009; 99:87–92.10.1257/aer.99.2.87
- Barden-O'Fallon JL, deGraft-Johnson J, Bisika T, Sulzbach S, Benson A, Tsui AO. Factors associated with HIV/AIDS knowledge and risk perception in rural Malawi. *AIDS Behav.* 2004; 8:131–40.10.1023/B:AIBE.0000030244.92791.63 [PubMed: 15187475]
- Bifulco R, Fletcher J, Ross S. The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health. *Am Econ J Econ Policy.* 2011; 3:25–53.

- Bühler C, Kohler H-P. Talking about AIDS. *Demogr Res Special*. 2003; 1:397–438.10.4054/DemRes.2003.S1.13
- Bunnell RE, Nassozi J, Marum E, Mubangizi J, Malamba S, Dillon B, Kalule J, Bahizi J, Musoke N, Mermin JH. Living with discordance: knowledge, challenges, and prevention strategies of HIV-discordant couples in Uganda. *AIDS Care*. 2005; 17:999–1012.10.1080/09540120500100718 [PubMed: 16176896]
- Conley TG, Hansen CB, Rossi PE. Plausibly Exogenous. *Rev Econ Stat*. 2012; 94:260–272.10.1162/REST_a_00139
- Crosnoe R. Friendships in Childhood and Adolescence: The Life Course and New Directions. *Soc Psychol Q*. 2000; 63:377–391.
- Delavande A, Giné X, McKenzie D. Measuring subjective expectations in developing countries: A critical review and new evidence. *J Dev Econ*. 2011; 94:151–163.10.1016/j.jdeveco.2010.01.008
- Delavande A, Kohler H-P. Subjective expectations in the context of HIV/AIDS in Malawi. *Demogr Res*. 2009; 20:817–874.10.4054/DemRes.2009.20.31 [PubMed: 19946378]
- Duke NN, Borowsky IW, Pettingell SL, Skay CL, McMorris BJ. Adolescent early death perception: links to behavioral and life outcomes in young adulthood. *J Pediatr Health Care*. 2002; 25:224–34.10.1016/j.pedhc.2010.03.004 [PubMed: 21700137]
- Dupas P. Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya. *Am Econ J Appl Econ*. 2011; 3:1–34. [PubMed: 22199993]
- Fischhoff B, Parker AM, Bruine De Bruin W, Downs J, Palmgren C, Dawes R, Manski CF. Teen expectations for significant life events. *Public Opin Q*. 2000; 64:189–205.10.1086/317762 [PubMed: 10984333]
- Fletcher J. Peer influences on adolescent alcohol consumption: evidence using an instrumental variables/fixed effect approach. *J Popul Econ*. 2011; 25:1265–1286.10.1007/s00148-011-0365-9
- Fletcher J. Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects. *Health Econ*. 2010; 484:466–484.10.1002/hec [PubMed: 19382102]
- Fortson JG. The gradient in sub-Saharan Africa: socioeconomic status and HIV/AIDS. *Demography*. 2008; 45:303–22.10.1353/dem.0.0006 [PubMed: 18613483]
- Giordano PC. Relationships in Adolescence. *Annu Rev Sociol*. 2003; 29:257–281.10.1146/annurev.soc.29.010202.100047
- Godlonton S, Thornton R. Peer effects in learning HIV results. *J Dev Econ*. 2012; 97:118–129.10.1016/j.jdeveco.2010.12.003 [PubMed: 22081739]
- Helleringer S, Kohler H-P. Social networks, perceptions of risk, and changing attitudes towards HIV/AIDS: new evidence from a longitudinal study using fixed-effects analysis. *Popul Stud (NY)*. 2005; 59:265–82.10.1080/00324720500212230
- Imbens GW, Angrist JD. Identification and estimation of local average treatment effects. *Econometrica*. 1994; 62:467–475.
- Jamieson PE, Romer D. Unrealistic fatalism in U.S. youth ages 14 to 22: prevalence and characteristics. *J Adolesc Health*. 2008; 42:154–60.10.1016/j.jadohealth.2007.07.010 [PubMed: 18207093]
- Kaler A. AIDS-talk in everyday life: The presence of HIV/AIDS in men's informal conversation in Southern Malawi. *Soc Sci Med*. 2004; 59:285–297.10.1016/j.socscimed.2003.10.023 [PubMed: 15110420]
- Kim J. School Socioeconomic Composition and Adolescent Sexual Initiation in Malawi. *Stud Fam Plann*. 2015; 46:263–279.10.1111/j.1728-4465.2015.00029.x [PubMed: 26347090]
- Kohler H, Behrman J, Watkins S. Social networks and HIV/AIDS risk perceptions. *Demography*. 2007
- Lavy V, Schlosser A. Mechanisms and Impacts of Gender Peer Effects at School. *Am Econ J Appl Econ*. 2011; 3:1–33. [PubMed: 22199993]
- Lundborg P. Having the wrong friends? Peer effects in adolescent substance use. *J Health Econ*. 2006; 25:214–33.10.1016/j.jhealeco.2005.02.001 [PubMed: 15964090]
- Maccoby EE. Gender and relationships: A developmental account. *Am Psychol*. 1990

- Macintyre K, Rutenberg N, Brown L, Karim A. Understanding perceptions of HIV risk among adolescents in KwaZulu-Natal. *AIDS Behav.* 2004; 8:237–50.10.1023/B:AIBE.0000044072.71361.b3 [PubMed: 15475673]
- Manski CF. Identification of Endogenous Social Effects: The Reflection Problem. *Rev Econ Stud.* 1993; 60:531.10.2307/2298123
- Millstein SG, Halpern-Felsher BL. Judgments about Risk and Perceived Invulnerability in Adolescents and Young Adults. *J Res Adolesc.* 2002; 12:399–422.10.1111/1532-7795.00039
- Moore G. Structural Determinants of Men's and Women's Personal Networks. *Am Sociol Rev.* 1990; 55:726–735.
- National Statistical Office. Malawi Demographic and Health Survey 2010. Calverton, MD: 2011.
- Nguyen QC, Hussey JM, Halpern CT, Villaveces A, Marshall SW, Siddiqi A, Poole C. Adolescent expectations of early death predict young adult socioeconomic status. *Soc Sci Med.* 2012; 74:1452–60.10.1016/j.socscimed.2012.01.006 [PubMed: 22405687]
- Paula Á, Shapira G, Todd P. How beliefs about hiv status affect risky behaviors: Evidence from malawi. *J Appl Econ.* 201310.1002/jae
- Prohaska T, Albrecht G, Levy J. Determinants of self-perceived risk for AIDS. *J Health Soc Behav.* 1990; 31:384–394. [PubMed: 2135939]
- Ross L, Greene D, House P. The “false consensus effect”: An egocentric bias in social perception and attribution process. *J Exp Soc Psychol.* 1977; 13:279–301.
- Smith KP. Why are they worried? Concern about AIDS in rural Malawi. *Demogr Res Special.* 2003; 1:279–318.10.4054/DemRes.2003.S1.9
- Smith KP, Watkins SC. Perceptions of risk and strategies for prevention: responses to HIV/AIDS in rural Malawi. *Soc Sci Med.* 2005; 60:649–60.10.1016/j.socscimed.2004.06.009 [PubMed: 15550312]
- Staiger D, Stock J. Instrumental Variables Regression with Weak Instruments. *Econometrica.* 1997; 65:557–586.
- Tavory I. The situations of culture: humor and the limits of measurability. *Theory Soc.* 2014:275–289.10.1007/s11186-014-9222-7
- Trinitapoli J, Yeatman S. Uncertainty and Fertility in a Generalized AIDS Epidemic. *Am Sociol Rev.* 2011; 76:935–954.10.1177/0003122411427672 [PubMed: 22536003]
- Watkins S. Navigating the AIDS epidemic in rural Malawi. *Popul Dev Rev.* 2004; 30:673–705.
- Watkins SC, Warriner I. How do we know we need to control for selectivity? *Demogr Res.* 2003; 9:109–142.10.4054/DemRes.2003.S1.4
- World Bank. [11.23.15] World Bank Education Statistics Database [WWW Document]. 2015. URL <http://data.worldbank.org/country/malawi>

Table 1

Summary statistics, full sample (n = 7,910)

	Mean	SD	Min	Max
<i>Individual-level characteristics</i>				
Subjective probability of HIV infection	16.04	12.92	0.0	100.0
Age	16.15	1.84	11.0	29.0
Ethnicity				
Chewa	0.50	0.50	0.0	1.0
Yao	0.07	0.26	0.0	1.0
Other ethnicity	0.42	0.49	0.0	1.0
Grade				
Grade 1	0.29	0.46	0.0	1.0
Grade 2	0.39	0.49	0.0	1.0
Grade 3	0.31	0.46	0.0	1.0
Family acceptance of respondent's hypothetical HIV infection	3.84	1.24	1.0	5.0
Parent's secondary education completed	0.50	0.50	0.0	1.0
Asset-based wealth index	0.00	1.00	-2.0	2.2
Any parent dead	0.24	0.43	0.0	1.0
Living with both parents	0.54	0.50	0.0	1.0
Family members' death due to HIV/AIDS	0.18	0.38	0.0	1.0
HIV/AIDS knowledge	17.28	1.76	8.0	20.0
<i>Classmate-level characteristics</i>				
Mean subjective probability of HIV infection	16.04	2.67	8.8	24.2
% Classmates with family members' death due to HIV/AIDS	0.18	0.05	0.1	0.4
% Female	0.50	0.08	0.2	0.7
% Chewa	0.50	0.23	0.2	1.0
% Yao	0.07	0.04	0.0	0.2
% Other ethnicity	0.42	0.20	0.0	0.7
% Parent's secondary education completed	0.50	0.19	0.0	0.8
Mean wealth index	-0.00	0.59	-1.0	1.0
% Any parent dead	0.24	0.06	0.1	0.4

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

	Mean	SD	Min	Max
% Living with both parents	0.53	0.09	0.2	0.7
Number of sampled classmates	101.38	55.64	20.0	277.0
Mean family acceptance of respondent's hypothetical HIV infection	3.84	0.25	2.5	4.4
Observations	7910			

Table 2
Balancing tests for the association between cohort-level instruments and individual characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Female	Age	Chewa	Yao	Self-reported academic performance	Parent's secondary education completed	Asset-based wealth index	Father's high occupational status	Living with both parents	Family acceptance of respondent's hypothetical HIV infection
Panel A.										
% Classmates with family members' death due to HIV/AIDS	-0.179 (0.187)	3.709* (1.477)	1.414** (0.513)	-0.214* (0.089)	0.213 (0.384)	-0.749+ (0.397)	-3.266** (1.187)	-0.625** (0.210)	0.298 (0.181)	0.342 (0.762)
F-test	0.916	6.306	7.597	5.782	0.308	3.559	7.571	8.858	2.711	0.201
F-test p-value	0.3448	0.0173	0.0096	0.0226	0.5838	0.0684	0.0097	0.0056	0.1088	0.6562
N	7948	7948	7936	7936	7926	7945	7931	7943	7948	7948
R-squared	0.002	0.178	0.020	0.002	0.014	0.006	0.028	0.005	0.003	0.000
Panel B.										
% Classmates with family members' death due to HIV/AIDS	0.145 (0.182)	0.659 (0.851)	-0.106 (0.104)	0.010 (0.100)	0.649 (0.431)	0.305 (0.211)	0.109 (0.239)	-0.133 (0.183)	0.023 (0.143)	0.793 (0.811)
F-test	0.635	0.600	1.039	0.010	2.267	2.089	0.208	0.528	0.026	0.956
F-test p-value	0.4323	0.4443	0.3137	0.9214	0.1418	0.1588	0.6513	0.4719	0.8704	0.3359
N	7948	7948	7936	7936	7926	7945	7931	7943	7948	7948
R-squared	0.016	0.292	0.209	0.020	0.039	0.126	0.339	0.054	0.029	0.008

Additional controls for Panel A: complete set of grade dummies.

Additional controls for Panel B: complete set of grade dummies and school dummies.

Robust standard errors are clustered at the school level.

The F-statistics are for the effect of % Classmates with family members' death due to HIV/AIDS.

+ $p < 0.1$.

Kim

100.0 > d

'10.0 > d
**
'50.0 > d
*

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3
Analysis of peer effect on subjective probability of HIV infection, OLS, IV, and IV/FE results, full sample

Sample Specification	(1)		(2)		(3)	
	Full	OLS	Full	IV	Full	IV/FE
<i>Endogenous peer effect</i>						
Mean subjective probability of HIV infection	0.439*** (0.081)		0.420 (0.277)		0.650*** (0.137)	
N		7910		7910		7910
R-squared		0.075		0.075		0.053
F-statistic (First-stage)				5.748		10.885

Robust standard errors are clustered at the school level.

Additional controls not presented here but included in the regression include: gender, age, ethnicity, grade level, family acceptance of respondent's hypothetical HIV infection, parental education, asset-based wealth index, parental death, co-residence with both parents, family members' death due to HIV, percent female, percent Yao people, percent other ethnicity, percent of parents who completed secondary school, mean wealth index, percent with any parent dead, percent living with both parents, number of sampled classmates, and mean family acceptance of respondent's hypothetical HIV infection.

+ $p < 0.1$,
* $p < 0.05$,
** $p < 0.01$,
*** $p < 0.001$

Table 4

Falsification tests for subjective probability of HIV infection

Sample Specification	(1)	(2)	(3)
	Full IV/FE Baseline	Full IV/FE Grade above	Full IV/FE Grade below
<i>Endogenous peer effect</i>			
Mean subjective probability of HIV infection	0.650*** (0.137)		
g+1 Mean subjective probability of HIV infection		-0.471 (0.309)	
g-1 Mean subjective probability of HIV infection			0.063 (0.475)
N	7910	5435	5582
R-squared	0.053	0.060	0.063
F-statistic (First-stage)	10.885	7.393	3.385

Robust standard errors are clustered at the school level.

All models include the same set of controls as in Table 3.

Column 1 presents baseline results. Columns 2 and 3 show results of the placebo treatment, when the cohorts are rotated forward (2nd graders' mean subjective probability of HIV infection replaced by 3rd graders') and backward (2nd graders' mean subjective probability of HIV infection replaced by 1st graders'). 3rd graders and 1st graders are dropped in Columns 2 and 3, respectively.

+ $p < 0.1$,
* $p < 0.05$,
** $p < 0.01$,
*** $p < 0.001$

Table 5
Analysis of peer effect on subjective probability of HIV infection, OLS, IV, and IV/FE results, by gender

Panel A.		(1)	(2)	(3)	(4)	(5)	(6)
Sample	Male	Male	Male	Male	Female	Female	Female
Specification	OLS	IV	IV/FE	OLS	IV	IV/FE	IV/FE
<i>Endogenous peer effect</i>							
Mean subjective probability of HIV infection	0.466*** (0.119)	0.851 ⁺ (0.437)	1.180** (0.418)	0.396** (0.120)	-0.056 (0.504)	0.143 (0.462)	
N	3949	3949	3949	3961	3961	3961	3961
R-squared	0.069	0.066	0.035	0.068	0.062	0.051	0.051
F-statistic (First-stage)		5.608	10.157		5.670	11.622	
Panel B.		(1)	(2)	(3)	(4)	(5)	(6)
Sample	Male	Male	Male	Female	Female	Female	Female
Specification	OLS	IV	IV/FE	OLS	IV	IV/FE	IV/FE
<i>Endogenous peer effect</i>							
Mean subjective probability of HIV infection among same-sex classmates	0.296** (0.093)	0.480** (0.183)	0.575* (0.224)	0.349** (0.108)	-0.264 (1.090)	-0.185 (1.045)	
N	3949	3949	3949	3961	3961	3961	3961
R-squared	0.068	0.067	0.041	0.068	0.054	0.052	0.052
F-statistic (First-stage)		11.100	5.173		1.596	1.142	

Robust standard errors are clustered at the school level.

All models include the same set of controls as in Table 3.

Note that the instrument in Panel B is different than the instrument in Panel A because in Panel B the proportion of classmates with family members who died of HIV/AIDS is calculated separately by gender, while in Panel A men and women are pooled.

⁺ $p < 0.1$,

^{*} $p < 0.05$,

^{**} $p < 0.01$,

^{***} $p < 0.001$

Table 6

Analysis of peer effect on subjective probability of HIV infection, OLS, IV, and IV/FE results, by levels of HIV/AIDS knowledge

Sample	(1)	(2)	(3)	(4)	(5)	(6)
HIV/AIDS knowledge level	Full	Full	Male	Male	Female	Female
Specification	High	Low/Medium	High	Low/Medium	High	Low/Medium
	IV/FE	IV/FE	IV/FE	IV/FE	IV/FE	IV/FE
<i>Endogenous peer effect</i>						
Mean subjective probability of HIV infection	0.387 (0.372)	0.901 *** (0.265)	0.461 (0.702)	1.541 * (0.684)	0.283 (0.535)	0.037 (0.666)
N	4003	3907	2031	1918	1972	1989
R-squared	0.055	0.057	0.046	0.043	0.060	0.055
F-statistic (First-stage)	9.503	12.050	8.153	11.647	11.317	11.525

Robust standard errors are clustered at the school level.

All models include the same set of controls as in Table 3.

+ $p < 0.1$,

* $p < 0.05$,

** $p < 0.01$,

*** $p < 0.001$