South African Journal of Economics Vol. 76:S1 May 2008

LINKING POVERTY AND INCOME SHOCKS TO RISKY SEXUAL BEHAVIOUR: EVIDENCE FROM A PANEL STUDY OF YOUNG ADULTS IN CAPE TOWN

TARYN DINKELMAN, DAVID LAM AND MURRAY LEIBBRANDT*

Abstract
Is there a link between household income and income stress, and risky sexual behaviour of young people? Anecdotal and qualitative evidence suggests this may be the case, but there is little quantitative research measuring this relationship. We use two waves of new data from the Cape Area Panel Study to investigate this link for 2,993 African and coloured youths aged 14 to 22 in 2002. In the process, we discuss one type of research design that could allow for a causal interpretation of the effect of income poverty on HIV risk. This design plausibly separates out the effect of income stress from the effect of living in a poor household by comparing behaviours across households with and without negative economic shocks, conditional on baseline income. Our results indicate that females in poorer households are more likely to be sexually active in 2002 and more likely to sexually debut by 2005. In addition, girls in households experiencing negative economic shocks are more likely to reduce condom use between 2002 and 2005. However, they are less likely to have multiple partners in 2002 or have transitioned to multiple partners by 2005. Males who experienced a negative shock are more likely to have multiple partners. Despite the tight research design for assessing shocks, the findings on the impacts of shocks do not generate clear recommendations for policy. There appears to be no systematic difference in condom use at last sex by household income levels or income shocks.
JEL Classification: I12, I32, J13
Keywords: Youth, sexual behaviour, household poverty, household shocks, community poverty

1. INTRODUCTION

Poverty and infectious disease have always been linked. The conditions of poverty are some of the very conditions in which diseases of cholera, tuberculosis and malaria thrive. Ill health in turn reinforces aspects of poverty by undermining labour capabilities and eroding human capital potential. At the end of the 20th century, HIV/AIDS emerged as the newest communicable disease to become part of this cycle of poverty. 63% of the 38.6

* Taryn Dinkelman and David Lam are at the University of Michigan, Murray Leibbrandt is at the University of Cape Town. A first version of this paper was written for the workshop “A Symposium for investigating the empirical evidence for understanding vulnerability and the associations between poverty, HIV infection and AIDS impact”. The authors thank USAID and HEARD (Health Economics and HIV/AIDS Research Division) for financial support during preparation of the paper and attendance at the workshop. Support for this research was provided by the U.S. National Institute of Child Health and Human Development (R01HD045581), the Fogarty International Center of the U.S. National Institutes of Health (D43TW000657), the Andrew W. Mellon Foundation and the Research Chairs Initiative of the Department of Science and Technology and National Research Foundation. The data used in this paper are publicly accessible at http://www.caps.uct.ac.za.

million people living with HIV live in sub-Saharan Africa, the world's poorest region, and in many countries the poorest individuals are most likely to be infected.[1]

However, the ways in which poverty heightens vulnerability to this disease, and the effects of HIV on poverty, are still not well understood. This paper addresses the first part of the gap: do conditions of income poverty increase behavioural vulnerability to HIV? Can we understand through what channels income may have an effect?

Poverty may raise the probability of contracting HIV in several ways: malnutrition can increase susceptibility to any disease,[2] poverty-related lack of education and information may be a barrier to individuals changing their behaviours, and specific sexual behaviours adopted by poor individuals in poor communities may directly increase vulnerability. We focus on this last channel, and look at the behaviours of new entrants into the sexual market: youths aged 14 to 22 years. We ask whether higher household incomes are associated with less risky behaviours for individuals (particularly females) at a point in time and over time, and whether there is a positive own income effect on number of sex partners (particularly for males).[3] We look for evidence of any relationship between community poverty levels and risky sexual behaviour. We also ask whether these associations can be claimed as causal. On the way to answering these questions, we describe one type of research design that could plausibly pin down such causal links.

Identifying a causal link between income and HIV risk is difficult for several reasons. First, very few data sets include detailed information on both sexual behaviours and household income. For example, the widely used Demographic Health Surveys do measure sexual and reproductive health behaviours, but not income. Living Standards Measurement Surveys that cover many developing countries typically measure income, but not sexual behaviour. Second, most of the data sources in which household assets and behaviour can be correlated are cross-sectional. Drawing conclusions about current sexual behaviour based on current income measures may lead to spurious conclusions about causality. This is especially the case for adults who have been in the sexual market for longer and may have experienced important disadvantage earlier in life. Third, unobserved individual and/or household characteristics are likely to be correlated with measures of sexual behaviour as well as with incomes. Such correlation will bias estimates of the effect of resources on sexual behaviours.

In a review of the usefulness of panel data Woolard and Leibbrandt (2006) explain that, such data provides the most promising possibility to address – if not completely solve – each of these issues. In a case such as this one, panel data gives the researcher observations on incomes and sexual behaviours of individuals at a point in time as well as observations on how these both change over time. It is this that gives panel data its analytical bite. We use new panel data from the Cape Area Panel Study (CAPS) that surveys individual youths aged 14 to 22 in Cape Town, South Africa in 2002 and again in 2005 (Lam *et al.*, 2006). Youths are particularly important for public health policy: before entering the sexual market, they represent the new generation that begins with no

[1] Data are from http://www.avert.org/worldstats.html. See Fenton (2004) for a background piece on links between poverty and HIV/AIDS.

[2] See Stillwaggon (2002) for a discussion of how poverty undermines health and thus heightens HIV/AIDS susceptibility.

[3] Household income is admittedly a crude indicator of poverty. We have experimented with household assets to proxy for the poverty of the household, and results are qualitatively the same.

HIV incidence. Preventing new infections among youths is arguably a top priority. Using the CAPS data, we investigate whether each of three self-reported measures of sexual behaviour – sexual debut, number of recent partners and lack of condom use at last sex – is affected by household income constraints and income shocks. We compare 2002 behaviours across youths in rich and poor households, and in households that have negative economic shocks, conditional on 2002 income levels. In addition, we consider whether individual transitions in sexual behaviours between 2002 and 2005 are related to lagged household income levels and lagged shocks.

Since CAPS covers new entrants into the sexual market, we will be most concerned with confounding effects arising from household level unobservable characteristics rather than individual unobservable characteristics. We argue that variation in economic shocks across households more closely approximates a random assignment of income than variations in household income alone. In principle then, the effect of these shocks is less likely to be confounded by unobservable features of households. In our data, households in all income quintiles experience economic shocks over the period related to death, illness and job loss. We find some evidence that these shocks may be a good way to isolate the effect of income on behaviour. However, it is not possible to separately identify whether risky behaviours are more likely due to the financial consequences of an economic shock, or because of other instabilities that occur after such a shock.

Our findings indicate significant improvements in some behaviours among young people observed between 2002 and 2005: condom use rises substantially for females and multiple partner incidence falls for both males and females. These changes seem to cross all age groups.

In all of the models, the effect of income should be interpreted as the combined effect of income and household type on behavioural choices. Out of the three behaviour measures that we consider, only sexual debut (for females) and multiple partners (for males) are significantly correlated with household income and negative economic shocks. Household income protects females (but not males) from early sexual debut, negative economic shocks increase the probability of early debut (for females), and higher levels of household income are associated with an increase in partners (for males). Where an individual lives also predicts sexual behaviour – males and females in high poverty communities are much more likely to be sexually active in 2002 and males in poorer communities are significantly more likely to report multiple partners in 2002. The sexual debut results are of particular interest, given what has been documented about the positive relationship between early sex and subsequent HIV status.

2. RELATED LITERATURE

Anecdotal and qualitative evidence suggest that poverty increases the risks of contracting HIV through the channel of more risky behaviour, and does so especially for youth and for women. Safe sex may be expensive for practical reasons; it can also imply welfare loss if economic resources provide compensation for the risk of unprotected sex. For example, small focus group surveys of young people (13-25 years) in Khutsong, South Africa revealed that the price of condoms was somewhat of a deterrent to safe sex. In addition, the young adults in this study agreed that a common reason for young women to have sex is economic. “Participants spoke of young women and females in Khutsong who engage in sexual relationships in exchange for lifts home from school, gifts and subsistence cash"; however, the understanding was that while the women expect gifts in the relationship, "they are not viewed as a source of income" (MacPhail and Campbell 2001).

Leclerc-Madala (2002) conducted qualitative research in another South African urban area (Durban), and concluded that "For women, and most especially young unmarried women, sexuality is conceptualised as a resource that can be drawn upon for material or economic advantage". Then, in a literature review of 45 qualitative (small focus group studies) and quantitative studies, Luke (2003) finds that age and economic asymmetries are common in sub-Saharan African, and that these asymmetries are associated with unsafe sexual behaviours and increased risk of HIV infection. She notes that the evidence on why women and young females search for older partners is mixed. Some studies report that females seek older men as marriage partners because they are more stable and more likely to support children. Other studies report that females depend on financial receipts from older partners during times of economic crisis or emergency only. Interestingly, several studies point out aspects of relationships that even young females have control over such as the start of new relationships and the duration of relationships. This agency does not appear to extend to aspects of safe sex within a particular relationship.

The evidence from this and other qualitative research reinforces the following picture: social culture facilitates the acceptance of multiple sex partners, especially for men. For women living in poverty, economic reliance on men (older or not) for long-term support or during transitory shocks can lead to risky partnerships and risky behaviour within partnerships.[4] On most of these issues, there is almost no convincing quantitative evidence that poverty is related to HIV risk in these direct, behavioural ways. This is largely related to the sensitive and detailed nature of the data required to show these links, as well as the difficulties of convincingly isolating a causal link between poverty and HIV risk.

There is some work that correlates measures of poverty and measures of sexual behaviour or actual disease-prevalence at an aggregate level. The South African National HIV prevalence, HIV Incidence, Behaviour and Communication Survey (2004) collected household survey data and individual information on risk behaviours and HIV status from a national sample of people. The report by the HSRC (Shisana *et al.*, 2005) documents much higher HIV prevalence rates in the poorest urban and rural informal areas. Youth (aged 15-24 years) and women in these poor areas also report the highest number of recent sexual partners and the lowest incidence of condom use. These patterns certainly suggest that vulnerability is higher in poorer areas.

Recent research conducted using Demographic and Health Survey data provides some conflicting evidence on the relationship between poverty and HIV status. In Kenya, the wealthiest women are almost three times more likely to be HIV positive than the poorest women (Johnson and Way, 2003).[5] In an unpublished study using DHS data from a

[4] See Caldwell (2000) for a discussion of other risk factors for HIV that have been part of African society for a long time. These include polygamy, high rates of migration, skewed sex ratios in cities, early age of sexual debut and the high value of fertility.

[5] Testing non-response for the entire sample was 19%. The authors note that "the effect of non-response on the representativeness of the Kenya DHS HIV data has been undertaken and is available from the authors. In short, it was found that non-response to the survey did not significantly bias the prevalence estimates".

range of African countries, Donnelly (2006) reports on the findings that in almost all countries, the wealthiest people have highest HIV prevalence rates.[6] Together, these findings suggest that we do not fully understand the channels linking poverty to HIV risk. While poor people may exhibit more risky behaviours, HIV prevalence is apparently not uniformly higher in poorer communities across Africa.

There are a handful of papers that attempt to more formally model the market for unsafe sex, both within a commercial sex work context and within informal non-marital relationships. All of these papers focus on linking condom use to financial resources of male partners. Rao *et al.* (2003) estimate the compensating differential for safe sex among commercial sex workers in Calcutta, India. They exploit variation from a natural experiment in which sex workers were randomly exposed to HIV/AIDS information at different times. Early access to information is used as an instrumental variable (IV) for condom use. The authors find that sex workers who always use condoms incur large losses of over 60% per sex act. Their IV strategy is one way to get around the endogeneity of condom use, and their focus is on the difference in transfers that results after acts of protected versus unprotected sex.

In a related paper, Gertler *et al.* (2005) investigate the compensating differential for risky sex amongst sex workers in two Mexican states. They use panel data to estimate sex worker-fixed effects models of the price of unprotected sex and, therefore, they compare the price that the same sex worker charges to different clients for sex with and without a condom. They measure the risky sex premium at 23%, and 46% for more attractive sex workers. Their empirical strategy looks across partnerships for the same woman to deal with the endogeneity of condom use, and their focus is also on the size of the transfer in relation to the type of sex.

Luke (2006) takes this idea into an informal market. Using her own data on men aged 21-45 years in Kisumu, Kenya, she presents convincing evidence that there is a market for risky sex even in informal relationships. By comparing the use of condoms (her measure of safe sex) by the same man across different female partners with different levels of transfers using a male-fixed effect model, she finds that partnerships involving larger transfers are strongly and significantly associated with lower probabilities of safe sex. She does not find a difference in this trade-off for younger compared to older women, and concludes that a market for risky sex exists for all women who are willing to trade off higher incomes for greater disease and pregnancy risk. This is somewhat in contrast to prevailing sentiment that young women are the group most vulnerable to sexual exploitation in conditions of poverty. Luke (2006) suggests that the reason why women are willing to make such a trade off may be related to poverty and to the consumption demands of these women. Her paper concludes with the call for more evidence to be marshalled on this point.

Our paper weighs in on this issue directly by asking the question: to what extent is household income poverty and income instability related to risky sexual behaviours of new entrants into the sexual market? We broaden the focus beyond condom use, and also consider the impact of resources on the timing of sexual debut and the likelihood of multiple partners among young adults. We are particularly interested in the timing of sexual debut for these young people, as several researchers have established that early age

[6] The study on which this article was based is not yet in the public domain.

of sexual debut is strongly predictive of later HIV status (Gregson *et al.*, 2002; Pettifor *et al.*, 2004).

3. ECONOMETRIC MODEL

(a) Identifying Income Effects

To obtain a clean measure of the effect of income on sexual risk behaviours of young adults, we would like to randomly assign high or low income to households and then observe differences in behaviours across these households. Such an experiment would eliminate the positive feedback effects (endogeneities) that make it difficult to separate out causality between income, poverty and HIV risk. However, this experiment is not easily found or undertaken in the real world. We discuss three models that one might want to estimate (and which we do estimate) bearing in mind the nature of the experiment we would like to run. We point out which of the models comes closest to this ideal, and what can be concluded when we do not attain this ideal research design.

Suppose we have a cross section of data on individuals i in households j. To relate *per capita* household income I_j to sexual risk behaviours y_{ij}, we could specify the following model:

$$y_{ij} = \alpha_0 + \alpha_1 X_{ij} + \alpha_2 I_j + \gamma_j + \varepsilon_{ij} \tag{1}$$

where X_{ij} is a set of individual-level covariates (age, education, race), γ_j is an unobserved household fixed effect and γ_{ij} is a random error term. γ_j includes all aspects of a particular household that influence y_{ij} and I_j which the econometrician cannot measure. For example, γ_j could be parental ability that is negatively related to the behavioural risk measure y_{ij} (more able adults may raise children better able to avoid risk behaviours) and positively related to household income (more able adults earn more in the labour market). We will refer to γ_j as household type.

An OLS regression of (1) would compare behaviours of young adults in high-income households to young adults in low-income households. However, high- and low-income households are likely to be systematically different in many unobserved respects. If γ_j really is parental ability, we would expect the positive correlation between γ_j and I_j to bias the OLS estimator of α_2 upwards. There are other stories for γ which would imply a downward bias. For example, ambitious parents may earn lots of money but not have time to supervise kids. All we can conclude from estimating the model in (1) is that $\hat{\alpha}_2$ is the total effect of household income and household type on youth sexual behaviours.

A second strategy is to compare sexual behaviours of young adults in households with and without negative economic shocks, conditional on household income level. Formally, the model becomes:

$$y_{ij} = \alpha_0 + \alpha_1 X_{ij} + \alpha_2 I_j + \alpha_3 SHOCK_j + \gamma_j + \varepsilon_{ij} \tag{2}$$

where $SHOCK_j$ is a dummy variable for a negative household income shock that occurs before the survey date. As long as there is sufficient variation in the incidence of shocks across households and these shocks are not correlated with household type γ_j or with household income I_j then α_3 will be identified even if α_2 is biased.

The assumption that low-income households are as likely to experience a negative income shock as high-income households may be a strong one. We would expect that some types of shocks are more likely to occur in richer households (*e.g.* loss of a job, failure of a business) while others could be more likely in poor households (*e.g.* a serious illness). In the next section, we show that although the incidence of some shocks does vary across household income quintiles, there is less systematic variation than one might expect. There are enough shocks of similar type in rich and poor households to compare shock and non-shock households. However, it is not clear that negative shocks also have no direct relationship with γ_j which is what we need to consistently estimate α_3.

If some households are more prone to economic shocks and these are also the households where children are getting less parental guidance, then $\hat{\alpha}_3$ will be a biased estimate of α_3. Households with low values of γ_j will have high values of $SHOCK_j$, and so the omitted household effect biases the coefficient on the shock variable upwards. In this case, we must again interpret the coefficient as the combined effect of household type and economic shock on behaviours.

To test whether these economic shocks are a clean measure of an income shock or if they are picking up some of the effect of γ_j on behaviours, we check to see whether risky behaviour measures are significantly related to future shocks in the household, conditional on the prior shock. If there is a significant relationship between behaviour in period t and a shock that occurs after time t, this is some evidence that the households in which youths exhibit risky sexual behaviours are also simply prone to economic shocks.[7]

Specifically, we estimate the following model:

$$y_{ijt} = \alpha_0 + \alpha_1 X_{ijt} + \alpha_2 I_{jt} + \alpha_3 SHOCK_{jt} + \alpha_4 SHOCK_{jt+1} + \gamma_j + \varepsilon_{ijt} \quad (3)$$

where $SHOCK_{jt+1}$ is a shock in a future period. We should find a significant coefficient on the prior shock and not on the future shock variable in order to support the claim that shocks are not correlated with γ_j. If this claim stands up to this test, we can be more confident that this research design brings us closer to the ideal experiment than simply comparing behaviours across rich and poor households.

Since the effects of poverty are probably transmitted over time, we would ideally like to measure the effect of household income on future changes in behaviours: sexual debut, increasing number of recent partners or decreasing condom use. Using the panel nature of the CAPS data, we can begin to do that. The model in (1) can be altered to measure the change in sexual behaviour over the 3-year period as these changes relate to baseline levels of household income:

$$(y_{ijt+1} - y_{ijt}) = \alpha_0 + \alpha_1 X_{ijt} + \alpha_2 I_{jt} + \gamma_j + \varepsilon_{ijt} \quad (4)$$

where the dependent variable is now measuring a change in behaviour. If correctly specified, this model would compare transitions in young adult behaviour across high- versus low-income households, when this income is measured early on in a child's life. α_2 would identify the effect of living in a low-income household on later transitions in sexual behaviours. The problem remains that while γ_j is correlated with sexual transitions, it is also likely to be correlated with lagged household income and perhaps with economic shocks. OLS estimation of model (4) will still provide a biased estimator of α_2. We can include prior period shocks into (4) and, under the assumption that shocks are

[7] This is similar to a test of the orthogonality assumption applied to economic shock variables in Duryea *et al.* (2007).

uncorrelated with γ_j we could recover the effect of the shock on behaviour. The next CAPS wave is not yet complete and so we cannot do the false experiment using future shocks, as we can for the 2002 data. In the interim, findings from the model run in levels will be used to guide the interpretation of the coefficients in the transitions models.[8]

For each transition model that we estimate, we condition on a particular behaviour in 2002. For example, to model sexual debut, we estimate the above regression for individuals who had not yet had sex by 2002; to model an increase in the number of sex partners, we restrict to the sexually active youths who have 0 or 1 partner only in 2002; to model an increase in risky sex (no condom at last sex), we restrict to the set of people who reported using a condom at last sex in 2002. Because the samples over which we estimate the transition probits are different from the samples used in the levels equations, the results are not directly comparable. However, if the levels equation is picking up a true relationship between poverty and vulnerability (whether this is due to income or some combination of income and household type) we should expect to see similar signed coefficients in the transition results, especially if some of the effects of poverty are transmitted over time.

In estimating the model in levels or in changes, we will also include measures of community poverty obtained from the 2001 Census. The inclusion of these variables will tell us whether behaviours vary across neighbourhoods with a larger proportion of poor households, but will not explain why because neighbourhoods with high poverty rates are also neighbourhoods with a number of other social problems that we are not measuring.

(b) Measurement Error

A final issue that arises in studying reported sexual behaviour is measurement error in the dependent variable. Misreporting is more likely, the more sensitive the survey question. While there is little evidence of systematic differences in reports of sexual debut, age of sexual debut, condom use or age of partner across men and women, there is evidence of a discrepancy in reports of number of sex partners. Smith (1992) observes that, in a set of developed countries, it is standard to find men reporting a higher number of sexual partners than women. Nnko *et al.* (2004) report the same finding for a closed population in rural Tanzania. If the sample is not closed, these discrepancies could be due to non-coverage of some populations, non-response or misreporting.

It is important to note that the 2005 wave of CAPS was carefully constructed to try to minimise the biases in responses to these sensitive questions. For the 2005 survey, respondents were asked a series of questions about each of their 10 most recent relationships – once these questions were completed, the two facing pages of the survey were sealed together. Respondents were able to fill in answers to the survey by hand if they preferred and seal the pages on completion. Out of all applicants, 14% chose this option. Across those who did and who did not respond themselves, there was no systematic difference in the number of sex partners reported in 2005 – although the gap between male and female reports in number of sex partners does fall when the survey is self-

[8] Note that even a household fixed effects model would not eliminate the omitted variables bias issue, as changes in income within a household over time are also likely to be correlated with household type. This would be equivalent to allowing the error term to have three components: $\gamma_j + \gamma_j^* t + \varepsilon_{ijt}$. In addition, sibling fixed effects will not identify the effect of income on behaviour, because 2002 income is the same for the siblings living in the same household in this year.

administered.[9] There was no systematic difference in any of the sexual behaviour variables across those who did and did not self-report answers in this section.

One way to check for consistency of reports by the same person is to compare two reports of the total number of sex partners given in CAPS 2005. Youths were first asked "With how many different people have you had sexual intercourse in your whole life?" In a later section, youths were asked to record detailed information about their 10 most recent sexual partners. We summed the total number of partners with information and regressed the total reported partners on the summed measure of partners for those individuals with 10 or fewer total lifetime partners.[10] If reports were consistent and there was no measurement error, we should have found a zero constant term and a slope coefficient close to 1. However, as can be seen in the results contained in Table A1 in the Appendix, we find a significantly positive constant term for both females and males, and a significant slope coefficient of less than 1 for each group. The constant is larger for males, indicating that they are reporting more total partners than they provide information on. There was no difference in these results across individuals who self-reported and those who responded to the interviewer.

This is clear evidence that at least one of the number of partner measures is incorrect. Since we cannot tell which one is more correct, we try to avoid severe measurement error issues by collapsing the multiple partners variable into 0 (if zero or one partner) and 1 (if more than 1 partner). We do this and re-estimate regressions for the two measures of multiple partners in 2005. The results are shown in Table A1 in the Appendix and it can be seen that the constant is statistically insignificant and the slope coefficient is over 0.88 (and not statistically different from 1). Using the information on multiple partners in this way means our results are less likely to be driven by reporting biases in the number of sex partners.

4. DATA AND VARIABLE DESCRIPTIONS

The CAPS follows the lives of a large and representative sample of adolescents in Cape Town as they undergo the multiple transitions from adolescence to adulthood. The study began in 2002, with almost 5,000 young people between the ages of 14 and 22 being interviewed as part of wave 1 of CAPS. In 2003 and 2004, most of these young adults (YA's) were re-interviewed in waves 2A and 2B of the project. In 2005, the entire sample was targeted for re-interview in wave 3 of the survey. CAPS covers a range of aspects of adolescence, including schooling, entry into the labour market, sexual and reproductive health, and experiences within families and households. Most data are collected from the young people themselves, but data are also collected from parents and other older household members (Lam *et al.*, 2006). This paper uses data from waves 1 (2002) and 3 (2005) of the survey and variables pertaining to individual sexual and reproductive health and household incomes and income shocks.[11]

[9] Results not reported here. See Tourangeau and Smith (1996) for a discussion of how self- versus other-administered questions about number of sex partners elicits different responses in the USA.

[10] Less than 1% of the sample reported more than 10 lifetime partners.

[11] We do not use data on what young people know about HIV/AIDS. Anderson and Beutel (2006) report that levels of HIV/AIDS knowledge are very high in the 2002 CAPS data.

Table 1. Composition of CAPS 2002 (wave 1) and 2005 (wave 3) sample

Population group	Number of households	Number of young adults	Unweighted %	Weighted %	% YA's followed up in 2005	Number of YA's with completed hh surveys
2002						
Black/African	1,443	2,151	46	28		2,151
Coloured	1,399	1,980	42	53		1,980
White	445	593	13	19		
Total	3,287	4,724	100	100		4,131
2005						
Black/African	1,043	1,519	43	27	71	1,410
Coloured	1,221	1,664	47	59	84	1,583
White	260	336	10	14	57	
Total	2,524	3,519	100	100		2,993

Note: Weighted is calculated using youth weights which correct for sample design, household and youth non-response in 2002.

Table 1 shows the breakdown of the original sample interviewed in wave 1 with weighted and unweighted percentages, as well as response rates for these young adults in 2005. Note that the survey design oversampled African youths (Lam *et al.*, 2006). Therefore, it is important to use weights in most analyses to correct for this over sampling. Almost two thirds of the (weighted) sample are coloured, just over one quarter is African and about 14% are white. The total number of YA's with completed individual questionnaires in 2002 was 4,724. In the 2005 wave, 3,519 of these individuals were re-surveyed, although not all of these youths had completed household surveys in each year.

Initial non-response as well as attrition between waves was very high for white youths, with only 57% of the originally surveyed White sample being re-interviewed in wave 3. Since the white sample is small to begin with, and attrition rates are substantially higher, we only examine behaviours for African and coloured youth in this paper. Attrition among the coloured and African subsamples was much lower; 84% of coloured youths and 71% of Africans did not attrit in the 3 years between the first and third wave. Table A2 in the Appendix provides results for a probit model predicting attrition between waves 1 and 3 for the African and coloured youth. Older youths are more likely to not be found in wave 3, as are youths with initially higher levels of education and living in higher income households.[12]

The basic sample for the remainder of the paper is the set of African and coloured youths with matched and completed household and individual surveys in both waves. There are 2,993 individuals living in just over 2,000 households in this group. For each individual, we use information on basic demographics (age, sex, education and race), sexual behaviours and relationship histories and household level variables including *per capita* income and the presence of negative economic shocks. In both waves, young adults answered questions on sexual and reproductive health. Data were collected on whether the youth had sexually debuted, at what age this occurred, the number of different sex partners in the 12 months prior to the survey year, the total number of sex partners, and condom and contraceptive use at first and last sex. We use this information in a single year (2002) and construct measures of transition using the differences in 2002 and 2005 reports.

[12] The variables significantly correlated with attrition are many of the same variables that predict initial non-response in the first wave. See Lam *et al.* (2006) for a discussion of first wave non-response.

The set of sexual relationship variables used as outcomes in the probit models are: SEXEVER (whether the individual had sex by 2002), MULTIPLE (the number of partners in the year before 2002, with 0 representing none or 1 partner and 1 representing more than 1 partner), NOCONDOM (no condom reported at last sex in 2002), ΔSEXEVER (=1 if the youth becomes sexually active between 2002 and 2005, =0 if remains not sexually active), ΔMULTIPLE (=1 if the youth increase sex partners from 0 or 1 in 2002 to more than 1 in 2005, =0 if the number of sex partners remains 0 or 1 in 2005); ΔCONDOM (=1 if the youth reports condom use at last sex in 2002 but not in 2005, =0 if youth reports condom use at last sex in both years).

In each year, a knowledgeable respondent reported data on total monthly household income as a continuous number or in a bracket. Income amounts were imputed for those households who responded in brackets, using the midpoint of the bracket. The log of *per capita* household income was calculated using household size reported in that year. For households with no income information, we constructed a dummy variable if the data was missing and assigned these households a 0 for the log of household *per capita* income.[13] Whites and coloureds are significantly more likely to record missing income, as are individuals in the Western Cape, in urban areas, and people under 30. Missing income values were high for those with higher level occupations (conditional on working), but overall were higher for the unemployed compared to the employed. There is something of a conventional wisdom that wealthier households are more likely to not report an income value, although this is difficult to verify in practice.

Households also provided information on negative economic shocks in the 2 years prior to 2002, and in the years between the 2002 and 2005 survey. Questions were asked about specific shocks: death, job loss, loss of a grant or loss of other financial support in the household. The month and year of these shocks was recorded as well as the main coping mechanisms employed by the household. We examine the incidence of shocks below, and focus on those that appear to hit a majority of households: death, job loss and loss of a grant or other financial support in the household. We code the variable $SHOCK_t = 1$ if the household reported at least one of these shocks in the survey year *t*, 0 otherwise. It is possible that households can insure against income loss due to death shocks, and that some shocks may free up income for other purposes. However, interpreting these shocks as negative income shocks seems reasonable for these households – in 2005, over 70% of households experiencing one of these shocks reported that the negative impact was moderate or large.[14]

A final set of variables that we include are community poverty rates and youth test scores from a math and literacy test administered in the first wave. The community poverty rates are computed as the proportion of households living below the poverty line (R9,600 per household per year) in 2001, calculated from 2001 Census data at the sub-place level. These poverty rates may capture some neighbourhood effects on behaviour, although we cannot unpack what these specific effects might be with these data. The test score measures are included to proxy for individual ability that plausibly

---

[13] It is not unusual to find a high percentage of missing income values in household surveys in South Africa. Ardington *et al.* (2006) report that 16% of individuals (as opposed to households) have missing income values in the 10% sample of the South African 2001 Census.

[14] The question was not asked in 2002.

affects the future prospects of young adults. They also probably capture some of the effect of living in households with more able and involved parents.

5. RESULTS

(a) Descriptive Statistics

Table 2 describes the individual and household characteristics of young adults in the matched sample. Means are weighted, and each observation is an individual. The household level means are interpreted as the proportion of young adults living in households of this type rather than the proportion of households in the sample with this feature.

The average age of the sample in 2002 is 17.8 years, youths have on average 9.3 years of education and Africans constitute about 30% of the (weighted) sample. On average, youths live in sub-places where about one in four households are in poverty. This masks wide variation in actual poverty rates across communities.

In 2002 the log monthly *per capita* household income was 6.2 (about R533) on average. This average masks large variance across and within the African and coloured YAs with one standard deviation either side spanning R133 to R1340. The mean of the log of *per capita* household income does not change between years (2005 values are not reported in the table), although the proportion of households without income values increases from 4% to almost 20% of the sample. For all of our analysis, we will use income reported in 2002 rather than 2005 and so will not be concerned with the high rate of non-response to the income questions in 2005.

Table 2. Descriptive statistics for the sample of CAPS panel respondents

X-variables	**Mean**	**Std. Dev.**
Female	0.518	
Age in 2002	17.785	(2.47)
African	0.304	(0.46)
Years of education in 2002	9.269	(2.12)
Biological mother in house in 2002	0.764	
Biological father in house in 2002	0.480	
Mean test score (z-score) in 2002	−0.015	(0.89)
Log monthly per capita household income (Rands) in 2002	6.278	(1.00)
Per capita household income missing in 2002	0.048	
Community poverty rate from 2001 Census	0.252	(0.16)
Any shock between 2000-2002	0.359	
Any shock between 2002-2005	0.265	
N individuals	2,993	
N households	2,183	

Notes:

1. All means are weighted by the 2002 individual youth weight.
2. The sample consists of African and Coloured young adults who were observed in 2002 and 2005 with completed questionnaires.
3. Female, African, biological mother and father in house, missing income and shock variables are coded as dummy variables.
4. Community poverty rate is the fraction of the Census sub-place earning below an annual poverty line, as measured in the 2001 South African Census.
5. Households experience a shock if there is a death, job loss, loss of a grant or loss of other financial support in the household.

Table 3. Percentage of CAPS respondents in three categories of sexual behavior, 2002 and 2005

Age	African female		African male		Coloured female		Coloured male	
	2002	2005	2002	2005	2002	2005	2002	2005
Ever had sex								
17–18	60.4	71.6**	59.2	64.6	22.2	30.7**	35.9	37.4
19–20	84.4	87.8	79.5	87.8*	52.1	56.3	61	68.8
21–22	88	96.4***	86.7	88.1	70.7	63.4	67.5	77.2*
Total	76.5	86.8***	74.9	80.1*	44.7	51.1**	51.7	61.8***
N	511	529	418	411	524	594	451	534
Condom use at last sex								
17–18	58.9	77.3***	75.3	79.9	30.5	33.7	74.3	82.8
19–20	51.5	70.5***	76.5	85	28.1	40.7*	74.9	61.9**
21–22	41.4	67.3***	68.8	87.5***	17.9	30.5**	55.9	63.3
Total	50.4	71.7***	72	84.7***	24.9	33.5**	68.2	65.3
N	371	420	296	293	221	272	214	294
Multiple sex partners in last year								
17–18	23.1	12.7*	56.8	39.6**	16.3	4.7*	48	25.1**
19–20	20.5	7.2***	52.5	41.9	4.7	2.4	48	26.9***
21–22	22.3	12.2**	63.3	31.9***	8.9	5.6	43.2	19.2***
Total	21.8	10.8***	56.7	36.5***	7.7	3.7*	47.8	24.5***
N	339	406	270	277	194	246	186	253

1. Asterisks indicate significant level for test of equality between 2002 and 2005 percentages: *p < 0.1, **p < 0.05, ***p < 0.01.
2. Sample includes all African and coloured respondents in 2002 who were followed in 2005. Statistics are calculated over the set of individuals aged 17–18 (for example) in 2002, and then for those aged 17–18 in 2005.
3. "Ever had sex" =1 if young person reported ever having had sex in a given year, =0 otherwise.
4. "Condom use at last sex" =1 if young person reported using a condom at last sex in the year before the interview, =0 otherwise.
5. "Multiple partners in last year"=1 if young person reported more than one sex partner in the 12 months before the survey, =0 otherwise (*i.e.* not necessarily concurrent partners).
6. Only young people who have made their sexual debut in a given year are included in the definitions of condom use and multiple sex partners.H: 0

Death and job loss are the two highest incidence shocks and we group them together in the variable "any shock". As can be seen in Table 2, over 36% of young adults experienced at least one of these shocks in the period 2000-2002 and over 26% experienced one of these shocks in the period 2002-2005.

(b) Sexual Behaviour and Changes Over Time

Table 3 reports average sexual behaviours in 2002 and 2005 by race, gender and age. Also, it reports tests of whether the 2002-2005 changes are statistically significant. There are considerable and interesting variations by race and age both in the levels and in the changes.

As is to be expected, the probability of reporting sexual debut increases by age for both genders and race groups. Within this common trend, coloured youth and coloured females in particular are much less likely to report ever having sex than African youth at each age. Coloured females are much less likely to report ever having sex by age 20 than any other group. The difference between African females and males is less marked with African females being more likely to report having ever had sex than their male counterparts in each wave.

In like manner, condom use differs markedly by race and gender. In both years, the proportion is much higher for males than females and marginally higher for African than coloured males. African females are very much more likely to use a condom than their coloured comparators. In 2002, 24.9% of coloured females used a condom at last sex while 72% of African males reported condom use. With the exception of African males, the probability of condom use at last sex decreases with age.

The patterns for multiple sex partners in the last year are the same as those which we have just described for condom use. However, the magnitudes are lower, ranging, in 2002, from 7.7% for coloured females to 56.7% for African males.

Taking means over the entire sample may confound aging effects with real changes in behaviour for a given age group. For example, as young people age, we would expect more of them to reporting ever having had sex. To check that all of these changes are not being driven by the aging of the sample, we restrict the sample to those aged 17-22 in both 2002 and 2005. Thus, we are actually comparing 17- to 22-year-old youth in 2002 with the same age group in 2005. Table 3 shows that more 17- to 22-year-old youth are likely to report having had sex before in 2005 than in 2002. The increase in probability of reporting sexual debut is large at each age, particularly for African males aged 18-22 and for all African females.

Turning to condom use in the second panel of Table 3, we note that condom use increases for 17- to 22-year-olds between 2002 and 2005. This increase is much stronger for African than coloured youth. The increase in condom use by African youth is a real change that is good news. Dinkelman *et al.* (2007) show that condom use in 2005 is higher for each age among African females. The very low levels of condom use by coloured women may not necessarily be a cause for concern if these women are also in monogamous partnerships. However, these low levels of condom use may be worrying, given that about 25% of coloured males still report multiple partners in the year before 2005.

In the third panel of Table 3, we see a lower proportion of males and females reporting two or more partners by 2005. There is a definite reduction in the incidence of multiple partnerships within this age group. This is not attrition bias (*i.e.* we are not losing the most sexually active) because the individuals in each picture appear in both waves. This is consistent with a search story in which, as young people age, the incidence of multiple partnerships falls as searching yields more stable partnerships.

In sum then, our basic descriptive statistics indicate some positive changes in behaviour: higher condom use, fewer youths reporting more than one recent partner. There are also some ambiguous changes: a higher proportion of youths having sex at every age. By construction of the age cohorts we have shown that these effects are due to more than aging.

(c) Household Income and Household Income Shocks

The argument for using household shocks to look at the effect of an income reduction on behaviour rests on the strong assumption that these shocks are not systematically related to unobserved household type. One empirical fact we would like to be true is that these shocks are not systematically related to observables either: that is, poor households are not more likely to experience some shocks than richer types. Figure 1 provides some evidence that this is a defensible assumption, especially for job shocks. The graphs show the proportion of households (one observation per household) in our matched sample that reports a death, illness, or job loss shock in 2000-2002 or 2002-2005.[15] These proportions are graphed by quintile of 2002 household *per capita* income.

[15] Other shocks were omitted because they were of much lower incidence in the sample. They do not form part of our shock measure.

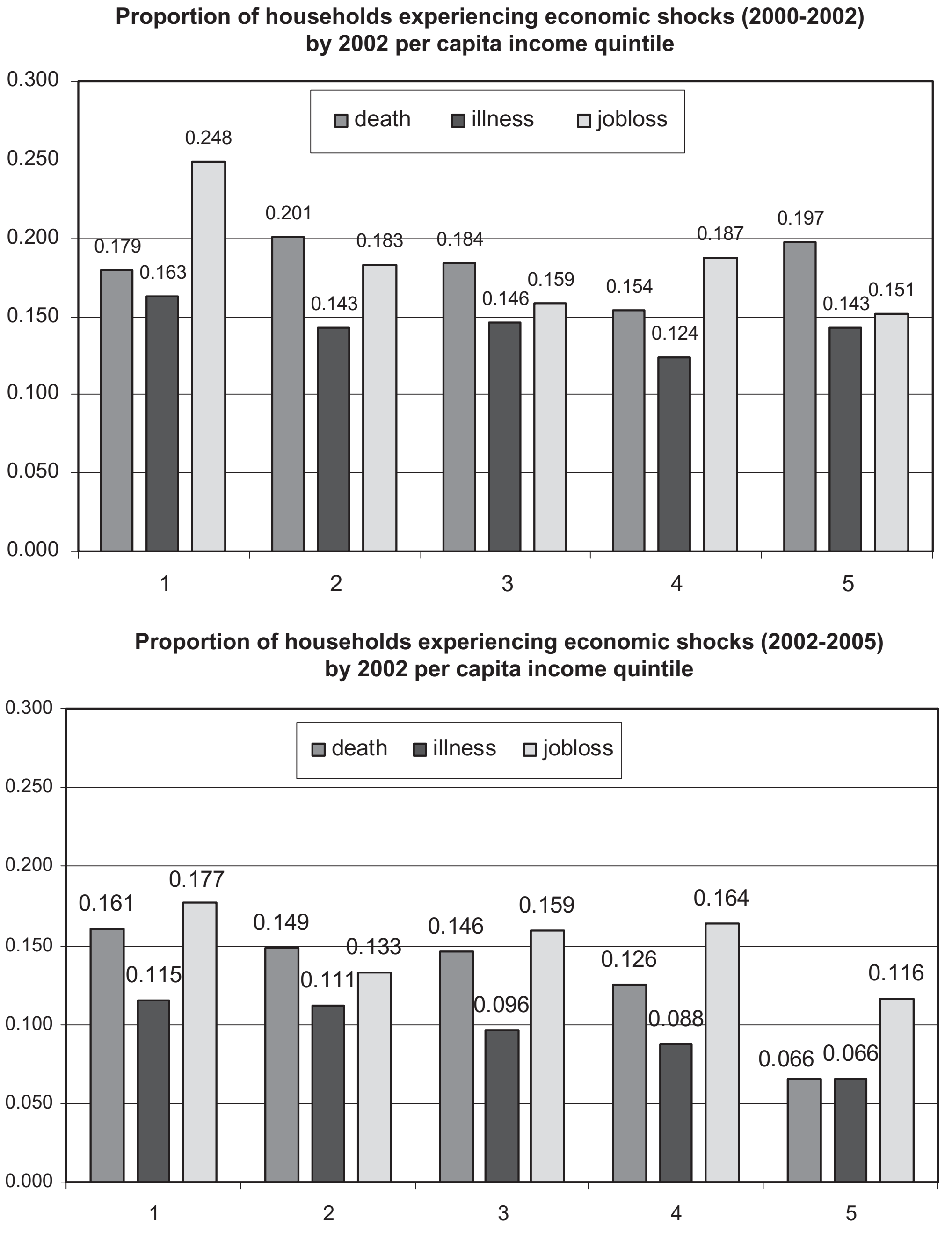

Figure 1. Incidence of income shocks (2000-2002; 2002-2005) by 2002 income quintile

What is striking is that these shocks affect a substantial proportion of even the richest households. Death and illnesses in 2002-2005 are most prevalent among the lower quintile households, but this is less systematic for the 2000-2002 shocks. Job shocks do not seem to have a systematic relationship to household income quintile in either time period. Over 15% of households in each quintile experience a job loss before 2002, and 10% or more have a job shock between 2002 and 2005. Over time, the incidence of all of these shocks seems to have decreased for all quintiles. The graphs give us more confidence in the assumption that these shocks are not related to unobserved household characteristics. They also show that it will be possible to compare households with and without shocks at a given level of income.

(d) Probit Results

In Table 4 probit results (marginal effects calculated at the means of each of the variables) are presented, for females and males separately, for reports of sexual debut by 2002 and for sexual debut between 2002 and 2005. We discuss in detail here the set up of the right-hand side variables as well as the eight estimations in the table because this same approach is replicated for condom use and for multiple partners in Tables 5 and 6.

The first column in each panel ((1) and (5), respectively) shows results with dummy controls for African (1)/coloured (0), mother at home (1), father at home (1), eight age categories (not shown), and a continuous control for the individual math and English test

Table 4. Predictors of sexual activity in 2002 and sexual debut between 2002 and 2005 (Marginal effects from probit estimations)

X-variables	**Females**				**Males**			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	2002	**2002**	**2002**	**2002/05**	**2002**	**2002**	**2002**	**2002/05**
African	0.138	0.135	0.138	0.308	0.125	0.122	0.126	0.076
	(0.053)***	(0.053)***	(0.053)**	(0.058)***	(0.056)**	(0.056)**	(0.056)**	(0.074)
Education	–0.058	–0.058	–0.059	0.014	–0.002	–0.002	–0.002	0.002
	(0.011)***	(0.011)***	(0.011)***	(0.020)	(0.012)	(0.012)	(0.012)	(0.018)
Mother at home	–0.087	–0.087	–0.086	–0.077	–0.001	–0.002	–0.004	–0.034
	(0.038)**	(0.038)**	(0.038)**	(0.054)	(0.041)	(0.041)	(0.041)	(0.061)
Father at home	–0.055	–0.055	–0.055	–0.050	–0.047	–0.047	–0.046	–0.013
	(0.034)	(0.034)	(0.034)	(0.044)	(0.035)	(0.036)	(0.036)	(0.049)
Test z-score	–0.014	–0.014	–0.013	–0.045	–0.064	–0.065	–0.064	–0.048
	(0.021)	(0.021)	(0.021)	(0.031)	(0.022)***	(0.022)***	(0.021)***	(0.030)
Log pc hh income	–0.046	–0.046	–0.049	–0.05	–0.003	–0.002	–0.004	–0.011
	(0.019)**	(0.019)**	(0.019)***	(0.026)*	(0.021)	(0.021)	(0.021)	(0.029)
Community poverty rate	0.379	0.379	0.374	0.082	0.297	0.297	0.293	0.462
	(0.162)**	(0.162)**	(0.162)**	(0.210)	(0.176)*	(0.176)*	(0.176)*	(0.232)**
Shock, 2000-2002		0.035	0.035	0.034		0.04	0.039	0.005
		(0.033)	(0.033)	(0.043)		(0.036)	(0.036)	(0.047)
Shock, 2002-2005			0.002	0.053			–0.027	0.098
			(0.036)	(0.045)			(0.038)	(0.050)**
N	1,589	1,589	1,589	858	1,338	1,338	1,338	691
y-bar	0.39	0.39	0.39	0.48	0.44	0.44	0.44	0.51

Notes:

1. Robust standard errors in brackets, clustered at the household level. ***p < 0.001, **p < 0.05, *p < 0.1.
2. In columns (1)-(3) and (5)-(7) the outcome variable is whether the individual reported ever having sex by 2002.
3. In columns (4) and (8) the outcome variable is an indicator for whether the respondent first started having sex between 2002 and 2005; =0 if respondent had never had sex by 2005. The sample used for these regressions includes everyone who had never had sex by 2002.
4. All X-variables are measured in 2002, except for the household shock that occurs between 2002 and 2005.
5. Other included variables: eight age dummies and a dummy variable for households with missing income data.

Table 5. Predictors of multiple sex partners in 2002 and an increased number of partners between 2002 and 2005 (Marginal effects from probit estimations)

X-variables	Females				Males			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2002	2002	2002	2002/05	2002	2002	2002	2002/05
African	0.067	0.072	0.064	0.072	0.028	0.027	0.016	0.138
	(0.055)	(0.056)	(0.056)	(0.029)**	(0.086)	(0.088)	(0.089)	(0.130)
Education	0.003	0.002	0.002	0.001	0.000	−0.002	−0.001	0.013
	(0.010)	(0.010)	(0.010)	(0.003)	(0.016)	(0.016)	(0.016)	(0.020)
Mother at home	0.021	0.022	0.022	−0.001	−0.003	0.004	0.008	−0.056
	(0.032)	(0.032)	(0.033)	(0.010)	(0.059)	(0.059)	(0.059)	(0.068)
Father at home	0.007	0.007	0.006	−0.009	−0.031	−0.03	−0.024	0.111
	(0.037)	(0.037)	(0.038)	(0.009)	(0.053)	(0.052)	(0.052)	(0.064)*
Test z-score	−0.039	−0.038	−0.035	−0.007	−0.023	−0.023	−0.021	0.008
	(0.018)**	(0.018)**	(0.018)*	(0.006)	(0.032)	(0.031)	(0.032)	(0.035)
Log pc hh income	0.007	0.007	0.011	0.002	−0.02	−0.025	−0.032	0.024
	(0.020)	(0.019)	(0.019)	(0.005)	(0.030)	(0.030)	(0.031)	(0.038)
Community poverty rate	0.123	0.12	0.135	−0.049	0.108	0.093	0.064	0.151
	(0.142)	(0.144)	(0.145)	(0.031)	(0.245)	(0.249)	(0.251)	(0.296)
Shock, 2000-2002		−0.054	−0.052	0.017		0.093	0.087	0.065
		(0.030)*	(0.030)*	(0.012)		(0.051)*	(0.051)*	(0.071)
Shock, 2002-2005			0.018	0.002			0.09	0.142
			(0.033)	(0.008)			(0.053)*	(0.089)
N	591	591	591	403	517	517	517	183
y-bar	0.160	0.160	0.160	0.040	0.510	0.510	0.510	0.210

Notes:

1. Robust standard errors in brackets, clustered at the household level. **$p < 0.05$, *$p < 0.1$.
2. Outcome variable in (1)-(3) and (5)-(7) is whether the individual reported more than one partner in 2002, =0 if no partner or 1 partner. Only individuals who report ever having had sex before in 2002 are included in the analysis sample.
3. Outcome variable in columns (4) and (8) is an indicator for whether the respondent increased the total number of sex partners in the last year between 2002 and 2005; =0 if respondent kept the same number of partners or reduced partners. The sample used for these regressions includes all those reporting they were sexually active in 2002.
4. All X-variables are measured in 2002, except for the household shock that occurs between 2002 and 2005.
5. Other included variables: eight age dummies and a dummy variable for households with missing income data.

scores (z-scores). Log *per capita* household income in 2002 and the community poverty rate are included as the income variables of interest. A dummy variable which is not shown controls for missing household income. The second column in each panel ((2) and (6), respectively) adds shocks in the 2000-2002 period to this model and the third column in each panel ((3) and (7), respectively) adds shocks after 2002 and before 2005. Columns (4) and (8) report estimations of whether or not females and males who had not made their sexual debut by 2002 had done so by 2005.

Females in better off households are significantly less likely to report ever having had sex by 2002 compared to poorer households: for a 10% increase in log *per capita* household income, the probability of reporting ever having sex by 2002 decreases by 0.5 percentage points. This effect is about the same as the effect of an additional year of schooling. Adding controls in the levels equations does not reduce the coefficient on income. Looking at the transition into being sexually active between 2002 and 2005 (column (4)), there is some evidence that higher income households may protect females from making this sexual transition. However, this presumes that the missing income dummy captures richer households. Recall that the income coefficient (on log *per capita*

Table 6. Predictors of condom use at last sex in 2002 and decrease in condom use between 2002 and 2005 (Marginal effects from probit estimations)

X-variables	Females				Males			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2002	2002	2002	2002/05	2002	2002	2002	2002/05
African	0.311	0.305	0.311	−0.362	0.269	0.27	0.269	−0.274
	(0.068)***	(0.068)***	(0.068)***	(0.125)***	(0.067)***	(0.066)***	(0.067)***	(0.078)***
Education	0.042	0.042	0.043	−0.063	−0.005	−0.006	−0.005	−0.01
	(0.013)***	(0.013)***	(0.013)***	(0.026)**	(0.013)	(0.013)	(0.013)	(0.016)
Mother at home	0.096	0.096	0.095	−0.021	0.084	0.086	0.083	0.016
	(0.043)**	(0.043)**	(0.043)**	(0.080)	(0.054)	(0.054)	(0.054)	(0.063)
Father at home	−0.007	−0.006	−0.005	0.030	0.020	0.020	0.018	0.076
	(0.048)	(0.049)	(0.049)	(0.084)	(0.044)	(0.044)	(0.044)	(0.053)
Test z-score	0.023	0.023	0.023	0.063	0.040	0.040	0.039	−0.013
	(0.028)	(0.028)	(0.028)	(0.049)	(0.027)	(0.027)	(0.027)	(0.033)
Log pc hh income	0.024	0.024	0.019	0.037	−0.004	−0.006	−0.003	−0.041
	(0.024)	(0.024)	(0.024)	(0.039)	(0.026)	(0.026)	(0.025)	(0.030)
Community poverty rate	0.056	0.056	0.043	0.360	−0.802	−0.798	−0.785	0.288
	(0.213)	(0.213)	(0.213)	(0.365)	(0.219)***	(0.218)***	(0.219)***	(0.254)
Shock, 2000-2002		0.049	0.05	0.138		−0.049	−0.051	0.008
		(0.044)	(0.044)	(0.076)*		(0.044)	(0.044)	(0.054)
Shock, 2002-2005			0.002	0.064			0.031	−0.016
			(0.044)	(0.072)			(0.048)	(0.060)
N	654	654	654	256	576	576	576	347
y-bar	0.37	0.37	0.37	0.400	0.70	0.70	0.70	0.270

Notes:

1. Robust standard errors in brackets. ***$p < 0.001$, **$p < 0.05$, *$p < 0.1$.
2. Outcome variable in (1)-(3) and (5)-(7) is whether the individual reported more than one partner in 2002, =0 if no partner or 1 partner. Only individuals who report ever having had sex before in 2002 are included in the analysis sample. Totals differ from Table 5 since some respondents did not report information on number of partners in the past year.
3. Outcome variable in columns (4) and (8) is an indicator for whether the respondent reported using a condom at last sex in 2002 and not in 2005; =0 if condom use behavior was the same across waves. The sample used for this regression includes all respondents who were sexually active in 2002.
4. All X-variables are measured in 2002, except for the household shock that occurs between 2002 and 2005.
5. Other included variables: eight age dummies and a dummy variable for households with missing income data.

household income and on the missing income dummy) should be interpreted as the net effects of two forces: household income and household type.

The community poverty rate coefficient is revealing. For a 10% increase in community poverty, females are almost 4% more likely to report ever having sex by 2002. The community poverty rate coefficient is remarkably consistent across models and the size of this coefficient swamps all of the other coefficients in every female regression (1)-(3). However, it is not statistically significant in equation (4) implying that females from poorer communities who were not sexually active in 2002 were no more likely to report sexual debut between 2002 and 2005 than analogous females from better-off areas.

In (2) the economic shock by 2002 variable has a positive coefficient that does not disappear when we add in the future shock variable in (3). However, neither the 2002 nor the 2005 shock variables are significant in these equations. Recall that including the future shock in the 2002 probit model is an indirect test of the hypothesis that shocks are uncorrelated with unobserved household effects. The data pass this test, and we can be more confident that we are really picking up the effect (or lack thereof in this case) of shocks on females' sexual debut. Looking at transitions between 2002 and 2005, economic shocks in this period predict an increase in the chances of making a sexual debut by 2005, conditional on household incomes and community poverty rates. However, again, the coefficient is not statistically significant. We are unable to implement the indirect test that the shocks are picking up a household effect until future waves of data become available.

One of the most striking results in Table 4 is the difference between household predictors of sexual debut for females and males who live in the same households. Columns (5) through (7) show that males in rich households are not significantly less likely to report being sexually active than males in poor households. On the other hand, community poverty rates are particularly good predictors of sexual activity: for a 10% increase in community poverty, the probability of sexual debut by 2002 increases by 4.6%. This effect is somewhat smaller for the probability of debut between 2002 and 2005. The economic shocks do not seem to matter (statistically) for male sexual debut.

Since females and males are not differentially likely to live in rich or poor households or in households with high or low shock propensity, it is puzzling why we should find different results on household income for females and males. One conclusion we could draw is that household unobservables matter differentially for the sexual debut decisions of females than for males.

Table 5 presents results for females and males, separately, that were sexually active by 2002 and reported more than one partner in 2002. The estimates for females and males are shown in columns (1)-(3) and (5)-(7), respectively. In columns (4) and (8) we report estimates of whether females and males who were sexually active by 2002 with 0 or one partner in 2002 increased their partners to more than one in 2005.

Our earlier descriptive analysis reported that African females were significantly more likely than coloured females to report multiple partners in the past year; while African and coloured men did not have notably different propensities for reporting multiple partners. The fact that the African coefficient is insignificant in all level estimations implies that this case is much weaker when appropriate controls are in place. That said, the African coefficient is positive, large and highly significant in the change estimation for females reported in column (4). This implies a significantly larger increase in having multiple sexual partners for African females compared to coloured females over the 2002/2005 period holding other factors constant.

Sexually active females in higher income households are *more* likely to report multiple partners in 2002, although these coefficients are not statistically significant from zero. Males in richer households are *less* likely to report multiple partners in 2002, but, as with females, their coefficients are not statistically different from zero. When we restrict the estimation to sexually active youth with 0 or one partner in 2002 and predict which of them will increase their partners by 2005, we see that household income predicts a positive but insignificant change for males, while the household income coefficient is effectively zero for females. Thus, this too is certainly not definitive and, of course, this coefficient should still be interpreted as a biased income effect that conflates the effect of income and household type on behaviour. Thus, males living in high ability households with high incomes may be able to attract more partners over time, but it is not clear that this is because of economic resources or because their resources reflect other desirable aspects of a partner.

The community poverty coefficients on multiple partners are positive but insignificant for both females and males. Comparing coefficients on the community poverty rates in the transition equations for males and females yields an interesting puzzle. Females in poorer communities in 2002 are significantly less likely to report multiple partners in 2005; while males in poorer communities in 2002 are more likely to report multiple partners in 2005. Although neither coefficient is statistically different from zero, we tested the difference between male and female community poverty coefficients and rejected that they were equal to each other at the 5% level ($p = 0.0232$).[16]

If females in poor communities are matching with males in poor communities, it is not clear how these coefficients can be different from each other. It may be the case that the sample of females for whom we are not measuring effects (those who begin with multiple partners in 2002 and particularly those in poorer areas) are the ones that are matching with these males. We cannot measure the extent of positive assortative matching on community poverty rates with our data. In addition, the sample of females who report multiple partners in 2002 is very small, so drawing inference from probit models with this sample becomes difficult.

The 2002 shock variable is significant in the probit models of both females and males (equations (2) and (6)) and it retains its significance with the inclusion of the shock between 2002-2005 variable (equations (3) and (7)). However, the female and male coefficients have opposite signs. The 2002 shock reduces the probability of multiple partners for females and increases the probability for males. In the case of males, the shock 2002-2005 variable is significant too. We are unable, therefore, to reject the hypothesis that we are picking up unobserved household effects in this case. Further lack of clarity arises from the fact that the shock variable has a positive but insignificant sign for both females and males in the increases in sexual partner equations ((4) and (8), respectively). Thus, the picture with regard to shocks and multiple partners is far from definitive.

Table 6 shows a set of estimates for the last outcome; condom use at last sex. The results suggest that this outcome is not significantly different across rich and poor households for males and females. We tested whether this was the case, and were unable to reject that these coefficients are the same in the male and female equations. Community poverty rates are significantly related to the probability of reporting no condom use at last sex for males but not for females. This difference was significant across males and females at the 1% level (p-value of the test is $p = 0.0194$). Across columns (5)-(7) a 10% increase in community poverty is associated with about an 8% increase in probability of not using a condom for males in 2002; none of the economic shocks are significantly related to low condom use or reducing condom use over time for either females or males.

For the most part, none of the household level variables or economic variables significantly predict lack of condom use in 2002 or a decrease in condom use over time. Education seems to positively predict condom use for females, it does not for males. That the lack of condom use among youths is not systematically different by household income may imply that all youths are behaving in similar ways regardless of income. However, it may also mask different reasons for unsafe behaviour across rich and poor kids.

6. CONCLUSION

Young people in South Africa transition into a sexual marketplace characterised by a generalised HIV/AIDS epidemic and so understanding what influences their behaviours

[16] We also rejected that all of male and female coefficients in each equation were equal to each other: the p-value of this test was $p = 0.0314$.

is a high priority area for research. Getting subsequent groups of youths to protect themselves would do much to stem the tide of the epidemic. However, getting a close look at the sexual behaviours of young people and how they relate to their household resources is difficult with most data sources.

Using a unique panel data set of almost 3,000 young adults in Cape Town, we show that over time there are definite improvements in some types of behaviour for African and coloured youths. As these 14- to 22-year-olds, they are much more likely to report condom use and less likely to report more than one partner in the past year. These are not just aging effects; real changes are evident across cohorts too.

In looking at how these behaviours are different across rich and poor households, we find that low levels of household income and an economic shock (a death, job loss or loss of a grant or other financial support) increases the risk of a sexual debut, conditional on age. This is consistent with qualitative work on the effect of resource constraints on risky sexual behaviour. The relationship between household resources, shocks and sexual debut is weaker for males, which is something of a puzzle. Household income is positively and significantly associated with males having an increasing number of partners over time. While we cannot separate out a true income effect from a household type effect, this result is consistent with anecdotal and anthropological evidence that wealthier males have more partners. Community poverty rates significantly predict earlier sexual debut for females and males, and higher rates of unprotected recent sex for males. Just what these community poverty rates are proxying for is difficult to unpack. What the significant behavioural differences across richer and poorer communities do indicate is that in addition to other channels, youths in poorer areas are vulnerable to HIV/AIDS through their own risky behaviours.

Importantly, our data enable us to use economic shocks to approach the ideal experiment of a random assignment of income across households. This research design is possible with panel or retrospective data; particularly when measuring transitions in behaviour. Identification relies on the assumption that shocks are not correlated with unobserved household type. We showed that in most cases the 2002 data do not fail the false experiment that future shocks predict past behaviour, which provides some justification for this research design. Having provided this tighter test for assessing the impact of shocks, the only behaviour that was sensitive to shocks was the prediction of multiple partners. For males a shock increased the probability of multiple partners (but failed the future shock test) while for females it decreased this probability. However, for females a negative shock reduces the probability of condom use all other things held constant.

The observation that many behaviours differ significantly by household and community income levels (and across males and females within these households and communities) is also important. As our discussion of identification points out, we cannot take these correlations to be causal. However, the findings do challenge us to think about the sources of bias in our econometric framework as well as consider alternative identification strategies that may effectively deal with these types of biases.

APPENDIX

Table A1. Do CAPS respondents give consistent reports on number of sex partners?

X-variables	FEMALES			MALES		
	Enumerated number of partners	Enumerated number of partners	More than one partner	Enumerated number of partners	Enumerated number of partners	More than one partner
	(1)	(2)	(3)	(4)	(5)	(6)
Reported total number of partners	0.688 (0.015)***	0.687 (0.015)***	0.862 (0.014)***	0.652 (0.018)***	0.653 (0.018)***	0.903 (0.017)***
Privacy		−0.021 (0.050)			−0.074 (0.118)	
Constant	0.394 (0.033)***	0.398 (0.034)***	0.003 (0.009)	0.583 (0.056)***	0.587 (0.056)***	0.01 (0.014)
N	1,096	1,095	1,096	812	811	812
R^2	0.64	0.64	0.77	0.61	0.61	0.77

Notes:
1. Robust standard errors in brackets. ***p < 0.001.
2. Standard errors clustered at the 2002 household level.
3. Each column is from a different regression: columns (1), (2), (4) and (5) have as their outcome variable the total number of partners enumerated by the respondent during the survey. The outcome variable in columns (3) and (6) is a dummy variable for whether the respondent enumerated more than one partner in the prior 12 months.
4. The "reported total" is the number of sex partners reported in response to the question "How many partners have you had in total in the last 12 months?"
5. "Privacy" is an indicator variable equal to one if the individual chose to self-report the sexual partnerships questionnaire, equal to zero if they allowed the interviewer to record answers.
6. Observations are restricted to those who report a total of 10 or fewer partners in the last 12 months.

Table A2. Predictors of non-attrition between 2002 and 2005: probit results

X-variable	Probit coefficient
Female	−0.090 (0.049)*
African	−0.488 (0.063)***
Years of education	0.043 (0.014)***
Log pc hh income	−0.104 (0.029)***
Prevalence of shack dwellers	−0.041 (0.091)
Work	−0.027 (0.072)
Missing work 2002	−0.187 (0.302)
Missing education 2002	−0.066 (0.419)
Observations	4,124
Log-likelihood	−2,285.65

Notes:
1. Robust standard errors in parentheses *significant at 10%; ***significant at 1%.
2. All x variables are measured in 2002. Dependent variable is =1 if young adult was found (*i.e.* did not attrit) and re-interviewed in 2005, otherwise 0.
3. Other included variables: age dummies for ages 15 to 17, and a constant.

REFERENCES

ANDERSON, K. G. and BEUTEL, A. G. (2006). Race, gender, and perceived HIV/AIDS risk among South African youth. University of Oklahoma: Unpublished manuscript.

ARDINGTON, C., LAM, D., LEIBBRANDT, M. and WELCH, M. (2006). The sensitivity to key data imputations of recent estimates of income poverty and inequality in South Africa. *Economic Modelling*, 23: 822-835.

CALDWELL, J. C. (2000). Rethinking the African AIDS epidemic. *Population and Development Review*, 26: 117-135.

DINKELMAN, T., LAM, D. and LEIBBRANDT, M. (2007). Household and community income, economic shocks and risky sexual behaviour of young adults: Evidence from the Cape Area Panel Study 2002 and 2005. *AIDS*, 21(suppl 7): S49-S56.

DONNELLY, J. (2006). HIV hits Africa's rich hardest, study says: Analysis disputes long-held beliefs. Boston Globe, June.

DURYEA, S., LAM, D. and LEVISON, D. (2007). Effects of economic shocks on children's employment and schooling in Brazil. *Journal of Development Economics*, 84(1): 188-214.

FENTON, L. (2004). Preventing HIV/AIDS through poverty reduction: The only sustainable solution. *The Lancet*, 364: 1186-1187.

GERTLER, P., SHAH, M. and BERTOZZI, S. (2005). Risky business: The market for unprotected commercial sex. *Journal of Political Economy*, 113 (4): 518-550.

GREGSON, S., NYAMUKAPA, C. A., GARNETT, G. P., MASON, P. R., ZHUWAU, T., CARAL, M., CHANDIWANA, S. K. and ANDERSON, R. M. (2002). Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *The Lancet*, 359: 1896-1903.

JOHNSON, K. and WAY, A. (2003). Risk factors for HIV infection in a national adult population: Evidence from the 2003 Kenya Demographic and Health Survey. *Journal of Acquired Immune Deficiency Syndromes*, 42: 627-637.

LAM, D., SEEKINGS, J. and SPARKS, M. (2006). The Cape Area Panel Study: Overview and Technical Documentation for Waves 1-2-3. University of Cape Town, December 2006.

LECLERC-MADALA, S. (2002). Youth, HIV/AIDS and the importance of sexual culture and context. *Social Dynamics*, 28: 20-41.

LUKE, N. (2003). Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa. *Studies in Family Planning*, 34: 67-86.

——— (2006). Exchange and condom use in informal sexual relationships in urban Kenya. *Economic Development and Cultural Change*, 54: 319-348.

MACPHAIL, C. and CAMPBELL, C. (2001). I think condoms are good but, Aai, I hate those things: Condom use among adolescents and young people in a South African township. *Social Science and Medicine*, 52: 1613-1627.

NNKO, S., BOERMA, J., URASSA, M., MWALUKO, G. and ZABA, B. (2004). Secretive females or swaggering males? An assessment of the quality of sexual partnership reporting in rural Tanzania. *Social Science and Medicine*, 59(2): 299-310.

PETTIFOR, A., VAN DER STRATEN, E., DUNBAR, M. S., SHIBOSKI, S. C. and PADIAN, N. S. (2004). Early age of first sex: A risk factor for HIV infection among women in Zimbabwe. *Epidemiology and Social Aids*, 18: 1435-1442.

RAO, V., GUPTA, I., LOSHKIN, M. and JANA, S. (2003). Sex workers and the cost of safe sex: The compensating differential for condom use among Calcutta prostitutes. *Journal of Development Economics*, 71(2): 585-603.

SHISANA, O., REHLE, T., SIMBAYI, L. C., PARKER, W., BHANA, K. A., CONNOLLY, C., JOOSTE, S. and PILLAY, V. (2005). South African National HIV Prevalence, HIV Incidence, Behaviour and Communication Survey, 2005. *Technical Report*, Human Sciences Research Council, Cape Town, HSRC Press.

SMITH, T. W. (1992). Discrepancies between men and women in reporting number of sexual partners: A summary from four countries. *Social Biology*, 39: 203-211.

STILLWAGGON, E. (2002). HIV/AIDS in Africa: Fertile terrain. *Journal of Development Studies*, 38: 1-22.

TOURANGEAU, R. and SMITH, T.W. (1996). Asking Sensitive Questions: The impact of data collection mode, question format and question context. *The Public Opinion Quarterly*, 60(2): 275-304.

WOOLARD, I. and LEIBBRANDT, M. (2006). Planning for the South African National Income Dynamics Study (NIDS): Lessons from the international experience. Report produced for the Presidency and the Joint Economic AIDS and Poverty Programme (JEAPP), February.