DO KENYAN TEENAGERS RESPOND TO HIV RISK INFORMATION? A PROCEDURAL REPLICATION OF DupaS (2011)

Fernando Martel García, Ph.D.

Submitted as a replication proposal to 3ie Replication Window 3: HIV Prevention
Contents

Preliminary info 3

Principal Investigator ................................................. 3

Replicated Article ......................................................... 3

Contact Details ............................................................... 3

Introduction 3

Replication objectives and research questions 5

What are the dynamic general equilibrium effects – and potential for unintended consequences – of the relative risk intervention? ............................. 5

Were the effects driven by risk avoidance, as the study claims, or moral suasion and stigma? ............................................................. 6

Replication plan 8

Procedural replication ....................................................... 8

Theory of change analysis .................................................. 9

Tentative time frame ......................................................... 12

Conclusion 12

Copyright © 2014 Fernando Martel García
Introduction

Unprotected sex between teenage girls and adult, riskier, men more than five year their senior remains an important channel of HIV transmission in sub-Saharan Africa. Cross-generational sex may also explain why HIV prevalence is three times higher amongst teenage girls compared to boys, and why a quarter of all annual HIV / AIDS infections in sub-Saharan Africa affect those under 25 (Dupas 2011, pp 1–2). At the same time a majority of African children now receive some primary education, which opens up the possibility of using school-based behaviour change interventions to reduce transmission rates cheaply and at scale (Dupas 2011, p 1). Here I propose to replicate Dupas’s (2011) field experiment in Kenya investigating the equivalence of school-based risk reduction interventions, that promote safer sex practices like same-age partner selection amongst teenagers, versus risk avoidance interventions, that promote sexual abstinence until marriage. The study was published in the American Economic Review, and has been cited 157 times (Google Scholar, accessed on 8/10/14).

Dupas’s (2011) study is a blocked, cluster-randomized, controlled, equivalency trial using a $2 \times 2$ factorial design. Study participants are students in 328 primary schools from a convenience sample of seven administrative divisions in the school districts of Bungome and
Butere/Mumias, in Kenya’s Western province. The control arm is the absence of treatment, namely the national HIV / AIDS curriculum introduced in 2001 but seldom taught (Dupas 2011, p. 7). The first intervention arm, rolled out to 164 primary schools between February and May of 2003 and affecting all school grades, is a teacher training (TT) program designed to boost teaching of the national curriculum, with its focus on abstinence (Dupas 2011, p. 7). The second intervention arm, rolled out to 71 schools between July and October in 2004 and affecting only 8th grade students, is a relative risk (RR) information campaign that includes “a 10-minute video used to trigger the discussion on ‘sugar daddies’” (Dupas 2011, p. 29), and information on the local profile of HIV prevalence disaggregated by age and gender. Student participants were cluster-randomized by school, with 127 schools randomly allocated to the control arm; 36 to RR; 125 to TT; and 35 to both TT and RR (5 schools in the TT program lacked an 8th grade class for the RR intervention and so are excluded form the analysis). Secondary analyses were performed on an observational sample of additional “control” cohorts using difference in difference estimators.

The objective of Dupas’s (2011) study is to estimate, in reduced form, the equivalence of risk reduction versus risk avoidance behavioural interventions, as proxied by the RR and TT programs respectively. Key outcomes of interest are defined at the individual level. These include “the incidence of unprotected sex between teenage girls and male partners five or more years older; and the incidence of unprotected sex between teenage girls and teenage boys” (Dupas 2011, p. 12). Neither is measured directly. Instead childbearing data is used as an objective proxy for unprotected sex, complemented by self-reported sexual behaviour data from the subset of study cohort members attending secondary school in 2004. The extent to which participants, teachers, and enumerators were blind to treatment status is unreported.

Dupas (2011) finds that the risk reduction intervention “led to a 28 percent decrease in teen pregnancy, an objective proxy for the incidence of unprotected sex” (Dupas 2011, p. 1), and to a reduction in the age difference between childbearing partners. Based on self-reported sexual behaviour data from a subset of participants attending secondary school in 2004, the study also finds that teenage girls in the RR treatment substitute away from sex with older men towards condom-protected sex with same-age men, thus implying a reduction in teenagers’ risk of HIV infection. And it finds that risk avoidance strategies are ineffective in reducing teenage pregnancies, both alone or in combination with risk reduction strategies.

The study merits replication because it addresses a hugely important topic, touches on an ongoing policy controversy, and has stark policy implications. First, the focus on partner selection suggests an important avenue for preventing cross-generational transmission of HIV / AIDS, a major concern in sub-Saharan Africa. Second, the choice between risk reduction versus risk avoidance interventions is controversial because inducing sexual abstinence until marriage is difficult and likely not cost effective, while risk reduction interventions may induce more sexual activity, thus exposing some groups to additional risks. Third, the study suggests that Kenya’s standard curriculum on risk avoidance is ineffective. A better curriculum would include information on risk reduction, including information about the relative risk of HIV infection by partner’s age. The benefits of an improved curriculum can
be significant, as school-based interventions operate at scale and can be very cost effective.

Replication objectives and research questions

My first objective is to assess the procedural reliability of the original study. That is, establish whether the published findings can be reproduced using the study’s own data and methods, and evaluate the quality of the reporting and the risk of bias from the procedures used. My second objective is to investigate the extent to which policy recommendations are supported by the study findings. Specifically I consider two potential problems in the way the study findings are interpreted: First, a possible ecological fallacy that may be masking harmful side effects, and, second, a problem with construct validity that questions whether the risk reduction information was the real driver of the observed effects.

The concern with the possibility of an ecological fallacy stems from the way aggregate outcome data are used to make inferences about individual behaviour. In turn, this may be masking potential harmful side effects that were a major motivation of the study (Dupas 2011, p. 3). (Aggregate data only allows us to place bounds on individual behaviour which we can explore via simulation.) Meanwhile, the concern with construct validity stems from the fact that the risk reduction intervention involved moral suasion, as well as information on HIV prevalence disaggregated by age and gender, the true focus of the study. Although the study dismisses the possibility that the moral suasion may be driving the results some patterns in the data may suggest otherwise. I expand on these two research questions below.

What are the dynamic general equilibrium effects – and potential for unintended consequences – of the relative risk intervention?

A key claim of the study is that “teenage girls who received information on the relative riskiness of older partners substituted away from older partners and toward condom-protected sex with same-age partners” (Dupas 2011, p. 29). Yet this conclusion does not appear warranted by the theory, experimental design, data, or analysis, and may instead be driven by an ecological fallacy. Indeed, the treatment may have had the unintended consequence of exposing some teenage boys and girls in the treatment group to a higher risk of HIV.

The risk reduction treatment group witnessed an average decline in sex with older men, an average increase in sex with same age partners, and stable pregnancies with same age partners. From these observations the study concludes that teenage girls substituted “away from older partners and toward condom-protected sex with same-age partners”. Yet these average patterns do not tell us whether the exact same girls previously engaged in sex with older men are now the ones engaged in condom-protected sex with same age partners; nor can they rule out the possibility that these girls are now having unprotected sex with teenage boys. If so, their new teenage partners may be facing a higher risk of HIV, as these girls were
at higher risk of HIV from cross-generational sex. In addition, girls who become sexually active because of treatment will also face a higher risk of HIV infection relative to control (where they remain inactive).

Such collateral damage may be transitory or permanent. For example, when the program is first introduced, and assuming it is sustained, there might be a one-off transition cost as girls that had previously engaged in cross-generational sex, and thus were more likely to be HIV positive, turn their attention to same age partners. At the same time, if the program induces some girls to become sexually active, then it has a permanent collateral damage, as the rate of sexual activity (some of it unprotected) may increase relative to control. However, if the intervention was a one-off, then the transition costs of exposing some boys in the treatment group to girls previously engaged in cross generational sex would have been for nothing; while the permanent collateral damage of a sustained program will only be transitory.

The fundamental problem with Dupas’s (2011) study is that it sets out to answer questions about individual teenage sexual behaviour with a research design that cannot answer individual-level questions. Inference about individual behaviour requires a within-person research design, like collecting individual-level panel data on baseline and endline measures. Instead the study measured anonymous outcome variables only at endline, and at times only within a self-selected sub-sample of the experimental subjects. Using these aggregate outcomes to make inferences about individual-level behaviour can thus result in an ecological fallacy.

Given the original study’s research design and data I cannot demonstrate whether there was collateral damage or not. What I propose to do is show the possibility thereof using an agent-based model. Leveraging the experimental results with some additional parameters on teenage sexual behaviour from the literature can illustrate the realm of possibilities, improving our understanding of epidemiological dynamics, potential winners and losers, and the effect of one-off versus sustained behavioural interventions. Such simulations can also highlight potentially important policy and ethical implications. In addition, agent-based models force the researcher to confront the micro-foundations of the experiment: individual decision-making algorithms (Rubinstein 2013, § 3). Learning how teenagers make risky decisions can inform better behaviour change interventions, and ought to be a central focus of any behaviour change research program.

Were the effects driven by risk avoidance, as the study claims, or moral suasion and stigma?

The original study attributes the observed changes in aggregate behaviour to the provision of data on HIV incidence by gender and age, but this interpretation is not warranted by the research design or data. The problem here is participants in the risk reduction arm received a composite treatment. This included “a 10-minute video used to trigger the discussion on
‘sugar daddies’” (Dupas 2011, p. 29), and information on the local profile of risk by age and gender. The “video did not mention HIV, but portrayed “sugar daddies” as predators.” (Dupas 2011, p. 29). The risk data displayed on the blackboard are reproduced in Table 1. In comparison, students in the risk avoidance treatment arm only received data on the average Kenyan prevalence rate (12 percent).

<table>
<thead>
<tr>
<th>Age</th>
<th>15–19</th>
<th>20–24</th>
<th>25–29</th>
<th>30–39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>22</td>
<td>36</td>
<td>35</td>
<td>32</td>
</tr>
<tr>
<td>Male</td>
<td>4</td>
<td>13</td>
<td>28</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 1: Percentage HIV prevalence rates provided to students in the risk reduction treatment arm.

The question is whether the changes in observed behaviour in the risk reduction arm are driven by updating of beliefs about relative risks, as the study concludes, or by the (moral) stigma attached to “sugar daddies” in the video. The study argues that “It is possible, though unlikely, that the video alone would have had a similar effect as that observed.” (Dupas 2011, p. 29). But pinning all of the effect on relative risk raises some possible contradictions. On the one hand teenage girls appear to react strongly to information that the prevalence rate amongst 25-29 years old men is almost treble the national average. On the other hand teenage boys do not appear to react as strongly to information that the HIV prevalence amongst teenage girls is about twice the national average. Or maybe they do. Perhaps girls switch to younger partners, and boys switch towards more protected sex, as the study suggests. But if protected sex offers assurances against risk for boys, why would teenage girls not simply use protected sex with older men, who may also offer other material benefits?

There are two problems here. First, one can question the construct validity of a the risk reduction intervention, as it appears to involve more than providing information on relative risk. (Another way to put this is that the experiment does not meet the exclusion restriction, in that unintended aspects of the interventions may be driving the results.) Second, if we allow for the composite nature of the RR treatment arm, what Shadish, Cook, and Campbell (2002) refer to as a molar treatment, then we need a more explicit model of moral reasoning and decision making to begin to understand the effect (on moral reasoning see Haidt (2001)). Specifically, there appears to be some implicit heterogeneity in the way boys and girls react to the risk information and moral suasion. Perhaps the moral suasion induced risk avoidance (abstinence from sex with older men), while the risk information induced risk reduction (using protected sex with teenage partners).

As with the previous hypothesis the original study’s research design and data will not allow me to answer these questions directly. Instead I will use agent-based simulations and calibrations to get a better understanding of the possibilities. My goal is to question the interpretation of the original study’s findings, and point towards a richer behavioural model with sharper predictions for future experimental testing. In a Bayesian context questioning works as follows. The original author starts with an implicit prior for a particular model of risk avoidance behaviour. The data support that model, so presumably we should gain
more confidence that the risk avoidance model is correct. However, what I argue is that
the heterogeneous behaviour of boys and girls described above is not very consistent with a
simple model of risk avoidance. And a risk avoidance model that would be consistent with
these data is, in my view, quite unlikely \textit{a priori}. So either the data support a model that
is very unlikely \textit{a prior}, or they support another model, as yet undefined, that also explains
these data but is more likely \textit{a priori}. To continue with the Bayesian analogy, the process
of searching for this model is similar to what some Bayesians refer to as model checking
\textit{(Gelman et al. 2013 \S 6)}. However, because my analysis is \textit{post hoc} the proposed model
should be validated using new data.

\section*{Replication plan}

\subsection*{Procedural replication}

To meet the first replication objective – assessing the procedural reliability of the original
study – I will perform a \textit{procedural replication} \textit{(Martel García 2014, pp. 15-17)}.
Intuitively, a procedural replication is like an in-depth peer review; a case-study of how a scientist
went about her research, what errors (if any) were made, and what lessons can be learned for
interpreting, planning, and evaluating similar research. It proceeds in two steps. First, a
\textit{pure replication} to infer the exact procedures and technologies, or scientific standard, used
in producing the study outputs. Second, a critical evaluation of the inferred standard.

The goal of procedural replications is to improve \textit{research practice}, or how scientists go
about doing science. This is accomplished by diagnosing and documenting sources of errors
and omissions, and by generating checklists to improve the reliability of scientific studies
and the effectiveness of peer reviews. The objective is not to prescribe “what to do”, which
likely differs across applications, so much as to inform researchers, reviewers, and editors of
“what to look out for”. Furthermore, the primary goal of procedural replications is not to
make substantive contributions, though such contributions are possible. Changing theories,
concepts, measures, instruments, data sources, model specifications, estimators, or inference
procedures to advance new findings confuses where a replication ends, with where a standard
new scientific study begins (see Figure \ref{fig:replication}). Indeed, replication studies that focus on making
substantive contributions often relegate the details of the replication – what went wrong,
and lessons learned – to a footnote, thus forsaking an opportunity to inform and improve
research practice. At worst, they create incentives for uncontrolled specification searches.

As part of the procedural replication I will perform a \textit{pure replication} \textit{(Martel García
2014, pp. 14-15)} to check whether the replication file obtained from the original author is a
faithful representation of the original research study. A pure replication is successful if the
exact same results reported in the original study, including any errors and omissions, can be
reproduced using the inputs in the replication file. That is, a pure replication helps uncover
the exact procedures and technologies, or scientific standard, used in producing the study
Figure 1: Procedural replications help improve how scientists do science, resulting in more reliable scientific knowledge.

outputs. However, it does not evaluate the inferred scientific standard.

To evaluate the trial in a systematic and reproducible fashion I will rely on the formal definitions and replication procedures outlined in Martel García (2014). In terms of the critical evaluation I will consider both the quality of the reporting (itself a pre-requisite for evaluating the trial) and the risk of bias. To evaluate the reporting I will use the Consolidated Standards of Reporting Trials (CONSORT) extension to cluster randomised trials (Campbell et al. 2012) and the BestBETS Trials Checklists. For example, a quick inspection suggests the original article provides few details about blinding, randomization procedures, sample size calculations, power, participant flow, cluster sizes, intra-cluster correlation, pre-specified hypotheses, or registration details. Of these, the lack of a participant flow diagram makes the results particularly hard to interpret. To evaluate the risk of bias I will rely on the Cochrane Collaboration Risk Assessment Tool (Higgins et al. 2011). Even so, the lack of agreement on formal evaluation criteria (Olivo et al. 2008), and the undesirability of simple rating scales (Higgins et al. 2011), makes individual judgement unavoidable.

Theory of change analysis

To meet the second replication objective – investigating the extent to which policy recommendations are supported by the study findings – I will perform a theory of change analysis (Brown, Cameron, and Wood 2014). Intuitively this is a way to explore model uncertainty by considering the implications of alternative models that a priori are considered likely, are consistent with the findings, yet have different policy implications. For example, the second replication question above examines the construct validity of the relative risk intervention, and explores different models of decisions making. These can yield different interpretations

1Available at [http://bestbets.org/ca/pdf/trial.pdf](http://bestbets.org/ca/pdf/trial.pdf)
of study findings, with different policy implications.

In answering both replication questions I will rely on agent-based models (ABMs). These are computational models of how individual heterogeneous agents interact in a given environment. Typically they are used to investigate the emergent properties of complex systems, including epidemiological dynamics driven by complex interactions between heterogeneous, boundedly rational, individuals. Recently, efforts have been made to combine ABMs with experimental data (Janssen and Ostrom 2006), including research where human subjects participate in role-playing games to help parametrize agent-based models (Barreteau, Bousquet, and Attonaty 2001; Gurung, Bousquet, and Trébuil 2006). At the same time, there is increasing interest in the identification of dynamic general equilibrium effects from experimental data (Abbring and Heckman 2008, 2007; Acemoglu 2009; Coady and Harris 2004; Duflo, Hanna, and Ryan 2012). Indeed, experiments offer a great opportunity to test and calibrate structural models of behaviour.

Mathematical models of sexually transmitted infections, and HIV in particular, have become increasingly popular (Baggaley and Fraser 2010; Cassels and Goodreau 2011). The primary reason is that interventions designed to reduce transmission at the individual level have complex effects across time and space at the population level. Put simply, the population effects we care about, such as incidence, prevalence, and survival, can seldom be estimated or understood using short-run, partial equilibrium, sample estimates from field experiments. To get a grasp of these there needs to be a dialog between modelers and experimenters at all stages of the research process, not least because the statistical power of the intervention in part depends on these complex dynamics. Indeed, such collaboration has become a requirement in the National Institutes of Health’s Methods for Prevention Packages Program to reduce HIV.

The difficulties in extrapolating the partial equilibrium findings from an experiment like Dupas’s (2011) to the broader population are well illustrated by Hallett et al.’s (2007) stratified population model. At a population level the model shows how behavioural interventions designed to influence cross-generational sex may not be very effective in reducing lifetime HIV risk or endemic prevalence – even when the interventions are highly successful in reducing cross-generational sex. At best such interventions reduce risks for teenage girls in the short-term, while increasing risk for teenage boys. By contrast, the model predicts that the use of condoms with all partners, irrespective of age, is more effective in improving population outcomes. Other mathematical models that have studied cross-generational sex include Anderson et al. (1989, 1992) and Garnett and Anderson (1994).

Garnett (2002) and Shalizi (2006) describe various types of mathematical models, and describe their advantages and disadvantages. In practice, there is no “best model”, only different models for different purposes. For example, Hallett et al.’s (2007) model predicts what might happen if we could exogenously manipulate people’s sexual behaviour. Even so, despite ignoring the details of choice behaviour itself, the model is still useful. It tells us that even if the intervention has a big effect on cross-generational sex, it may still have a very limited impact on the population outcomes we care about. In addition, it tells us that
promoting condom use with all partners, irrespective of age, is predicted to have a large effect on population outcomes. However, if we are to change behaviour in the real world, then at some point we need to get down to the details of individual choice behaviour, and that will require a different model.

For example, Dupas (2011) discusses the choice of sexual partners on the basis of income, information, and risk perceptions among other. These are variables that we can manipulate in order to influence behaviour. Indeed, the principal intervention in Dupas’s (2011) experiment was designed to change risk perceptions – and behaviour – through the provision of information. Yet none of these variables are present in Hallett et al.’s (2007) model. That is, Hallett et al.’s (2007) model tells us the effect on the population of changing individual behaviour exogenously, but it does not tell us anything about how we might intervene to bring such changes about. Yet, presumably, a major goal of a research program on sexually transmitted infections is to understand the choice equations driving individual behaviour.

Agent-based models (ABMs) are specially useful when our goal is to study both, individual behaviour, and population outcomes. Indeed, if individuals are highly heterogeneous, and interact in complex ways, then ABMs are the modeling framework of choice, as traditional population level mathematical models, including computable general equilibrium models that rely on representative agents, can seldom handle the complexity (Shalizi 2006). In addition to these computational advantages, ABMs have other practical advantages. For example, they emphasize computer programming over mathematical skills. They are parametrized at the level of individual behavior heuristics, a level with which many qualitative researchers in public health are intimately familiar. And they are ideally suited to the analysis of behavioural axioms relevant to the study of HIV. In this fashion ABMs can provide tractable insights in complex systems where the use of prospect theory, say, may have proved mathematically intractable.

My plan of attack is to use parameter estimates from Dupas’s (2011) experiment, supplemented with additional information from the peer-reviewed literature on teenage sexual behavior in Africa to calibrate an ABM of teenage sexual behaviour in a prototypical experimental school (e.g. Delavande and Kohler (2014); Duflo et al. (2006); Duflo, Dupas, and Kremer (2011); Greenwood et al. (2013); Magruder (2011); Paula, Shapira, and Todd (2013)). This includes modeling the choice behaviour of teenage girls and boys, and older men, their interaction, and the spread of HIV. The goal is not to provide accurate quantitative predictions so much as provide qualitative insights from heterogeneity, complex interactions, and dynamics implicit in Dupas’s (2011) analysis; and use these to reinterpret the original study’s findings and conclusions (for a full blown ABM model calibrated to make quantitative predictions of HIV see McCormick et al. (2014) among other). Such a model can shed light on the dynamic effects of the intervention on teenage boys and girls as the experimental cohort moves through primary and secondary school, and onto graduation; the sensitivity of findings to various choice algorithms and parameters; and what calibrations are consistent with the study findings and recommendations. All the analysis will be done using the free open source ABM software Netlogo and statistical software R (R Core Team 2013).
As with the rest of the replication proposal, the approach in this section is exploratory. That is, the goal is not to refute existing findings and policy advice so much as explore alternative causal models and explanations, and question the original study findings. Adjudicating between alternative causal models is left to future research, as it would require a new scientific study, registered protocol, and analysis plan. The objective is to get a better sense of model uncertainty and the policy implications thereof; and to demonstrate how agent-based modeling can be used alongside field experiments to better investigate dynamic, general equilibrium aspects of causality.

**Tentative time frame**

<table>
<thead>
<tr>
<th>Months</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – 2</td>
<td>Pure replication and evaluation of the original study’s reporting and risk of bias</td>
</tr>
<tr>
<td>3 – 3</td>
<td>Formulate problem and set objective for agent-based simulations</td>
</tr>
<tr>
<td>4 – 5</td>
<td>Build conceptual model, review literature for parameter information</td>
</tr>
<tr>
<td>6 – 6</td>
<td>Conceptual validation of model using system theories and assumptions</td>
</tr>
<tr>
<td>7 – 9</td>
<td>Translate conceptual model into computer model</td>
</tr>
<tr>
<td>10 – 12</td>
<td>Run simulations, operational validation, report results</td>
</tr>
</tbody>
</table>

**Conclusion**

Unprotected sex between teenage girls and adult, riskier, men more than five year their senior remains an important channel of HIV transmission in sub-Saharan Africa. School-based behaviour change interventions offer a potentially cheap, and scalable avenue for inducing safer sex practices, and reducing cross-generational HIV transmission. Dupas’s (2011) study is a blocked, cluster-randomized, controlled trial that uses a $2 \times 2$ factorial design to investigate the equivalence of school-based *risk reduction* interventions, that promote safer sex practices like same-age partner selection amongst teenagers, versus *risk avoidance* interventions, that promote sexual abstinence until marriage. The study finds that the risk reduction intervention “led to a 28 percent decrease in teen pregnancy, an objective proxy for the incidence of unprotected sex” (Dupas 2011, p. 1), while risk avoidance strategies are ineffective in reducing teenage pregnancies, both alone or in combination with risk reduction strategies. The study was chosen for replication because it addresses a hugely important topic, touches on an ongoing policy controversy, and has stark policy implications.

In this replication study I set myself two objectives. First, to assess the procedural re-
liability of the original study. That is, whether the published findings can be reproduced using the study’s own data, and whether the inferred procedures risk introducing bias and are adequately reported. Second, to investigate the extent to which policy recommendations are supported by the study findings. Specifically I consider two potential problems in the way the study findings are interpreted: First, a possible ecological fallacy that may be masking harmful side effects, and, second, a potential problem with construct validity that questions whether the risk reduction information was the real driver of the observed effects. Throughout I plan on using agent-based model to explore these dynamic, general equilibrium aspects of causality.

Besides a more explicit accounting of the uncertainty surrounding the original study, and its implication for HIV policy, this procedural replication can yield important lessons for research practice. First, the ethics of anonymity ought to be reconsidered when it compromises the goals of the study, specially when conducting the study may expose some subjects to additional risks. Second, IRB boards and funding agencies should insist that studies about individual-level behaviour use research designs the can answer individual-level questions. Third, researchers ought to pay more attention to dynamic general equilibrium effects, and consider experiments as an opportunity to test and calibrate structural models of behaviour.

References


Barreteau, O., F. Bousquet, and J.M. Attonaty. 2001. “Role-playing games for opening the
black box of multi-agent systems: method and lessons of its application to Senegal River

Brown, Annette N., Drew B. Cameron, and Benjamin D. K. Wood. 2014. “Quality evidence
for policymaking: I’ll believe it when I see the replication.” *Journal of Development
Effectiveness*: 1-21. Published online August 8 2014.

Campbell, Marion K, Gilda Piaggio, Diana R Elbourne, and Douglas G Altman. 2012.

Cassels, Susan, and Steven M Goodreau. 2011. “Interaction of mathematical modeling and
social and behavioral HIV/AIDS research.” *Current opinion in HIV and AIDS* 6 (March):
119–123.

Coady, David, and Rebecca Lee Harris. 2004. Evaluating targeted cash transfer programs:
a general equilibrium framework with an application to Mexico. Research reports 137
International Food Policy Research Institute (IFPRI).

Delavande, Adeline, and Hans-Peter Kohler. 2014. HIV/AIDS-related Expectations and
Risky Sexual Behavior in Malawi. Technical report.


HIV/AIDS Prevention : Evidence From A Randomized Evaluation In Western Kenya.”


1–34.

disease epidemiology.” *Sexually transmitted infections* 78 (1): 7–12.

and activity stratified model of HIV transmission in heterosexual populations.” *Mathemat-

Taylor & Francis.


