Three Risk-elicitation Methods in the Field:
Evidence from Rural Senegal

August 18, 2015

ABSTRACT: In the past decade, it has become common to use simple laboratory games and decision tasks as a device for measuring risk preferences in the developing world. In this paper, we build on existing taxonomies for risk elicitation and discuss pros and cons of using such methods in developing-country contexts. We use three distinct risk-elicitation mechanisms (the Holt-Laury task, the Gneezy-Potters mechanism, and a non-incentivized willingness-to-risk scale) and subjects from rural Senegal. Our study provides some guidance to researchers wishing to use risk-elicitation mechanisms in the rural developing world.

Contact: Gary Charness, University of California, Santa Barbara, 2127 North Hall, Santa Barbara, California 93106-9210, 1-805-893-2412.

Angelino Viceisza, Spelman College, 350 Spelman Lane SW, Atlanta, Georgia 30314, 1-404-270-6055

1. INTRODUCTION

In the developing world, welfare evaluation of any proposed policy with risky outcomes should take into account risk preferences (e.g, Harrison, 2011). Suppose a non-governmental organization implements a pilot-seed-variety program. Researchers are often requested to evaluate the impact of such a program, by for example designing a randomized controlled trial (RCT). It is common to wonder how the impact of this RCT varies with farmers' risk preferences. To answer this question, researchers turn to risk-preference data. If these are imprecise or biased, incorrect inferences and ineffective policies can emerge.

It has become common in recent years for researchers in development economics to utilize tools developed by experimental economists, in order to measure important traits, attitudes, and characteristics of the poor in rural settings in the developing world. Some examples include Ashraf (2009), Attanasio et al. (2011); Castillo and Carter (2011), Delavande et al. (2011), Giné et al. (2010), Harrison et al. (2010), and Hill et al. (2012).[1] The risk attitude of the population is one characteristic with clear importance for policy recommendations.[2]

Many techniques, ranging from simple to complex, have evolved for testing risk preferences. Our aim in this paper is to illustrate how three rather different mechanisms play out in rural Senegal. While we do not claim that this environment is fully representative of the developing world, we nevertheless build on existing taxonomies of risk-elicitation instruments and feel that we provide useful insights concerning the potential pros and cons of these mechanisms and their effectiveness, as well as document in detail some procedural issues.

In this study, we implement three elicitation devices in rural Senegal, including two choice-based mechanisms and a non-incentivized survey question. One device is the well-

---

[1] For additional examples, refer to the discussions by Viceisza (2012, 2015).
[2] Binswanger (1980) was the first to elicit risk preferences, using a lottery-choice framework with farmers in India.

known method used in Holt and Laury (2002), where people make a series of 10 choices between two systematically varied alternatives. A second device is an adaptation of a method used by Gneezy and Potters (1997), where there is a simple choice of how much to invest in a risky asset with a positive expected profit from investing. The third mechanism we use is a non-incentivized survey question taken from Dohmen et al. (2011), which asks the respondent to report her willingness to take risk on a scale of ten.

A summary of our main results: (1) There is considerable inconsistency in the choices made by individuals responding to the Holt-Laury (HL) questionnaire. (2) The results with the Gneezy-Potters (GP) approach show only modest differences from previous work. (3) Many women state they would take on the very highest degree of risk on the willingness-to-take risk (WTR) question. This is very different from results in most previous studies, but seems to match the environment in which these women live.

To our knowledge, few (if any) studies compare three different risk-elicitation instruments in the field (and possibly not even in the lab). It is this literature at the intersection of experimental and development economics that we seek to complement with our study. We provide interesting results, assess differences across the mechanisms, and provide detail regarding the issues involved in working in developing countries. Development economists wishing to gather data on risk preferences should find this useful; furthermore, the spirit of our methodological results may also apply to the elicitation of other preferences in this environment.

The remainder of the paper is structured as follows. In Section 2, we propose a taxonomy for risk-elicitation instruments and describe the results from previous studies using our elicitation methods. Section 3 describes the implementation and contains the experimental and survey results. We conclude in Section 4.

2. TAXONOMY OF ELICITATION METHODS AND EXISTING EVIDENCE

There have been many experimental studies on risk preferences, as this is one of the building blocks of economic theory and analysis. For reviews of previous work, we refer the interested reader to Offerman et al. (2009), Dave et al. (2010), and Crosetto and Filippin (2013). Charness, Gneezy, and Imas (2013) provide an in-depth discussion of the general issues involved with eliciting risk preferences. In this section, we provide a taxonomy of different methods, discuss some existing evidence, and use this to motivate the choice of instruments in the next section. When discussing existing evidence, we primarily focus on the main articles that use the elicitation methods that we implement.

Taxonomy of Risk-Elicitation Methods

One approach to classifying risk-elicitation approaches is to categorize them according to whether or not the instrument is (1) incentivized and (2) complex. Incentivized instruments associate choices with salient rewards, in contrast to non-incentivized instruments such as survey-based risk questions. Complex instruments are based on multiple price lists (MPL's, a series of choices between lotteries with varying probabilities or payoffs) or multiple risk questions; simple instruments have only one choice or response, with fixed probability where applicable. While other criteria could be used to organize risk-elicitation instruments, we choose incentives and complexity since they are relatively general and still provide structure.

Based on these criteria, Table 1 shows four broad categories of instruments with related sample studies: (1) incentivized-complex, (2) incentivized-simple, (3) non-incentivized-complex, and (4) non-incentivized-simple.

[Table 1 about here]

Each of these categories has advantages and disadvantages. While incentivized studies elicit incentive-compatible responses, researchers must carry substantial rewards into the field. This can be risky and time-consuming. Also, incentivized instruments are typically based on more abstract lottery choices rather than survey-based WTR questions (framed in day-to-day or general contexts). So incentivized questions can potentially cause more confusion than simple survey-based questions. Similarly, while complex instruments have the advantage of eliciting multiple responses, they take more time to explain and may present cognitive difficulties. Furthermore, it is difficult to deal with inconsistent subjects with complex instruments. So, it is unclear that any given method is superior to another; one might wish to employ a combination of methods.[3] Table 2 presents a list of studies conducted in developing countries.

[Table 2]

In this study, we implement three elicitation devices with villagers in rural Senegal. The instruments were chosen to represent (1) incentivized-complex, (2) incentivized-simple; and (3) non-incentivized simple. We did not employ a measure that represents non-incentivized-complex (NIC), since this study was designed to collect qualitative data on how subjects perceive "risk" in order to be able to formulate questions along the lines of domain-specific risk tasks (a prime example of NIC). The Holt-Laury task was chosen as a mainstay of risk elicitation. The Gneezy-Potters task, also widely used, was chosen for its incentivized simplicity. Finally, the WTR scale was chosen for its pure simplicity. Furthermore, were it to give fruitful results, this task could be extremely useful, since researchers (typically pressed for time and resources) prefer to use a task that is simplest to explain.

---

[3] Delavande et al. (2010) consider approaches to eliciting expectations in developing areas and conclude that issues such as comprehension of probability concepts and excessive time costs for collection are not insurmountable.

Holt-Laury

Holt and Laury (2002) have participants make 10 choices of either option A or option B, one for each row. Table 3 shows the choices participants faced in the low-payoff treatment. They also conducted treatments in which the payoffs are 20, 50, or 90 times the ones shown, as well as one with hypothetical payoffs. Higher payoffs lead to fewer risky choices, while hypothetical payoffs lead to more risky choices.

[Table 3 about here]

Option B dominates option A in the last row and one must have very strong risk-seeking preferences to choose option B in the first row. If a person chooses option A in the first row, there should be exactly one switch point. If we observe switching back and forth, this would seem to indicate either a lack of comprehension or unmotivated responses.

Most Western studies have inconsistency rates between 10 and 15 percent. But there is gloomier evidence in settings more similar to ours. Lammers, Lau and Verbon (2006) find inconsistency rates of 66.5% in South Africa; Galarza (2009) finds a rate of 52% in Peru, while Doerr et al. (2011) find a rate of 39 percent in Ethiopia. Jacobson and Petrie (2009) and Engle-Warnick et al. (2011) use an instrument that is different from Holt-Laury (but can also be used to assess inconsistencies) and find a rate of 52 percent in Rwanda and Peru respectively.[4] See Table A1 in the online Appendix for more detail.

Gneezy-Potters

Gneezy and Potters (1997) use a simple investment task, which is modified in Charness and Gneezy (2010). Each person is endowed with 100 units. Any or all could be invested in a

---

[4] However, exceptions include de Brauw and Eozonou (2011), who find an inconsistency rate of 14 percent in Mozambique, and Ihli et al. (2013), who implement a modified HL mechanism and find an inconsistency rate of 6%.

risky asset with a 50 percent chance of success, and a payoff of 2.5 times the investment if successful; whatever is not invested is kept. This mechanism typically gives a full range of values and seems rather easy to understand.[5] A summary of the results of other studies using this approach is provided in Table A2 in the online Appendix. There is considerable variation in the investment rate, with developing-world villagers more risk-averse than people in Western societies. The overall investment levels range from 44.67 to 70.86 among student populations and from 23.08 to 50.03 among villagers. Females are more risk averse than males in 11 of these 12 studies around the world, with no difference in the other.

Non-Incentivized Survey (WTR scale)

Dohmen et al. (2011) utilize a question from the German Socio-Economic Panel, which asks respondents to give an assessment of their general willingness to take risks on a 0-10 scale. This approach differs from the others in at least two important ways: (1) there are no financial incentives provided, and (2) this question is not specific to financial risk-taking. As there are many forms of risk-taking (e.g., physical, financial, and emotional), this question may pick up traits orthogonal to the issue of concern; however, they do find that this measure matches up well with the results from five additional questions on risk-taking in different specific contexts and an experimental validation. Still, a large amount of research in psychology suggests serious differences in risk attitudes across domains.[6] Their Figure 1 shows a good spread of risk attitudes, with a strong peak at five (the average value in the range).

To our knowledge, only a few studies employ multiple risk-elicitation methods using the

---

[5] A drawback is that this device cannot distinguish between risk-neutral and risk-seeking behavior; the latter comprises about 10% of the population (Holt and Laury, 2002)
[6] See the discussion in Charness, Gneezy, and Imas (2013).

same sample in a developing country context.[7] Ihli et al. (2013) compare two different lottery

designs, one with fixed payoffs and changing probabilities and the other with fixed probabilities

and changing payoffs, among a sample of farmers in Uganda. While both methods reveal high

proportions of farmers who are classified as risk averse, the two elicitation methods yield

significantly different results. Second, Castillo et al. (2014), using a sample of rural households

in Peru, compare two multiple-price-list lotteries that differ on whether they included only

positive or both positive and negative payoffs and whether payoffs or probabilities were

changed. They find that average behavior is close to risk neutral and the reaction to losses is

uniform and very negative. Third, Nielsen and Zeller (2014) compare one MPL method to six

hypothetical methods for a sample of villagers in Vietnam. Except for one assessment method,

they find that risk preferences are not stable and that the influence of adverse real-life events on

changes in risk preferences varies with the elicitation method.

3. NEW EXPERIMENTAL EVIDENCE

Implementation

Our experiments were conducted in rural Senegal (in the regions of Thies and Diourbel) in

December 2010. Since the experiments were conducted as part of a larger project on linking

farmers to markets, we had access to a sample of farmer groups that are part of the federation of

non-governmental organizations of Senegal (FONGS), which represents such groups at the

national level. The participants were recruited through village-level farmer groups. While these

participants do not represent the entire developing world, they do have several characteristics in

---

[7] Our study is related to that of Dave et al. (2010), who compare (with Canadian subjects) a relatively-complex elicitation method to a coarser-but-simpler elicitation method. They find: "The simpler task may be preferred for subjects who exhibit low numeracy ... For subjects with higher numerical skills, the greater predictive accuracy of the more complex task more than outweighs the larger noise."

common with other populations in similar settings. For example, our subjects have relatively

low (high) education (poverty) levels and depend substantially on agriculture-related activities

for their survival. Furthermore, subjects are often recruited through NGOs or microfinance

organizations.

We conducted four sessions in two days in December, around the agricultural

commercialization season. While farmers tend to be richer during this period, our main

comparisons are across instruments. We are unaware of major events that may have shaped risk

attitudes around this period (see Callen et al. 2014 and Cameron and Shah, forthcoming).

Timeline: On one day, two sessions—one Holt-Laury and one Gneezy-Potters—were

held in a village in Diourbel. Each session was conducted with members of distinct farmer

groups (between subjects). On the next day, two identical (between-subject) sessions were held

in a village in Thies. In all sessions, we elicited the WTR question as part of a pre-survey, with

the option to revise the response after the other task was completed. A typical session lasted

between 2.5 and 3.0 hours.[8] The average payoffs were about twice the average daily wage for a

comparable sample of households. There were 45 participants in the HL sessions and 46

participants in the GP sessions. Since all responded to the WTR question, these data comprised

91 observations.

We provide very detailed information about the implementation process in the Appendix,

which may be a useful reference for researchers. Three main aspects of our experiment protocol

are noteworthy. (1) The experiments were conducted by a main experimenter in English and

---

[8] Given the length of time of these sessions, one might be concerned that data were contaminated by discussions
between people. But this was not the case; in fact, there was little or no free time during a session. A substantial
amount of time was involved in making sure that people understood at the individual level. Additional time was
required for the pre- and post-surveys, seating, introduction, and payment. Finally, subjects were separated by
dividers.

translated on the spot and line by line into Wolof (the main national language).[9] (2) We framed the HL and GP experiments in terms of "seeds" and "yields", since most subjects in rural areas can relate to concepts of risk in agricultural terms.[10] (3) For the WTR question, we described risk as a situation that could sometimes lead to a good event (high payoff/gain) or a bad event (low payoff/loss). Specifically, we asked: "On a scale of 0-10, with 0 being the least willing and 10 being the most willing, how willing are you to take risks." However, we note that "risk" is not a well-defined term in Wolof (so that we described outcomes), and so translations are inevitably imperfect. This points to a potential difficulty in using non-choice mechanisms in the developing world. Note that in all cases, we also asked subjects to report what risk means to them in their everyday environment, in terms of output risk, crop risk, price risk, and credit risk.

In our version of HL, a non-risky seed gave a payoff of FCFA 1,000 (800) in rainy (dry) weather, while a risky seed gave a payoff of FCFA 2,000 (100) in rainy (dry) weather. The probabilities were systematically varied and described as years of good and bad weather. While this framing arguably may have altered subjects' responses, it was adopted to maximize their understanding of probabilities and relate this concept to an issue with which they are familiar in the day-to-day environment.[11] A screenshot is shown in Figure A1 in the online appendix.

In our adapted GP task, non-risky seeds pay FCFA 100 per kilo regardless of the weather. Risky seeds paid FCFA 300 per kilo if the weather was good (rainy) and nothing if the weather was bad (dry). So, one receives 1,000 units if one purchases only non-risky seeds and corresponding amounts for purchases of other numbers of risky seeds in the event of either rainy

or dry weather (each 50 percent likely). We framed the decision in terms of how many "risky seeds" one wished to purchase. A screenshot is shown in Figure A2 in the online Appendix.

Holt-Laury: New Evidence

Figure 1 shows our data and also includes the original Holt-Laury 20x real-payoff data (all high-stake patterns are similar) for comparison. The horizontal axis displays the decision (1 through 10) and the vertical axis displays the probability of option A being chosen (across all respondents). The median number of safe choices is 5.0.

[Figure 1 here]

The Senegal curve shows relative insensitivity to changes in probability (and is not even monotonic). Furthermore, even in the last decision (where option B pays more with *certainty*), 40 percent still choose option A. To analyze choice inconsistency, we classify people into four types:

1. Respondents who first chose option A and switched once to option B. We see these subjects as truly understanding and consider them consistent.
2. Respondents who always chose option B. While these subjects are consistent, in principle we cannot rule out the possibility that they misunderstood.
3. Respondents who always chose option A. While these subjects are consistent, we suspect they did not understand, since one should always prefer option B in decision 10. So, we classify them as "inconsistent" in subsequent analysis.
4. Respondents who switch at least twice.

Under this classification scheme, 24.4 percent of participants appear to have possibly understood the task (type 1 or type 2), while another 24.4 participants always chose A (type 3). Finally, 51.1 percent of the participants switched columns at least twice. There is also a reasonable inference that people of type 2 (11.1 percent) did not really understand the task, since only extreme risk-seekers should choose option B in the first row. Since it is unclear that type 2 individuals really understood the task and only 11.1 percent were type 1, somewhere between

10

11.1 and 48.9 percent of participants understood the task. In line with the Filippin and Crosetto (2014) meta-study of HL results, men make only insignificantly fewer safe choices ($p = 0.394$, one-tailed test). We find that women are less consistent in their choices ($p = 0.030$, two-tailed test).[12]

Overall, our impression is that our formulation of the HL task was not well understood by the participants. While there may be better ways of presenting this task to people in this type of rural environment, it seems that using a relatively sophisticated mechanism is not effective here. Furthermore, this seems to be supported by some of the findings in Table 3.

Gneezy-Potters: New Evidence

Panel A of Figure 2 shows the distribution of risky *seeds* chosen in our Gneezy-Potters adaptation. The average number of risky seeds chosen in Senegal is 4.78 (of 10). On average male subjects chose 5.72 seeds, while female subjects chose 4.18 seeds, consistent with prior Gneezy-Potters results. However, we are reluctant to draw strong conclusions, as there may well be important differences across subject pools. There is a big spike at five risky seeds, as nearly one-third (15 of 46) purchased this number of seeds.

[Figure 2 about here]

While there are multiple differences across the populations, ruling out strong inferences, we nevertheless can make some comparisons to the data from two studies involving non-students (Charness and Villeval 2009 and Haigh and List 2005). There is no significant difference from our data for either study, although the proportion of the highest risk choices in Charness and

---

[12] Overall, in inconsistency in the HL task seems to be driven by people who choose the safe lottery more often and by women, even after controlling for cognitive measures such as education and the ability to calculate 15% of 2000.

Villeval (2009) is greater in Panel B.[13] The difference between male and female investment rates

is significant ($p = 0.034$, one-tailed t-test).

Non-incentivized survey: New Evidence

Figure 3 shows the distribution of subjects' levels of willingness to take risk on a scale of 1 to

10.[14] Dohmen et al. (2011) argue that such a measure generates the best all-around predictor of

risky behavior. There is a strong peak at 10 in our data (28.1 percent of the observations). This is

dramatically different than the 1.0 percent in the Dohmen et al. data. In fact, women were 50

percent more likely to choose 10 than were men (32.7 percent versus 21.6 percent), although this

difference is not statistically significant. Still, this pattern differs from findings in both

psychology and economics that women are financially more risk averse (see Charness and

Gneezy, 2012 and Croson and Gneezy, 2009) in settings around the world.[15] However, once

again, the populations are very different, so one must be cautious about interpretations. Women

are more willing to take risks than men in our WTR data ($p = 0.070$, two-tailed t-test).

[Figure 3 about here]

4. DISCUSSION

We complement existing evidence by testing three different risk-preference elicitation

mechanisms in rural Senegal. Our primary interest is pragmatic: What should a researcher do to

elicit risk preferences in a developing nation? If one is interested in policy implementation in

---

[13] To create Panel B, we placed the Charness and Villeval (2009) results into our bins (e.g., an investment of less than 10 of 100 tokens went to the left-most bin and an investment of over 90 tokens went to the right-most bin.
[14] We report the first WTR responses. There is little difference between the first and second WTR measures, as only nine of 89 people (10 percent changed their report). This suggests that there is little in the way of order effects. Two subjects were undecided between two of the integers, so we drop them from the analysis.
[15] However, Crosetto and Fillipin (2013) find that gender risk-preference differences appear only when the task is more likely to trigger loss aversion; Filippin and Crosetto (2014) find that these differences are smaller with the Holt-Laury task.

developing nations and knowledge of risk preferences is useful, it is vital to implement a mechanism that induces meaningful responses.

As we use three difference elicitation devices, we can examine whether there are significant correlations across WTR and the other risk measures. It turns out, that at least in the aggregate, there are none. Between HL and WTR, the correlation is 0.06 ($p = 0.73$) and between GP and WTR, it is -0.09. However, this looks a bit different when we break this down by gender, although no correlations are significant, perhaps due to the limited number of observations (ranging from 18 to 28). For WTR and HL, we have a correlation of 0.34 ($p = 0.16$) for men and -0.24 ($p = 0.25$) for women; for WTR and GP, we have a correlation of -0.30 ($p = 0.23$) for men and 0.19 ($p = 0.33$) for women.[16]

With the Holt-Laury mechanism, most respondents make inconsistent or dominated choices. Differing levels of comprehension may correlate with differences in preferences, making policy recommendations precarious. So comprehension is a serious issue, as there seems to be no good way to account for inconsistency without excluding a large portion of the data.[17] Thus, caution might be advisable concerning using a relatively sophisticated mechanism in a rural, developing-country environment.

Simply asking people about how prepared they are to take risks has low cost and easy implementation. However, here it delivers results that differ from most work on risk-preference elicitation, as there are far more choices of full risk tolerance for women in our data than in the German survey data. Indeed, a serious question is whether this is a meaningful result. In fact, it

---

[16] We do not wish to take a stand regarding which measure is "better". Nevertheless, the regressions in Tables A3-A5 suggest that WTR is the best for prediction purposes. Bear in mind that in order to frame WTR properly, one needs to conduct a qualitative study/pilot in order to detect how people perceive "risk", which we did on a small scale. We did not vary framed versus unframed here, but from pilot work in Senegal (e.g., Aflagah, Bernard, and Viceisza, 2015), we know that this can matter.

[17] Hirschauer et al. (2014) demonstrate that including inconsistent subjects in a HL analysis "will bias the mean as well as the variance of the risk attitudes".

is the women who handle the business end of matters in this environment and are more likely to engage in microfinance etc. and are thus more likely to take/face such risks. Tables A3-A5 in the online Appendix show that despite the degenerate distribution of WTR, when women report higher willingness to take risk, it correlates significantly with output risk in the expected direction. Although it correlates in the opposite direction for crop risk, but if the dependent variable captures past risk taking/facing experiences, this can easily be explained. In any case, this seems to be consistent with the real-life experience for women in this culture and may well have the most explanatory power in our data. Cultural influences may very well mediate differences in behavior across gender. For example, in Gneezy, Leonard, and List (2009), women choose the competitive environment far more frequently than men in a matrilineal society.

Finally, the adapted Gneezy-Potters task, involving equally likely alternatives and a fixed rate of return, led to results not dramatically different from those of previous work in developed and developing nations. However, it is impossible to detect confusion with this mechanism, and the choice pattern and the strong spike at 50 percent are consistent with arbitrary-choice behavior (although this 50-percent result does not hold when considering the choices of men and women separately). It could also be that choices are quite different with slight changes in the instructions, so that we see these results as indicative rather than conclusive.

Our study is an early attempt to gather data on the important methodological question of the effectiveness of different sorts of risk-elicitation mechanisms in developing nations. It is obvious that further research is needed. Nevertheless, our results provide some guidance to researchers in this area. Specifically, some possible rules-of-thumb when choosing risk-elicitation instruments are:

14

a. Conduct pilot studies.
b. Employ multiple risk-elicitation instruments.
c. When using complex instruments, consider the framing of one's instrument, particularly in relation to probabilities and visualization.
d. Prior to designing one's instruments, conduct a careful qualitative study aimed at understanding how one's subjects perceive "risk" and what "risk-taking" means in the day-to-day context.

ACKNOWLEDGEMENTS

REFERENCES

Aflagah, F., T. Bernard, and A. Viceisza (2015), "Communication and Coordination: Experimental Evidence from Farmer Groups in Senegal, Unpublished Manuscript.

Ashraf, B. N. 2009. "Spousal Control and Intra-Household Decision Making: An Experimental Study in the Philippines." *The American Economic Review* 99(4): 1245–1277.

Attanasio, O., A. Barr, J. Cárdenas, G. Genicot, and C. Meghir. 2011. "Risk Pooling, Risk Preferences, and Social Networks." UCL Working Paper, London UK.

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro. 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics* 112 (May): 537-579.

Bellemare, C., M. Krause, S. Kröger, and C. Zhang. 2005. "Myopic Loss Aversion: Information Feedback vs. Investment Flexibility." *Economics Letters* 87: 291–439.

Binswanger, H. P. 1980. "Attitudes toward Risk, Experimental Measurement in Rural India." *American Journal of Agricultural Economics* 62(August): 395-407.

Brick, K., M. Visser, and J. Burns. 2012. "Risk aversion: experimental evidence from South African fishing communities." *American Journal of Agricultural Economics* 94: 133–152.

Callen, M., M. Isaqzadeh, and C. Sprenger. 2014. "Violence and Risk Preferences: Artefactual and Experimental Evidence from Afghanistan" *American Economic Review* 104(1): 123-148.

Cameron, L., and M. Shah. "Risk-Taking Behavior in the Wake of Natural Disasters" *Journal of Human Resources (forthcoming)*.

Castillo, M., and M. Carter. 2011. "Trustworthiness and Social Capital in South Africa: Analysis of Actual Living Standards Data and Artefactual Field Experiments." *Economic Development and Cultural Change* 59(4): 695–722.

Castillo, M., C. R. Cotla, R. Petrie, and M. Torero. 2014. "Measuring the Preferences of the Poor: Evidence from a Large Randomized Test of Alternative Experimental Methods." Mimeo.

Charness, G., and G. Genicot. 2009. "Informal Risk Sharing in an Infinite-Horizon Experiment." *Economic Journal* 119(537): 796–825.

Charness, G., and U. Gneezy. 2010. "Portfolio Choice and Risk Attitudes: An Experiment." *Economic Inquiry* 48(1): 133–146.

_____. 2012. "Strong Evidence for Gender Differences in Experimental Investment." *Journal of Economic Behavior and Organization* 83(1): 50-58.

Charness, G., U. Gneezy, and A. Imas. 2013. "Experimental Methods: Eliciting Risk Preferences." *Journal of Economic Behavior and Organization* 87(3): 43-51.

Charness, G., and M. Villeval. 2009. "Cooperation, Competition, and Risk Attitudes: An Intergenerational Field and Laboratory Experiment." *The American Economic Review* 99(3): 956–978.

Chuang, Y., and L. Schechter. 2015. "Stability of Experimental and Survey Measures of Risk, Time, and Social Preferences Over Multiple Years." *Journal of Development Economics (forthcoming)*.

Cook, J., S. Chatterjee, D. Sur, and D. Whittington. 2013. "Measuring risk aversion among the urban poor in Kolkata, India." *Applied Economics Letters* 20(1), 1–9.

Crosetto, P. and A. Filippin, 2013. "A Theoretical and Experimental Appraisal of Five Risk Elicitation Methods." Jena Economic Research Paper No. 2013-009.

Dave, C., C. Eckel, C. Johnson, and C. Rojas. 2010. "Eliciting Risk Preferences: When is Simple Better?" *Journal of Risk and Uncertainty* 41(3): 219–243.

De Brauw, A., and P. Eozonou. 2011. "Measuring Risk Attitudes among Mozambican Farmers." IFPRI Harvest Plus Working Paper 6, Washington DC.

Delavande, A., X. Giné, and D. McKenzie. 2011. "Measuring subjective expectations in developing countries: A critical review and new evidence." *Journal of Development Economics* 94(2): 151-163.

Doerr, U., T. Mahmoud, and U. Schmidt. 2011. "Overconfidence and Risk Management of Ethiopian Farmers." Working Paper.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences." *Journal of the European Economic Association* 9(3): 522–550.

Dreber, A., and M. Hoffman. 2007. "2D:4D and Risk Aversion: Evidence that the Gender Gap in Preferences is Partly Biological." Stockholm School of Economics Working Paper, Stockholm.

Eckel, C. C. and P. J. Grossman. 2002. "Sex differences and statistical stereotyping in attitudes toward financial risk." *Evolution and Human Behavior* 23: 281-295.

Engle-Warnick, J., J. Escobal, and S. Laszlo. 2011. "Ambiguity Aversion and Portfolio Choice in Small-Scale Peruvian Farming." *The B.E. Journal of Economic Analysis & Policy* 11(1): 68.

Fellner, G., and M. Sutter. 2009. "Causes, consequences, and cures of myopic loss aversion – An experimental investigation." *Economic Journal* 119: 47-60.

Filippin, A. and P. Crosetto. 2014. "A Reconsideration of Gender Differences in Risk Attitudes," Working Papers 2014-01, Grenoble Applied Economics Laboratory (GAEL).

Galarza, F. 2009. "Choices under Risk in Rural Peru." Munich Personal RePec Archive Working Paper 17708, Munich.

Giné, X., P. Jakiela, D. Karlan, and J. Morduch. 2010. "Microfinance Games." *American Economic Journal: Applied Economics* 2(3): 60–95.

Gneezy, U., and J. Potters. 1997. "An Experiment on Risk Taking and Evaluation Periods." *Quarterly Journal of Economics* 112: 631–645.

Gneezy, U., K. Leonard, and J. List. 2009. "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society." *Econometrica* 77: 1637–1664.

Gong, B., and C-L. Yang. Forthcoming. "Gender Differences in Risk Attitudes: Field Experiments on the Matrilineal Mosuo and the Patriarchal Yi." *Journal of Economic Behavior and Organization*.

Haigh, M., and J. List. 2005. "Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis." *Journal of Finance* 60(1): 523–534.

Harrison, G., 2011. "Experimental Methods and the Welfare Evaluation of Policy Lotteries." *European Review of Agricultural Economics* 38(3): 335–360.

Harrison, G., J. List, and C. Towe. 2007. "Naturally Occurring Preferences and Exogenous Laboratory Experiments: A Case Study of Risk Aversion." *Econometrica* 75: 433–458.

Harrison, G., S. Humphrey, and A. Verschoor. 2010. "Choice under Uncertainty: Evidence from Ethiopia, India, and Uganda." *Economic Journal* 120 (March): 80–104.

Heinemann, F., R. Nagel, and P. Ockenfels. 2009. "Measuring strategic uncertainty in coordination games." *Review of Economic Studies* 76(1): 181–221.

Hill, R. V., and A. C. G. Viceisza. 2012. "A field experiment on the impact of weather shocks and insurance on risky investment." *Experimental Economics* 15(2): 341-371.

Hill, R., E. Maruyama, and A. Viceisza. 2012. "Breaking the Norm: An Empirical Investigation into the Unraveling of Good Behavior." *Journal of Development Economics* 99(1): 150-162.

Ihli, H., B. Chiputwa, G-F. Bauermeister, and O. Musshoff. 2013. "Measuring risk attitudes of smallholder farmers in Uganda: How consistent are results of different methods?" The Second International Agricultural Risk, Finance, and Insurance Conference Paper.

Hirschauer, N., S. Maart, and O. Musshoff. 2012. "Eliciting risk attitudes: How to avoid mean and variance bias in Holt-and-Laury lotteries." *Applied Economics Letters* 21(1): 35-38.

Holt, C., and S. Laury. 2002. "Risk Aversion and Incentive Effects." *The American Economic Review* 92 (5): 1644–1655.

Jacobson, S., and R. Petrie. 2009. "Learning from Mistakes: What do Inconsistent Choices over Risk Tell Us?" *Journal of Risk and Uncertainty* 38(2): 143–158.

Kimball, M.S., C. Sahm, and M. D. Shapiro. 2008. "Imputing Risk Tolerance from Survey Responses." *Journal of the American Statistical Association* 103(September): 1028-1038.

Kimball, M.S., C. Sahm, and M. D. Shapiro. 2009. "Risk Preferences in the PSID: Individual Imputations and Family Covariation." *American Economic Review Papers and Proceedings* 99(May): 363-368.

Lammers, J., M. Lau, and H. Verbon. 2006. "Perceived HIV- contamination risk, risk aversion and time preferences (A Laboratory Experiment in South Africa)." Conference paper.

Langer, T., and M. Weber. 2004. "Does Binding or Feedback Influence Myopic Loss Aversion? An Experimental Analysis." University of Muenster Working Paper, Muenster.

Lejuez, C.W., J. P. Read, C. W. Kahler, J. B. Richards, S. E. Ramsey, G. L. Stuart, D. R. Strong, et al. 2002. "Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART)." *Journal of Experimental Psychology: Applied* 8(2): 75–84.

Meier, S., and C. Sprenger. 2010. "Present-Biased Preferences and Credit Card Borrowing." *American Economic Journal: Applied Economics* 2(1): 193–210.

Nielsen, T., and M. Zeller. 2014. "The Impact of Shocks on Risk Preference Changes between Seasons for Smallholder Farmers in Vietnam." Working Paper.

Stockman, C. 2006. "Differences and Similarities in Health Risk Preferences across Two Subject Cohorts." California University of Pennsylvania Working Paper, California, PA.

Tanaka, T., C. Camerer, and Q. Nguyen. 2010. "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam." *American Economic Review* 100(1): 557–571.

Treibich, C. 2015. Are Survey Risk Aversion Measurements Adequate in a Low Income Context? Working Paper.

Viceisza, A. 2012. "Treating the field as a lab: A basic guide to conducting economics experiments for policymaking." *Food Security in Practice Series Technical Guide*. Washington, DC: International Food Policy Research Institute.

Viceisza, A. 2015. "Creating a Lab in the Field: Economics Experiments for Policymaking." *Journal of Economic Surveys (forthcoming)*.

Weber, E.U., A. R. Blais, N. E. Betz. 2002. "A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors." *Journal of Behavioral Decision Making* 15(4): 263–290.

Wik, M., T. Kebede, O. Bergland, and S. Holden. 2004. "On the measurement of risk aversion from experimental data." *Applied Economics* 36: 2443–2451.

Yesuf, M., and R. A. Bluffstone. 2009. "Poverty, risk aversion, and path dependence in low income countries: Experimental evidence from Ethiopia." *American Journal of Agricultural Economics* 91: 1022–1037.

Outline

The typical session comprised:

1) Survey 1
2) An introduction of the experimenter, translator, assistant(s), IFPRI, and the project
3) Instructions on the game (more below)
4) The game (sessions 1 and 3 were HL and sessions 2 and 4 were GP)
5) An opportunity to revise certain responses to survey 1
6) The lottery
   a. For HL, this comprised three draws:
       i. The first draw was to determine the row for which the subjects would get paid. In other words, this determined the distribution of the weather.
       ii. The second draw was to determine whether those who chose indifference would be treated as "magasin/storage A" or "magasin/storage B".
       iii. The third draw was to determine the weather, that is, whether it was good or bad.
   b. For GP, this comprised one draw that determined the weather.
7) Survey 2 (demographic survey)
8) Payment

Layout

1) The experiments were conducted in classrooms in the local schools. Boxes were used as dividers to give privacy.
2) The typical layout of the room was as follows:

FRONT OF ROOM (experimenter, translator, and white board)			
Seat 1	Seat 7	Seat 13	Seat 19
Seat 2	Seat 8	Seat 14	Seat 20
Seat 3	Seat 9	Seat 15	Seat 21
Seat 4	Seat 10	Seat 16	Seat 22
Seat 5	Seat 11	Seat 17	Seat 23
Seat 6	Seat 12	Seat 18	Seat 24
BACK OF ROOM (assistant experimenter/cashier)			

Survey 1

1) This survey was administered prior to anything else, that is, as soon as subjects walked into the laboratory.
2) The survey comprised the following questions:
   a. On the following scale of 1 to 10, please indicate how willing you are to take risks.
   b. In your day-to-day life, what do you consider to be a risky decision? Please describe using one or more examples.
   c. How much is 15% of 2,000 FCFA? If you don't know, put an **X**.

Introduction

1) The experimenter introduced himself, the translator and the two assistants. Typically, the main assistant experimenter was not introduced till the end.
2) The experimenter introduced IFPRI and the larger project, typically as follows:
   a. IFPRI is an institute in the United States.
   b. We are conducting a research project on farmer groups, their activities and so on.
   c. We have been holding discussions with farmers across many parts of Senegal. In particular, we have talked to farmers in Diourbel/Thies, but we have not been here before.
   d. For the upcoming task, we will pay you for the decisions that you make. We pay you for two purposes:
      i. Because you came here today and are spending your time with us. This is time in which you could be doing something else, so we pay you for that reason.
      ii. Also, we would like you to take this decision seriously, as you do any other decision in real life.

Instructions and Game

1) Game 1 (HL)
   a. The experimenter handed out the sheet of paper for the HL game.
   b. The experimenter first asked subjects what they thought the pictures on the form represented.
      i. This served as an icebreaker. It basically enabled subjects to start thinking about the material and the decisions they would be presented with during the session.
      ii. In some cases the storage was seen as a school and the good weather was perceived as clouds, but typically subjects soon realized that the task would have something to do with storage and good/bad weather.
   c. After this mini brainstorming, the experimenter explained the following steps:
      i. The brainstorming has shown that the task today has to do with storage and the weather.
      ii. Specifically, suppose there are two types of storage rooms (A/Abdu and B/Bara) that contain two different types of fertilizer/angrais (A/Abdu and B/Bara). We are going to ask you which of these two fertilizers you prefer.
      iii. How are these two fertilizers different? Let's focus on the first row of the first page.
         1. Fertilizer A in magasin Abdu
            a. The fertilizer in magasin Abdu gives FCFA 1,000 as income from production in times of bad weather and FCFA 800 in times of good weather.
               i. Explain payoff and how it is associated with good/bad weather.
               ii. Quiz people on how much the payoff is in times of good/bad weather.
         2. Fertilizer B in magasin Bara

    a. Now, let's look at the fertilizer in magasin Bara. What is different about it? Well, this fertilizer gives FCFA 2,000 as income from production in times of good weather but FCFA 100 in times of bad weather.

    b. So, the difference between the two fertilizers is that the one in magasin Bara pays MORE in times of good weather but LESS in times of bad weather.

        i. Similar to above, explanation and quiz.

3. Recap: So, we've seen that there are two types of fertilizer, the one in magasin Abdu and the one in magasin Bara. We also know that they're affected by the weather.

4. *What do we know about the weather?*

    a. As in real life, sometimes the weather is good and sometimes the weather is bad.

    b. These 10 numbers (1, 2, 3,…, 10) represent 10 years of weather.

    c. In the first row, 1 out of 10 years the weather is good and 9 out of 10 years the weather is bad.

        i. The number 1 represents the year that the weather is good.

        ii. The numbers 2, 3,…, 10 represent the years that the weather is bad.

        iii. The numbers in the columns of good/bad weather represent the years that weather can be good/bad.

        iv. So, note that the weather is the same for magasin Abdu and magasin Bara. *What is different is the income from production you get depending on weather being good or bad.*

    d. Questions/quiz for understanding

        i. How many years can the weather be good in row 1?

        ii. How many years can it be bad?

        iii. What is the income from production if the weather is bad?

            1. Depends on whether you buy Abdu or Bara.

        iv. Suppose you buy Abdu and the weather is good, what is your income from production? How about Bara?

        v. How about if the weather is bad?

d. This explains row 1. How are the other rows different from row 1?

    i. Notice that when we go from row 1 to row 2, the only aspect that changes is the number of years that weather can be good/bad. That is, the income from production does NOT change. However, in row 2 the number of years that weather can be good is 2 and the number of years that the weather can be bad is 8.

        1. Typically, the experimenter showed the years with the numbers 1, 2 in the left hand and 3, 4,…, 10 in the right hand.

ii. Now, what happens if we go from row 2 to row 3? Now, weather can be good 3 out of 10 years and bad 7 out of 10 years.

iii. This process was continued up to row 10.

1. At this stage, subjects typically smiled indicating their understanding that in row 10 the weather was always good.

e. So, we are going to ask you to make a decision for each of the rows: Abdu or Bara. If you do not know which one to choose, you can choose I for "indifferent."

f. Is this clear?

i. At this point, a row was selected to quiz subjects again. Questions were asked with regard to the probabilities and earnings.

ii. Then, subjects were informed that only one row would be selected for payment. The exact procedures for selecting the row and drawing/simulating the weather were typically explained when the lottery was drawn in order to avoid too much information prior to decisions being made.

iii. Then, decisions were made.

2) Game 2 (GP)

a. The game sheets (appendix) were handed out. Subjects were prompted on the images at the top as an icebreaker exercise. In both sessions they recognized them correctly as the two types of weather.

b. The experimenter asked the subjects to imagine that they are grain farmers and they are to be given 10 kilos of seeds to plant for the new season. They are told that they can take two types of seeds–from Abdu or from Bara. At this point the experimenter emphasized that they must take a total of 10. He did this by giving them examples of the possible combinations of the seeds that they could take.

i. *Imagine that you are a grain farmer and you are given 10 kilos of seeds for free for the coming season. You can choose between two different types of seeds to take. You can either take the seeds of Abdu or of Bara. It is important that you realize you can take as many kilos of Abdu and Bara seeds as you want as long as at the end you are taking 10 kilos in total—no more and no less. On your answer sheet you will see a place for you to write how many Abdu seeds and Bara seeds you will take.*

c. Next the experimenter explained how the seeds are different. Abdu seed is of higher quality than Bara seed but is more vulnerable to the weather. That is, when there is good weather the Abdu seed produces a harvest that sells for FCFA 300. When the weather is bad the harvest is so bad that it cannot be sold, eaten, or fed to the animals. On the other hand, the Bara seed does not respond to the weather and always gives FCFA 100 francs per kilo.

d. Next the experimenter proceeded to go through the columns for the Abdu seed and explain how different quantities of Abdu seed affect one's income from the harvest given good weather. What was emphasized through examples was that 300 times the number of kilos of the seeds determines the income, which is then provided for the subjects in the column 'xalis…. Abdu' on the side with good weather. This was done to the point where the experimenter felt comfortable with their understanding of the derivation of their income. Next the experimenter explained the bad weather columns for the Abdu seed, which was provided in the column 'xalis … Abdu on' the side

with bad weather. This was always zero. Again examples were given until the experimenter felt comfortable with their understanding.

e. The same procedure followed for the Bara seed: first with the good weather 'xalis… Bara' column and then with the bad weather 'xalis…Bara' column. It was emphasized that there was no difference between the columns.

f. The experimenter explained the total income for any given type of weather–by adding the columns of 'xalis… Abdu' and 'xalis… Bara'. It was explained that this number was indicated in the column 'li ngay … xalis'.

g. The experimenter asked subjects specific questions such as:

    i. *If the weather is good and one had 5 kilos of bara seeds how many Abdu seeds does one have? How much money does one make from these Bara seeds? ... from these Abdu seeds? In total?* Then he repeated for other combinations… (1 and 9, 3 and 7… etc. – each time varying the weather)

    ii. *Whenever one subject seemed to dominate by answering correctly in succession, the experimenter asked the translator to explain to him that we would like to hear from other people as well. These examples were repeated until the experimenter felt confident about the understanding of the subjects.*

h. The experimenter explained how weather was unknown at the time of the decision, how this was realistic, and how it was to be determined, from a box, with equal probability of the two types of weather. That is, the weather was to be determined from a box where drawing cards with numbers 1–5 would correspond to good weather, while drawing numbers 6–10 would correspond to bad weather. These numbers were the same as the ones used for the HL task.

i. Decisions were then made.

Lottery

The lotteries were conducted according to the procedures described previously. Typically, we let one of the subjects draw. Papers with numbers 1 through 10 were drawn from a bag.

Survey 2 (demographics)

1) This survey was administered after the main task and comprised the following questions:

    a. Education level

    b. Marital status

    c. Number of children

    d. Primary occupation

    e. How often do you find yourself short of cash?

    f. How much do you agree with the statement "Most people can be trusted"? (1=Strongly Disagree, 2=Disagree, 3=Slightly Disagree, 4=Neither Agree or Disagree, 5=Slightly Agree, 6=Agree, 7=Strongly Agree)

    g. How many cigarettes do you smoke per day?

    h. On a scale of 1 to 10, how patient do you consider yourself?

Payment

After all these steps were done, subjects were called by their seat number—one by one—to get paid in private by the assistant experimenter. They were also paid a fixed fee for showing up.

TABLES AND FIGURES

Table 1—Taxonomy of risk-elicitation approaches and related sample studies

		(2)	
		Complex I	Simple (S)
(1)	Incentivized (I)	IC: Brick, Visser, and Burns (2012) Holt and Laury (2002)* Tanaka, Camerer, and Nguyen (2010)**	IS: Binswanger (1980) Eckel and Grossman (2002) Gneezy and Potters (1997)*
	Non-Incentivized (NI)	NIC: Barsky et al. (1997) Kimball et al. (2008, 2009) Weber, Blais, Betz (2002)	NIS: Dohmen et al. (2011)* Lejuez et al. (2002) Treibich (2015)

*Indicates an instrument for which this study collects new data. **Note that to our knowledge Tanaka et al. (2010) is the only study that uses prospect theory and finds: "People living in poor countries are not necessarily afraid of uncertainty."

Table 2—Inconsistency rates reported by selected developing country field studies

Study	Method	Protocol	Subjects	Inconsistency	Implications
Lammers, Lau, Verbon (2006)	HL	- Paper-based - Neutral framing - Classroom explanation as part of experiment session - Real stakes	South African students	66.5%[a]	- Not cited due to preliminary nature of the analysis
Galarza (2009)	HL	- Paper-based - Mixed framing with e.g. safe (risky) lottery called sol (luna) - Classroom explanation as part of experiment session - Real stakes	Peruvian farmers	52%[a]	- Some evidence that multiple switching behavior explained by nonlinear probability weighting made in a context of random calculation mistakes - Majority of farmers characterized by prospect theory
De Brauw and Eozenou (2011)	HL	- Paper-based - Loaded framing with lotteries in terms of sweet potato varieties with varying yields - Individual explanation as part of survey - Hypothetical stakes	Mozambican farmers	14%[a]	- About three-fourths of the sample develops risk preferences by rank dependent utility - By making the CRRA assumption, they poorly predict risk preferences among those who are less risk averse
Doerr et al. (2011)	HL	- Paper-based - Neutral framing with lotteries displayed as balls in urns - Individual explanation as part of survey - Real stakes	Ethiopian farmers	39%[b]	- Risk taking of farmers seems to be unrelated to risk preferences but highly correlated to overconfidence measures
Jacobson and Petrie (JP, 2009)	Not HL*	- Paper-based - Neutral framing with lotteries called A and B - Individual explanation as part of survey - Real and hypothetical stakes	Rwandan farmers	52%[c]	- While risk aversion alone does not explain financial decisions, risk aversion and inconsistent choices interact in significant and sensible ways - Mistakes correlate with less than optimal behavior
Engle-Warnick et al. (2011)	Not HL**	- Paper-based - Neutral framing with lotteries displayed as color chips in a bag - Classroom explanation as part of experiment session - Real stakes	Peruvian farmers	52%[d]	- They combine experimental and survey data to find that ambiguity aversion predicts actual technology choices on the farm

Table notes:

[a] These rates are based on subjects who switched more than once in the HL task.

[b] This rate is based on a test of the common ratio effect, which is based on the Allais paradox, paired with the HL task.

[c] This rate is based on the number of people who make at least one mistake.

[d] This rate is based on the number of people who choose the dominated lottery at least twice. If we consider those who choose the dominated lottery at least once, the rate would be 76%.

[*] The format of the experiment is similar to HL, but with several key differences. HL keep payoffs constant and vary the probabilities of receiving the high and low outcomes. In JP, the probability is always 50–50, and the payoffs are varied. Also, HL present the lotteries all at once to the subjects. JP present lottery pairs sequentially. Finally, JP also present lotteries over losses.

[**] Their decomposition resembles the instrument in HL. Each row in the decomposition corresponds to a single binary choice between two alternative gambles. Beginning with the first row of choices and moving down, an expected utility maximizer will at some point switch from the left-hand side gamble with lower variance to the right-hand side gamble with a higher variance and slightly higher expected utility.

Table 3—Holt-Laury lottery choices

Option A	Option B	Expected payoff difference
1/10 of $2.00, 9/10 $1.60	1/10 of $3.85, 9/10 of $0.19	$1.17
2/10 of $2.00, 8/10 $1.60	2/10 of $3.85, 8/10 of $0.19	$0.83
3/10 of $2.00, 7/10 $1.60	3/10 of $3.85, 7/10 of $0.19	$0.50
4/10 of $2.00, 6/10 $1.60	4/10 of $3.85, 6/10 of $0.19	$0.16
5/10 of $2.00, 5/10 $1.60	5/10 of $3.85, 5/10 of $0.19	-$0.18
6/10 of $2.00, 4/10 $1.60	6/10 of $3.85, 4/10 of $0.19	-$0.51
7/10 of $2.00, 3/10 $1.60	7/10 of $3.85, 3/10 of $0.19	-$0.85
8/10 of $2.00, 2/10 $1.60	8/10 of $3.85, 2/10 of $0.19	-$1.18
9/10 of $2.00, 1/10 $1.60	9/10 of $3.85, 1/10 of $0.19	-$1.52
10/10 of $2.00, 0/10 $1.60	10/10 of $3.85, 0/10 of $0.19	-$1.85

Source: Holt and Laury (2002)

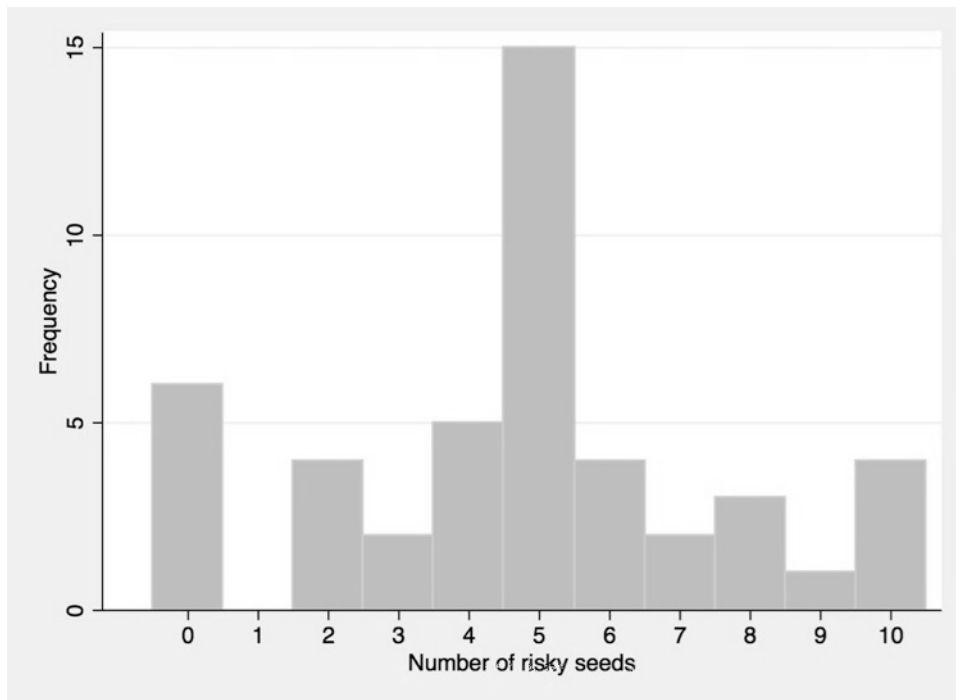**Figure 1—Senegal HL data (probability of choosing option A)**

26
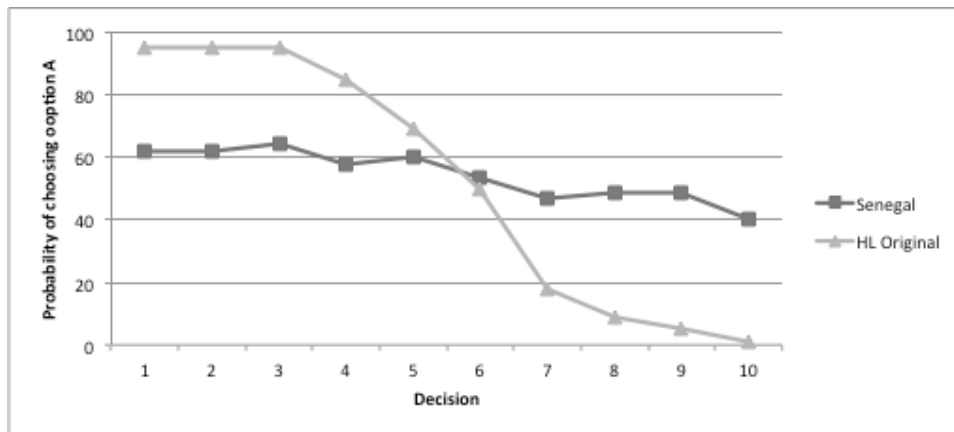
Figure 2—Histograms of GP data

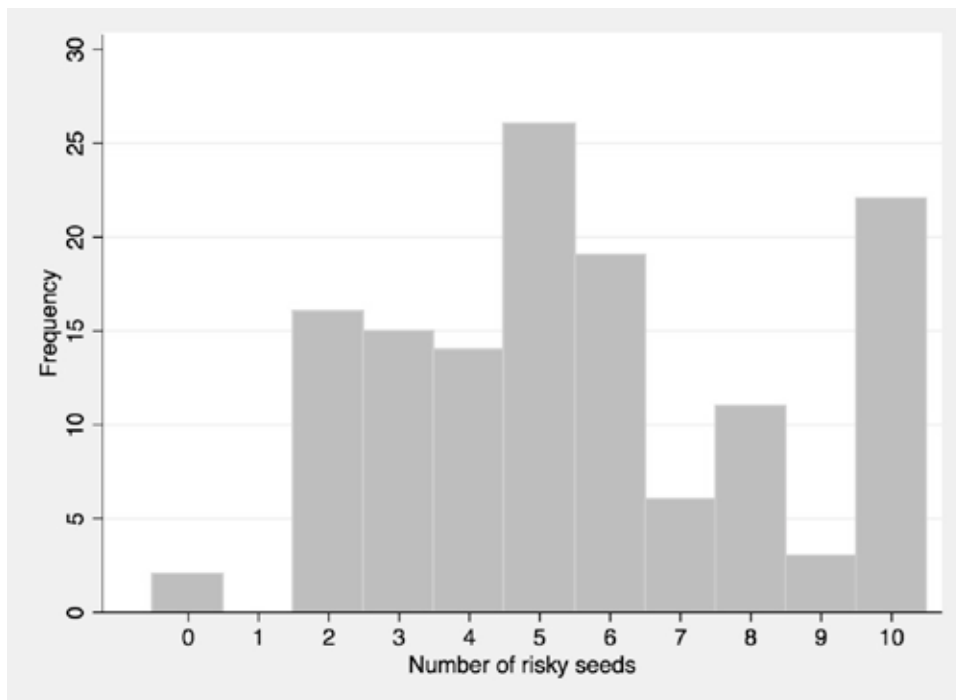Panel A: Senegal

Panel B: Charness and Villeval (2009)

Figure 3—Histograms of Senegal WTR data

Panel A: Senegal
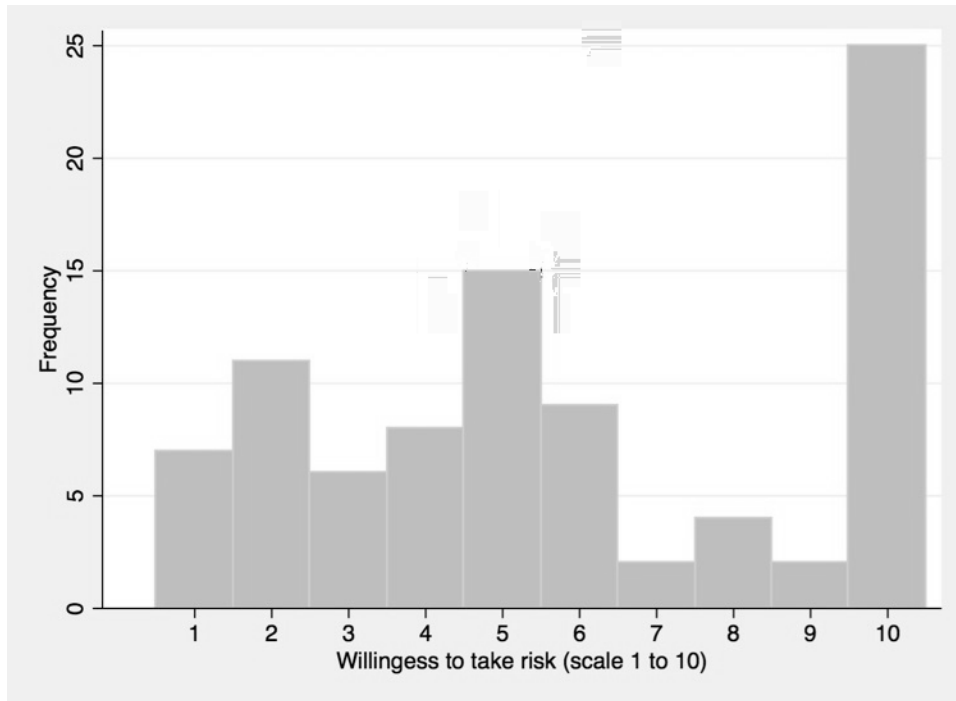
Panel B: Dohmen et al. (2011)
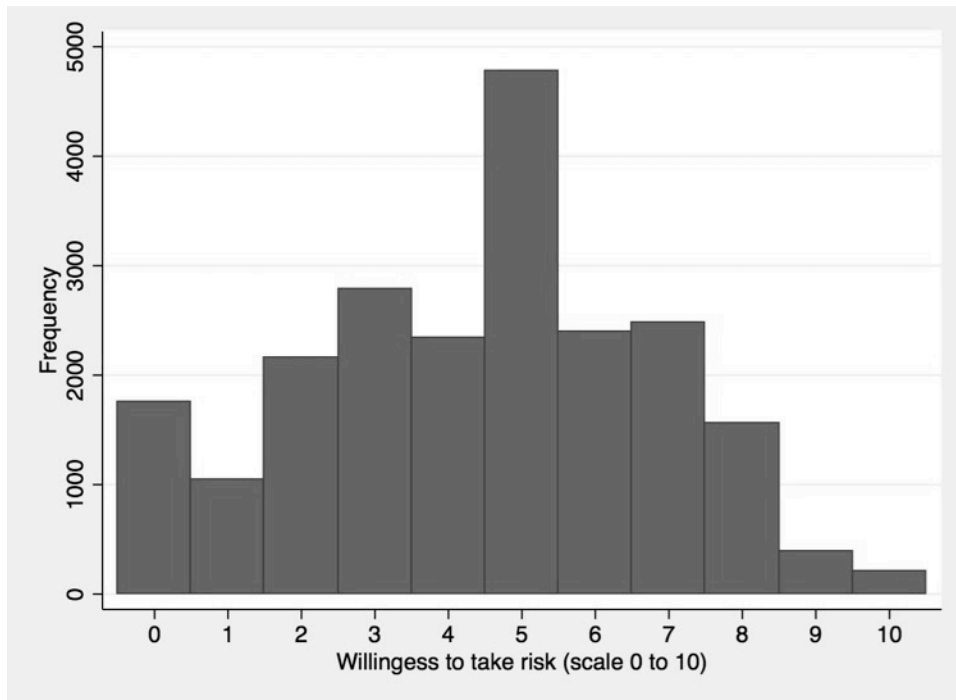
Table A1—Inconsistency rates reported by selected[+] developing country field studies

Study	Method	Protocol	Subjects	Rate	Implications
Lammers, Lau, Verbon (2006)	HL	- Paper-based - Neutral framing - Classroom explanation as part of experiment session - Real stakes	South African students	66.5%[a]	- Not cited due to preliminary nature of the analysis
Galarza (2009)	HL	- Paper-based - Mixed framing with e.g. safe (risky) lottery called sol (luna) - Classroom explanation as part of experiment session - Real stakes	Peruvian farmers	52%[a]	- Some evidence that multiple switching behavior explained by nonlinear probability weighting made in a context of random calculation mistakes - Majority of farmers characterized by prospect theory
De Brauw and Eozenou (2011)	HL	- Paper-based - Loaded framing with lotteries in terms of sweet potato varieties with varying yields - Individual explanation as part of survey - Hypothetical stakes	Mozam-bican farmers	14%[a]	- About three-fourths of the sample develops risk preferences by rank dependent utility - By making the CRRA assumption, they poorly predict risk preferences among those who are less risk averse
Doerr et al. (2011)	HL	- Paper-based - Neutral framing with lotteries displayed as balls in urns - Individual explanation as part of survey - Real stakes	Ethiopian farmers	39%[b]	- Risk taking of farmers seems to be unrelated to risk preferences but highly correlated to overconfidence measures
Jacobson and Petrie (JP, 2009)	Not HL*	- Paper-based - Neutral framing with lotteries called A and B - Individual explanation as part of survey - Real and hypothetical stakes	Rwandan farmers	52%[c]	- While risk aversion alone does not explain financial decisions, risk aversion and inconsistent choices interact in significant ways - Mistakes correlate with less than optimal behavior
Engle-Warnick et al. (2011)	Not HL**	- Paper-based - Neutral framing with lotteries displayed as color chips in a bag - Classroom explanation as part of experiment session - Real stakes	Peruvian farmers	52%[d]	- They combine experimental and survey data to find that ambiguity aversion predicts actual technology choices on the farm
Brick, Visser, and Burns (2012)	Not HL***	- Neutral framing of lotteries - Real stakes	South African fishers	41%	- They estimate risk attitudes and find that female fishers and rights holders are more risk averse than their male counterparts, while rights holders are found to be less risk averse relative to subjects without fishing rights.

Table notes:

[+] This table reviews some main studies that conduct MPL tasks in a developing country context. Due to space and given the fact that this is not a literature review, we do not summarize more studies. Additional references include but are not limited to: Wik et al. (2004) in Zambia; Yesuf and Bluffstone (2009) in Ethiopia; Cook et al. 2013 in India; and Chuang and Schechter (2015) who review evidence from multiple countries.

[a] These rates are based on subjects who switched more than once in the HL task.

[b] This rate is based on a test of the common ratio effect, which is based on the Allais paradox, paired with the HL task.
[c] This rate is based on the number of people who make at least one mistake.
[d] This rate is based on the number of people who choose the dominated lottery at least twice. If we consider those who choose the dominated lottery at least once, the rate would be 76%.
[*] The format of the experiment is similar to HL, but with several key differences. HL keep payoffs constant and vary the probabilities of receiving the high and low outcomes. In JP, the probability is always 50–50, and the payoffs are varied. Also, HL present the lotteries all at once to the subjects. JP present lottery pairs sequentially. Finally, JP also present lotteries over losses.
[**] Their decomposition resembles the instrument in HL. Each row in the decomposition corresponds to a single binary choice between two alternative gambles. Beginning with the first row of choices and moving down, an expected utility 30aximize will at some point switch from the left-hand side gamble with lower variance to the right-hand side gamble with a higher variance and slightly higher expected utility.
[***] Their main instrument presents subjects with a multiple price list, but with varying payoffs as opposed to varying probabilities similar to HL.

Table A2—Investment choices in other studies

Study	Participants	Periods	Avg. Male Investment (N)	Avg. Female Investment (N)
Langer and Weber (2004)	Finance students, Mannheim	30	64.62 (93)	58.70 (14)
Haigh and List (2005)	Professional traders, CBOT	9	58.30 (50)	55.59 (8)
Fellner and Sutter (2009)	Undergrads, Jena	18	57.44 (39)	49.04 (79)
Bellemare et al. (2005)	Undergrads, Tilburg	9	45.48 (95)	42.73 (40)
Charness and Genicot (2009)	Undergrads, UCLA	1	59.22 (41)	52.23 (53)
Dreber and Hoffman (2007)	Students, Stockholm	1	69.60 (92)	50.00 (55)
Gneezy, Leonard, and List (2009)	Villagers in Tanzania and India	1	50.00 (157)	50.06 (157)
Ertac and Gurdal (2010)[a]	Undergrads, Turkey	1	72.32 (79)	54.29 (49)
Gong and Yang (forthcoming)	Matrilineal villagers in China	1	53.9 (31)	32.5 (36)
Gong and Yang (forthcoming)	Patrilineal villagers in China	1	37.3 (37)	4.3 (28)
Charness and Villeval (2009)	Workers at French firms	1	62.73 (33)	53.2 (29)
Charness and Gneezy (2010)	Undergrads, UCSB	1	75.82 (137)	60.25 (64)

Note: [a] We include only the individual risk decisions where there is a positive expected return from investing in the risky asset. A similar gender difference applies in the other cases.

Table A3: Does HL predict risk taking?

VARIABLES	(1) risk_output	(2) risk_crop	(3) risk_price	(4) risk_credit
RA (HL proxy)	-0.010	0.007	0.053*	0.015
	(0.040)	(0.130)	(0.030)	(0.019)
HL consistency	-0.085	-0.059	0.327*	0.137
	(0.193)	(0.057)	(0.165)	(0.114)
WTR	-0.016	0.019	-0.014	0.024
	(0.045)	(0.018)	(0.036)	(0.022)
Gender	-0.224	0.417*	0.087	0.253
	(0.426)	(0.242)	(0.436)	(0.164)
Gender*RA	0.008	-0.027	-0.010	-0.002
	(0.052)	(0.018)	(0.046)	(0.022)
Gender*WTR	0.101*	-0.054*	-0.037	-0.028
	(0.055)	(0.030)	(0.053)	(0.024)
Constant	0.184	-0.033	0.300	-0.305
	(0.319)	(0.144)	(0.316)	(0.185)
Observations	43	43	43	43
R-squared	0.240	0.154	0.188	0.172
Adjusted R^2	0.088	-0.015	0.025	0.007

Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Session dummies included in all specifications.

Definition of dependent variables: After explaining what risk is, we asked respondents to report if they knew of any situations or decisions they usually encounter/take that can be seen as risky. We then coded the dummies risk_output, risk_crop, risk_price, and risk_credit based on similar keywords. In all cases, 1=yes if the respondent took risk in that dimension.

Definition of independent variables: RA=number of times a respondent chose the safe lottery; HL consistency=1 if the respondent was consistent in the HL task (0 otherwise); WTR=willingness to take risk on a scale of 1 to 10; Gender=1 if the respondent is female (0 otherwise).

Table A4: Does GP predict risk taking?

VARIABLES	(1) risk_output	(2) risk_crop	(3) risk_price	(4) risk_credit
GP	0.027	-0.002	0.011	0.011
	(0.046)	(0.0145)	(0.038)	(0.038)
WTR	-0.007	0.008	-0.030	-0.030
	(0.066)	(0.020)	(0.056)	(0.056)
Gender	-0.166	-0.090	0.565	0.517
	(0.467)	(0.159)	(0.410)	(0.412)
Gender*GP	-0.016	0.002	-0.023	-0.019
	(0.051)	(0.014)	(0.044)	(0.041)
Gender*WTR	0.024	-0.008	0.013	0.024
	(0.075)	(0.020)	(0.066)	(0.061)
Constant	0.857	0.090	-0.255	-0.398
	(0.542)	(0.159)	(0.494)	(0.475)
Observations	46	46	46	46
R-squared	0.269	0.074	0.196	0.290
Adjusted R^2	0.156	-0.069	0.072	0.180

Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Session dummies included in all specifications. All variables as defined in Table A3, except for GP=number of risky seeds chosen in the GP task.

Table A5: Does WTR predict risk taking?

VARIABLES	(1) risk_output	(2) risk_crop	(3) risk_price	(4) risk_credit
WTR	-0.021	0.017	-0.007	0.013
	(0.034)	(0.016)	(0.028)	(0.023)
Gender	-0.209	0.114	0.239	0.328*
	(0.237)	(0.0947)	(0.242)	(0.181)
Gender*WTR	0.080*	-0.034*	-0.036	-0.021
	(0.041)	(0.020)	(0.036)	(0.025)
Constant	0.270	0.055	0.384	-0.255
	(0.215)	(0.096)	(0.240)	(0.162)
Observations	89	89	89	89
R-squared	0.245	0.065	0.102	0.199
Adjusted R^2	0.190	-0.003	0.036	0.141

Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Session dummies included in all specifications. All variables as defined in Table A3.

Figure A1—Holt-Laury task screen shot

SESSION _____   DAY _____   SEAT _____

	Storage A				Storage B				Choice (A, B, I)
	good weather	yield	bad weather	yield	good weather	yield	bad weather	yield	
1	1	1000	2 3 4 5 6 / 7 8 9 10	800	1	2000	2 3 4 5 6 / 7 8 9 10	100	
2	1 2	1000	3 4 5 6 / 7 8 9 10	800	1 2	2000	3 4 5 6 / 7 8 9 10	100	
3	1 2 3	1000	4 5 6 / 7 8 9 10	800	1 2 3	2000	4 5 6 / 7 8 9 10	100	
4	1 2 3 4	1000	5 6 / 7 8 9 10	800	1 2 3 4	2000	5 6 / 7 8 9 10	100	
5	1 2 3 4 5	1000	6 / 7 8 9 10	800	1 2 3 4 5	2000	6 / 7 8 9 10	100	

SESSION _____   DAY _____   SEAT _____

	Storage A				Storage B				Choice (A, B, I)
	good weather	yield	bad weather	yield	good weather	yield	bad weather	yield	
6	1 2 3 4 5 / 6	1000	7 8 9 10	800	1 2 3 4 5 / 6	2000	7 8 9 10	100	
7	1 2 3 4 5 / 6 7	1000	8 9 10	800	1 2 3 4 5 / 6 7	2000	8 9 10	100	
8	1 2 3 4 5 / 6 7 8	1000	9 10	800	1 2 3 4 5 / 6 7 8	2000	9 10	100	
9	1 2 3 4 5 / 6 7 8 9	1000	10	800	1 2 3 4 5 / 6 7 8 9	2000	10	100	
10	1 2 3 4 5 / 6 7 8 9 10	1000		800	1 2 3 4 5 / 6 7 8 9 10	2000		100	

Figure A2—Gneezy-Potters task screen shot

Day_____ Session _____ Seat _____

_____Seeds **Abdu** _____ Seeds **Bara**

Good weather | 1 | 2 | 3 | 4 | 5

Bad Weather | 6 | 7 | 8 | 9 | 10

Kilos of seed Abdu	Kilos of seed Bara	Yield Abdu	Yield Bara	Total Yield	Yield Abdu	Yield Bara	Total Yield
0	10	0	1000	1000	0	1000	1000
1	9	300	900	1200	0	900	900
2	8	600	800	1400	0	800	800
3	7	900	700	1600	0	700	700
4	6	1200	600	1800	0	600	600
5	5	1500	500	2000	0	500	500
6	4	1800	400	2200	0	400	400
7	3	2100	300	2400	0	300	300
8	2	2400	200	2600	0	200	200
9	1	2700	100	2800	0	100	100
10	0	3000	0	3000	0	0	0