On the Advantages and Disadvantages of Subjective Measures.

Salar Jahedi and Fabio Méndez*
Department of Economics
University of Arkansas

June, 2013

Abstract

This paper utilizes data from a laboratory experiment in order to examine the advantages and disadvantages of subjective measures and their use in empirical analysis. Our results indicate that subjective measures correlate highly with the variables they are designed to capture, but that they also suffer from several systematic biases related to the economic and cognitive contexts. More importantly, our results suggest that subjective measures of broadly defined concepts convey additional information which is complementary to that obtained from facts and that the use of these measures may actually be preferable to the use of purely objective measures in some cases. Ironically, we find that subjective measures may be most useful in exactly those situations where they disagree with their objective counterparts.

*We would like to thank the Sutton Fund and the University of Arkansas for financial support. We also thank Kristine Brown, Cary Deck, Ori Heffetz, Youenn Loheac, and seminar participants at the University of Arkansas for valuable comments. Xirong Chen provided excellent research assistance. All errors are ours. Contact Info: Department of Economics, Sam M. Walton College of Business, Business Building 402, Fayetteville, AR 72701-1201, USA. Email: sjahedi@walton.uark.edu, fmendez@walton.uark.edu.

1 Introduction

Subjective measures are increasingly being used in empirical studies of many economic phenomena for which objective data is difficult to obtain.[1] When measuring corruption, happiness, racism, consumer satisfaction, or sexual behavior, for example, researchers resort to subjective data because the objective variables of interest are either actively being hidden by the parties involved, or are vaguely defined. Subjective measures are often derived from survey questions such as "How satisfied are you with your life?" or "How violent is your city?", which ask respondents to assess the variable of interest. The use of these measures for empirical analysis, however, has been confronted with strong skepticism (as pointed out by Manski 2004) and open calls for the use of purely objective measures instead (Olken 2009; Banerjee, Hanna, and Mullainathan 2012). It is thus, important to ask whether such mistrust is warranted.

Three main arguments against the use of subjective measures can be found in the literature. First, subjective measures have been shown to suffer from many systematic biases related to order, scale and halo-effects (Podsakoff et al. 2003), psychological factors (Bertrand and Mullainathan 2001; Redelmeier et al 2003), macroeconomic fluctuations (Donchev and Ujhelyi 2011), and others. Second, subjective measures have been shown to be uncorrelated (and even negatively correlated) with independent, objective measures related to the variable of interest (Olken 2009; Razafindrakoto and Roubaud 2010; Kaplan and Pathania 2010; Hardoon et al 2003). Third, subjective measures are difficult to aggregate and interpret because they are often expressed in ordinal scales (Rose-Ackerman 1999).

In this paper, we evaluate the validity of these arguments and examine the performance of subjective measures relative to that of objective measures. In doing so, we distinguish between two types of subjective measures: *general* and *specific*. Specific subjective measures are derived from survey questions that ask about well-defined concepts that can be observed in principle such as "the amount of money paid in bribes" or "the number of times you were racially discriminated against." General subjective measures are derived from questions that ask about broad concepts, such as "the level of corruption" or "the extent of racism," which comprise both explicit components that can be observed in principle (e.g. bribes paid by private investors) and implicit components that cannot be observed (e.g. investment projects aborted in order to avoid bribes).

[1] Kahneman and Krueger 2006, for example, note the increased use of perceptions in studies of life-satisfaction. They report that while only 5 studies using subjective measures were published between 1991 and 1995, more than 100 similar studies were published between 2001 and 2005. Others are also taking notice. Starting in 2005, the Gallup World Poll began maintaining a well-being index for 155 countries worldwide while in 2011, the United Kingdom decided to directly include some subjective well-being questions as part of their annual Integrated Household Survey.

This distinction, as the paper goes on to show, is important for comparing the relative performance of subjective and objective measures. Admittedly, when measuring well-defined concepts, the use of objective data, if it exists, is preferable. Specific subjective measures provide, at best, a noisy approximation of the facts. When measuring broadly defined concepts, however, the use of objective data may not always be preferable insofar as the objective data overlooks implicit components relevant to the variable of interest.[2] When measuring racism in the workplace, for example, an objective account of the racist acts that take place might not be as good as a general subjective measure of the overall extent of racism. To see this more clearly, think of an extreme situation in which racism is so strong that it forces all minorities to leave a certain environment. In that case, the total number of racist acts observed would necessarily be equal to zero and would not provide a good measure for the extent of racism that prevails. A general survey question, in contrast, may provide a more accurate representation if the respondents' answers are sensitive to both explicit racism (racist acts that take place) and implicit racism (racist acts that would have taken place if minorities had stayed in that environment).

In order to conduct our study, we set up a lab experiment with crime from which we obtain the corresponding objective and subjective measures necessary for our analysis. For the measurement of well-defined concepts such as the frequency of theft or the amount of money stolen, our results indicate that specific subjective measures correlate well with the objective facts they intend to quantify, but that they also suffer from several systematic biases related to both cognitive problems and economic conditions. In addition, we find that the individuals' assessments of the frequency of theft are polluted by the amount of money taken away from them and vice versa; even when participants had full access to the information necessary to answer questions correctly. These results suggest that the identification of individual parameters through the use of specific survey questions, as advocated by recent literature, might be more problematic than presently thought.

In contrast, our results suggest that the use of general subjective measures might be less problematic than presently thought. In particular, our results indicate that general subjective measures can effectively capture changes in both the explicit and the implicit components of the variable being measured and, therefore, that they can be better suited for the study of broadly defined concepts than objective measures. In fact, in our study, general subjective measures of crime were better correlated to the levels of crime exogenously introduced in the lab than were objective measures such as the total amount of money taken or the number of times a theft took place. At the same time, in accordance with previous studies, our results indicate that general subjective measures are influenced by many of the same biases found

[2]This point is demonstrated in a simple theoretical framework outlined in the Appendix.

in specific subjective measures. These biases include a recency bias (where the outcomes of the last five periods affect answers nearly twice as much as the outcomes of the first five periods); an intensity bias (where streaks of theft influence subjective perceptions more heavily than do dispersed theft acts of the same magnitude); and an income bias (where subjective answers are influenced by the level of payoffs received).

With these results at hand, we then re-evaluate two often-made claims about the perceived disadvantages of general subjective measures: (1) that general subjective measures are not as easy to interpret as objective measures because they are expressed in ordinal scales, and (2) that general subjective measures should be not be trusted because they move opposite to objective measures. We argue that neither of these claims are well justified; even though they are fully consistent with our data. First, we show that the inability to interpret the scale of general subjective measures is a problem that objective measures similarly face. Second, we show that subjective measures might move opposite to objective measures of related phenomena simply because of changes in implicit components that might not be accounted for in the objective measures.

We choose an experimental methodology intentionally because, in some ways, data gathered from laboratory experiments is better suited for this analysis than is naturally occurring data. An important advantage of the laboratory is that it allows one to collect objective data on incidences, free from any measurement error. The perfect account of the environment allows for a dutiful comparison of subjective measures to objective ones. For the case of crime in particular, where field data is prone to error, comparisons such as these are indeed difficult outside the lab. Another important advantage of the lab is that it enables one to observe variables that are typically unobservable. Thus, the laboratory experience allows one to obtain a comprehensive record of the variable one is interested in measuring; which includes both explicit components that can typically be observed (such as the amount of money stolen) as well as implicit components that are not observed so easily in practice but that contain useful information nonetheless (such as any actions taken in order to avoid theft).

In the experiment, subjects were given either the role of "investors" or "observers." Both, investors and observers, were randomly assigned into groups or "environments." In these environments, it was possible (but not necessary) for the observer to take money away from investors. The observer was informed of the specific frequency with which he could take money and the specific amounts he could take; where both of these variables changed exogenously across environments. He was then asked to make a decision on whether to take money or not. After his decision is made, the theft prospects faced by the investors (the amount they can expect to lose to the observer) are set for that environment. These theft prospects are what we refer to as the *exogenously determined crime level* prevalent in the environment. Subsequently, investors were asked to decide whether or not to invest in a series of projects with known expected payoffs. All investors were offered 10 investment opportunities per environment.

After all investments decisions were made, investors were asked to answer survey questions regarding the frequency with which money was taken away from them in that environment, and the amount that was taken. They were also asked to report their general assessments regarding the overall level of crime in the environment. From the investors' answers to these questions, we obtain the *subjective* measures necessary for the study. *Objective* measures of crime, in turn, were recorded during the experiment. Measures of explicit crime components, such as the frequency of theft and the amount of money taken, were collected alongside measures of implicit crime components, such as the number of investments aborted because of crime and the money that would have been taken if these investments were realized.

The paper makes a contribution to a recent literature that examines the validity of subjective measures (Bertrand and Mullainathan 2001; Dominitz and Manski 2004; Kahneman and Krueger 2006; Olken 2009; Barr and Serra 2010, Kaplan and Pathania 2010; Oswald and Wu 2010; Donchev and Ujhelyi 2011). The strategy that these papers have employed to validate subjective measures was to simply compare subjective measures with objective measures and to determine whether the two are correlated. In contrast to these papers, we argue that subjective measures can be useful even if there is little or negative correlation with the objective counterparts. In fact, we find that subjective measures are most useful when they disagree with objective measures. Our insight adds to the validity of subjective measures by showing that their validity can be independent of their relationship to objective measures. Importantly, unlike previous studies, the paper utilizes data gathered in the lab and, thus, it is the only one able to account for variations in exogenously determined levels of the variable being measured and its implicit components.

The paper also makes a contribution to the growing empirical literature that uses specific subjective data to approximate individual parameters such as the frequency of particular events or the amounts of money involved in those events (Seligson 2006; Renikka and Svensson 2006, Foster et al 2012). Specific subjective measures are increasingly welcome by researchers looking to test precise theoretical predictions; but we are not aware of any other study that systematically examines their properties. Finally, the paper makes a contribution to the empirical literatures that utilize general subjective measures as proxies of many broadly defined phenomena; as our results suggest that general subjective measures can effectively serve the empirical purposes they were meant to serve.

2 Experimental Design and Data Description

In this experiment, participants spend ten periods in an environment with an exogenously set level of theft before being placed into a different environment. At the end of each environment, participants are asked four questions regarding the different aspects of theft that were experienced. Participants' answers to these questions are used to derive subjective measures of crime. We then compare subjective measures to both objective measures and to exogenously determined levels of crime to determine their accuracy.

2.1 Methodology

The experiment was conducted at the Behavioral Business Research Lab at the University of Arkansas over 10 sessions and had a total of $N = 170$ participants. At the beginning of each session, all participants sat at separate computers and read the experimental instructions as they appeared on their computers screens.[3] Each player was subsequently assigned the role of "Investor" or "Observer," which they kept for the remainder of the experiment. Participants were asked to practice their role, and were given a rather intensive quiz to make sure that they understood the payoffs associated with each action. Any questions the participants had were answered by the experimenter privately.

Figure 1 describes the game-tree of the experiment. The participants were randomly paired into groups composed of four investors and one observer, and each group was assigned a random draw of three parameters, f, X, and r, which make up an "environment." The probability that an incurred investment has money stolen from it is f; the amount of money that is stolen from it is X; and the rate of return on an Investor's investment is r. In the first stage of each environment, the Observer was informed of the realized f and X of the environment, and was asked an once-and-for-all decision as to whether he would like to take an amount of X ECU[4] from a fraction f of those investors who chose to invest in that environment over the next 10 periods. It was possible (but not necessary) for the observer to take money away from investors.[5] If the observer did not take, he would receive only a flat fee, $F = 150$ ECU, for every period in that environment. If the observer decided to take, his payoff would increase with the number of investors that were approached for money. He was

[3]The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007). The instructions along with screenshots of the interface can be found in the web appendix, at http://comp.uark.edu/~sjahedi/measures/webappendix.pdf

[4]ECU stands for "Experimental Currency Unit" and can be converted to dollars at the end of the experiment at the rate of 100 ECU = $15.

[5]The primary role of the observer is to generate crime. Initially, we wanted a computer to play the role of the Observer. However, it was suggested to us that people's perceptions might differ when they are asked about a computer versus a living human.

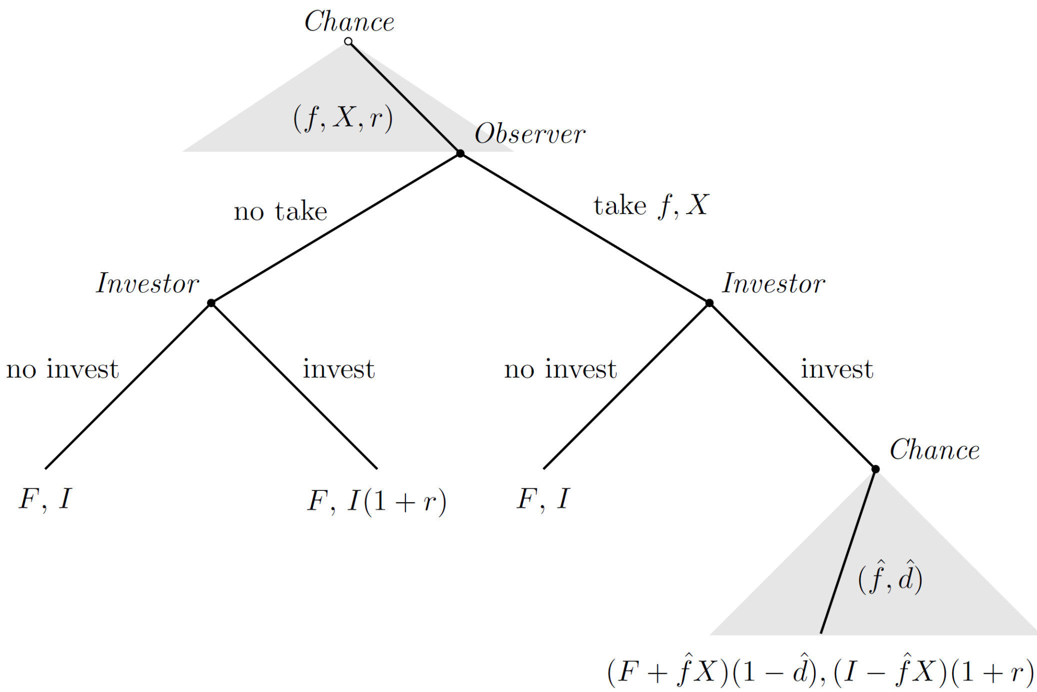


Figure 1: **The Investment Game.** Chance determines the parameters f, X, and r of the investment environment. The observer first decides whether or not to take a known amount X ECU from a probability f of investments (before they occur). After being told the observer's choice, the investor decides whether to invest I =100 ECU investment at a known return, r. Any theft is then realized, and payoffs are recorded.

aware that taking money was associated with a risk of getting caught; for each investor that was approached, there was a fixed probability $d = \frac{1}{100}$ that the observer would be caught and forfeit any money made in that period.[6] The variable $\hat{d}$ in Figure 1 is an indicator variable denoting those situations where the Observer was caught. After the observer chose whether or not to take money from investors, he made no more decisions for the remaining ten periods. Rather, after each period, the observer saw summary statistics regarding how many investors chose to invest and how much money was collected in total.

Once the observers made their decision, the investors had to make a choice of whether or not to invest in each of ten periods, with feedback after every period regarding the outcome of their investment decision. Before making each decision, investors were fully informed about the details of their environment: investors were told the frequency f in which money would be stolen from their investment, the amount of money X that the observer would

[6]We included a chance of getting caught so that investors and observers would both think that taking money was wrong. However, we chose this probability to be small enough that it was worthwhile for the observer to take money. This allowed us to create an environment with exogenous theft. It should be noted that only observers were alerted when they were caught; investor's did not receive feedback as to the number of observers that were caught.

steal, and the return on their investment, r. While f and X were constant across periods, the return varied slightly so that the Investor's task was not strictly repetitive. Investors were also given the expected value from investing and not investing in each period so to help in their decision-making. They subsequently made each of their ten investment decisions in turn. After every period, they were told how much money was taken from them (if any). If the investor chose not to invest, they kept their full endowment of $I = 100$ ECU. If the investor did invest, the computer would use a random draw from a Bernoulli distribution which took a value $\hat{f} \in \{0, 1\}$ with the probabilities $1 - f$ and f to decide whether the investor was approached for money. If the investor was not asked for money, he received his endowment plus the return, $100(1 + r)$. If the investor did invest and he was asked for money, he received a return on his endowment less the money taken, $(100 - X)(1 + r)$.

After the ten investment decisions are made, investors were asked four questions aimed at measuring both their experiences and their perceptions. It is important to note that in our experiment, the investors are pure victims of theft rather than active participants of crime (e.g., paying a bribe) and so the questions that we ask are not muddied by problems associated with the typical under-reporting problems faced in other studies. The first two questions are very general and abstract questions while the remaining two are more specific and verifiable. The questions are, in order:

> Q1. On an absolute scale from 0 to 100, how would you rate the current environment? Please enter an integer value (0 = Best — In this environment, any amount that I invest realizes its full return, with certainty; 100 = Worst — In this environment, any amount that I invest is taken away from me entirely, with certainty).
>
> Q2. Please rank the following statement: The fact that this environment allowed observers to take money away from investors was a serious problem for me. Please enter an integer value (0 = completely disagree; 100 = completely agree).
>
> Q3. How often were you faced with a situation in which money was taken away from you? Please enter an integer value (0 = never; 100 = always).
>
> Q4. How much money was taken away from you? Please enter an integer value (0 = none at all; 100 = all of it).

The formulation associated with the last two questions came about from studying the wording of actual surveys from the corruption literature.[7] Of course, we recognize that these

[7] Some examples include the following: "On a scale of 0 to 10, how serious do you consider the problem of corruption to be in the public sector?" is asked by Political and Economic Risk Consultancy to expatriate

questions may still have some ambiguity in their interpretation, but we decided to stick with them rather than invent our own. To address the potential ambiguity we include controls in our analysis for both the absolute and relative interpretations of these questions. For example, our regressions for Question 3 include the total number of times that individuals had money taken from them as well as the percentage of times that money was taken from them conditional on investing. We find that each matters individually. The formulation associated with the first two questions also heavily relied on the wording from questions in the corruption literature. Nevertheless, these questions are much more general and have more ambiguity than the latter two. In order to reduce some of the ambiguity, we included a scale that defined the endpoints. The scale used for these questions, especially Q1, were meant to capture the total amount of theft (both implicit and explicit) that exist in that environment. All analysis of the general questions used Q1, with robustness checks using Q2 relegated to the Appendix.

Finally, after these questions were answered, all players were again randomly assigned to groups of 4 investors and 1 observer, and again randomly placed into a new environment (defined by new draws of the parameters f, X, and r) where they were confronted with the same set of choices as before. Each session consisted of 8 environments, for a total of 80 periods per session. At the end of 8 environments, participants were paid. One of the 80 rounds was randomly chosen by the computer to determine each person's payoff. On average, participants received $30 for the 90 minute experiment.

2.2 Data description

The data contains the choices of 136 investors who made a total of 10080 investment decisions. Each investor made 10 decisions per environment and participated in a total of 8 environments.[8] An environment was defined by a particular combination of parameter values of f, X, and r. Each environment was randomly assigned a value of f from the set $\{f_1=0.25,\ f_2=0.5,\ f_3=0.75\}$, and a value of X from the set $\{X_1=16,\ X_2=24,\ X_3=48\}$. Given that some observers chose not to take money, however, the theft prospects faced by investors in any given environment also include the case for which $\{f_0,\ X_0\}=\{0,\ 0\}$. Finally, each

business executives. "In your industry, how commonly would you estimate that firms make undocumented extra payments or bribes connected with ... (exports, taxes, etc)?" is asked by World Economic Forum to business leaders and companies. "On average, what percentage of revenues do firms like yours typically pay per annum in unofficial payments to public officials?" is asked by the World Bank's BEEPS and WEEPS surveys to business managers.

[8]Two of the sessions had to be cut short, as some of the computers restarted in the middle of the experiment. For session 6, we gathered data from 8 investors across 4 environments. For session 7, we gathered data from 16 investors across 5 environments. In all other sessions, we had observations for all investors from all 8 environments.

Table 1: **Summary Stats.** There were a total of 30 types of environments that participants could be assigned. The first three columns list them in order of their return level and theft prospects ($f \cdot X$). The remaining columns give the environment average for each variable. Figure 6 in the Appendix shows the relationship of each variable Q1-Q4 with the variables they are intended to measure.

Parameters			**Observations**		**Objective Measures**		**Survey General**		**Survey Specific**	
f	X	r	N	Particip. Rate	Avg. Times \$ Taken	Avg. Amt. \$ Taken	Q1	Q2	Q3	Q4
0	0	[.01, .10]	36	94%	0	0	3.2	16.8	0	0
0.25	16	[.01, .10]	16	39%	0.9	15	43.1	69.1	16.6	15.3
0.25	24	[.01, .10]	24	34%	1	25	47.4	47.5	19.4	18.3
0.5	16	[.01, .10]	24	28%	1.3	20	61	60.2	19.2	14.1
0.25	48	[.01, .10]	28	23%	0.5	24	42.5	57	15.7	15.6
0.75	16	[.01, .10]	28	32%	2.4	38.3	62.7	63	28.9	18.3
0.5	24	[.01, .10]	52	15%	0.8	18.5	67.7	61.3	11.9	13.7
0.5	48	[.01, .10]	24	11%	0.5	24	62.6	78.5	20.8	18.3
0.75	24	[.01, .10]	32	18%	1.3	30.8	70.2	73.1	32.8	13.5
0.75	48	[.01, .10]	28	13%	1.1	54.9	78.8	90.3	31.3	22
0	0	[.11, .25]	44	98%	0	0	4	16.7	0.5	0.5
0.25	16	[.11, .25]	40	84%	2.2	34.8	21.9	43.4	14	14.7
0.25	24	[.11, .25]	52	64%	1.7	39.7	35.1	48.2	25.6	23.3
0.5	16	[.11, .25]	12	68%	3.5	56	35.6	38.8	24.1	22
0.25	48	[.11, .25]	28	45%	1.2	58.3	42.8	55	22.8	20
0.75	16	[.11, .25]	32	74%	5.9	94.5	38.1	50.2	52.4	40.3
0.5	24	[.11, .25]	56	47%	2.5	60	55.5	69	35.1	25.4
0.5	48	[.11, .25]	40	28%	1.5	69.6	59.6	64.6	24.3	27.8
0.75	24	[.11, .25]	56	35%	2.6	61.3	59.9	72.2	43.1	27
0.75	48	[.11, .25]	20	17%	1.4	64.8	72.8	71.1	23.6	22.3
0	0	[.30, .60]	32	98%	0	0	0.9	17.7	3.8	3.8
0.25	16	[.30, .60]	20	97%	1.9	30.4	10.7	38.5	20.6	18.2
0.25	24	[.30, .60]	28	95%	2.3	54	13.7	32.8	18.6	15.8
0.5	16	[.30, .60]	24	96%	5.6	89.3	23.1	30	44.9	27.4
0.25	48	[.30, .60]	40	68%	1.5	73.2	31	53.8	23.8	22.5
0.75	16	[.30, .60]	28	91%	6.9	110.3	45.3	42.8	64.8	39.1
0.5	24	[.30, .60]	56	96%	4.9	118.7	21.2	37.5	36.5	29.1
0.5	48	[.30, .60]	64	58%	2.7	129.8	51.7	68.9	46.6	40.1
0.75	24	[.30, .60]	16	91%	7	168	54.4	63.8	61.8	41.3
0.75	48	[.30, .60]	28	51%	3.4	164.6	66.1	69.9	44.2	40.4

Table 2: **Correlation Table.** A simple correlation table shows the correlation between the exogenously determined crime level (fX), and all six measures considered (two general survey questions, two detailed survey questions, and two objective measures).

	$f \cdot X$	Svy. Q1	Svy. Q2	Svy. Q3	Svy. Q4	Obj. Freq.	Obj. Amt.
$f \cdot X$	1						
General Survey: Q1	0.494	1					
General Survey: Q2	0.369	0.477	1				
Specific Survey: Frequency	0.237	0.246	0.239	1			
Specific Survey: Amount	0.250	0.221	0.263	0.604	1		
Objective: Frequency	0.128	0.056	0.029	0.536	0.487	1	
Objective: Amount	0.386	0.138	0.122	0.525	0.555	0.834	1

environment was randomly assigned a range for r which could be low, medium, or high. In environments with low returns, the value of r was picked randomly from the set [.01, .10] for each investment project; for medium returns, the value of r was picked randomly from the set [.11, .25]; and for high returns, the value of r was picked randomly from the set [.30, .60].

As a result, there were 30 different types of environments possible; these environments are described in the first three columns of Table 1. The fourth column of Table 1 shows the total number of investors that participated in that type of environment (N), while the fifth column shows the fraction of investments that transpired in that environment (Participation Rate). This table also presents summary statistics of the objective and subjective measures of crime in the corresponding environments. As explained before, the actual number of times investors had money taken away and the total amount of money taken constitute our objective measures. In turn, the investors' answers to the sets of questions Q1-Q2 and Q3-Q4 serve as our general and specific subjective measures, respectively.

Table 1 further illustrates two important properties of the experimental design regarding the parameter values f, X, and r. First, investing yields a lower expected payoff than not investing in about half of the investment projects (as a result, subjects in our experiment decided to invest in about 57% of their projects). This property is important because it allows one to differentiate between implicit and explicit components of crime. Noticeably, when investors decide not to invest the explicit components are zero (no money is taken at all), but the implicit displays are non-zero (because crime is acting to discourage investment). Second, there are multiple combinations of f and X that have the same expected payoff: $(\frac{1}{4}, 48)$, $(\frac{1}{2}, 24)$, and $(\frac{3}{4}, 16)$. This property is important because it allows one to examine which dimension of crime, f or X, matters more for the formation of people's perceptions.

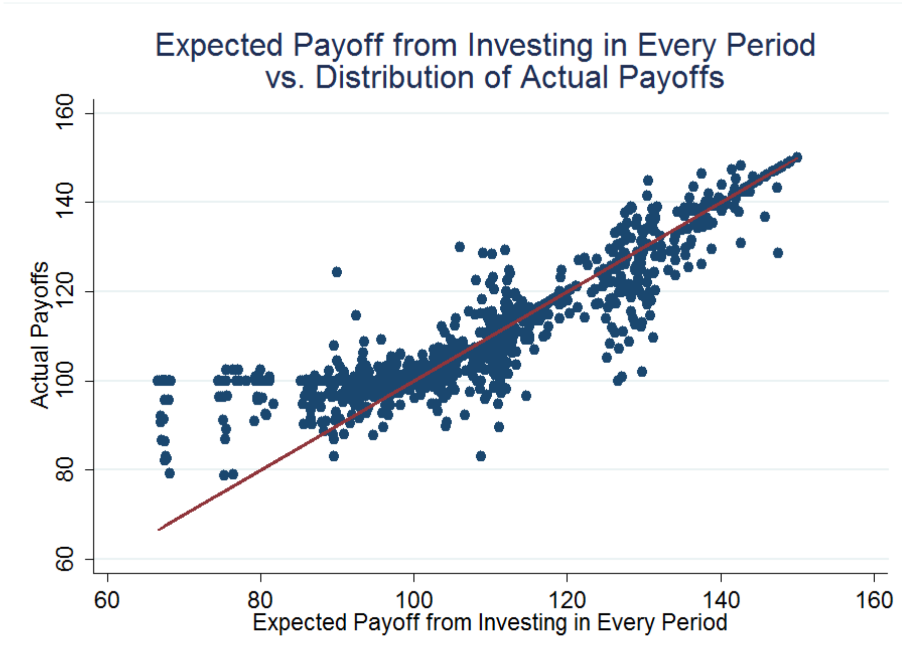


Figure 2: We find that nearly all participants are risk-neutral. When the value from investing is below that of not investing, the majority of people do not invest. When the value from investing is above that of not investing, the vast majority of people invest.

Finally, Table 2 provides a first look at how the exogenously determined levels of crime correlate with the different measures studied here. In our experimental setting the exogenously determined crime level is defined by the product of the two parameters that make up crime in an environment: $f \cdot X$. That is to say, we define the overall level of crime present in any given environment as the theft prospects faced by investors in that environment *before* they make their decisions. As shown in the appendix, however, similar results are obtained if alternative definitions are used.[9] In Table 2, the two general survey questions (Q1 and Q2) seem to perform at least as well, or better than, the specific survey questions or the objective measures. In fact, the average of the two correlation coefficients for the general survey questions (0.43) almost doubles the corresponding average for the specific survey questions (0.25) or the objective measures (0.24). In what follows, we address this matter more closely.

3 Results

First, we note that participants seem to understand the game well and that their investment patterns followed that of risk-neutral investors. This can be seen clearly in Figure 2. The 45

[9] In particular, we consider an alternate definition of crime which we label "Money Lost". For this definition, we add the total money that was taken from those that did invest to the foregone interest that was essentially taken from those that did not invest. The results do not change when this definition is used.

degree line represents the expected value from investing in the period and the dots represent the actual payoffs associated with investor decisions. When the expected value of investing was below the expected value of not investing, most people chose not to invest. However, when the expected value of investing was above the expected value of not investing, most people did invest.

Next, we now turn to evaluate how well subjective measures explain the variables they attempt to quantify. For the specific survey questions Q3-Q4, we compare responses to the survey questions with what actually happened. In turn, for the general survey questions Q1-Q2, we compare responses to the exogenously determined levels of crime, as defined before. For simplicity, the definition of crime in the experiment is reduced to theft (money taken away from investors). One may argue that this definition also suits concepts such as "the level of corruption", "the institutional quality", or the "business climate" observed in each particular environment. We agree that all these concepts are related and we make no attempt to disentangle them. Our contention is that, for the purposes of this paper, the specific title attached to the product $f \cdot X$ is inconsequential. What matters is that it constitutes a well defined target that allows one to judge the performance of different measures.

3.1 The accuracy of specific subjective measures.

We begin our analysis by studying how well specific survey measures correspond to the verifiable facts they are intended to portray. We asked participants one question about the frequency with which they were approached (Q3: How often were you faced with a situation in which money was taken away from you? (0 = never, 100 = always)) and one question about the amount of money that was taken from them (Q4: How much money was taken away from you? (0= none, 100= all)). As both of these questions are direct inquiries about the explicit theft acts, we can check whether people answer these questions correctly. It should be noted that before answering these questions, participants were reminded of the environmental factors f and X, they were shown a history box containing the number of times they were approached and the amount of money that was taken from them in the past 10 periods, and they were informed of the total percentage of theft acts that occurred to all investors in the environment.

We conduct OLS regressions for each question separately. The results are presented in Table 3 and Table 4, respectively. As the dependent variables we use the participants' numerical response to each survey question (Q3 in Table 3; Q4 in Table 4). The explanatory variables will depend on the question but generally include the factual measures that the survey question is asking about and also variables that can capture the biases in people's answers. In exploring biases, we focus on (1) factual measures that are unrelated to the question participants are being asked, (2) outcomes from past environments, (3) distant vs. recent events, and (4) the intensity of events. Remarkably, the conclusions are quite comparable across questions. In all of the estimations, the errors are clustered at the participant level as participant's responses may be correlated across environments.

The main explanatory variables used in Table 3 include the absolute number of times that the individual was approached in that environment, the relative number of times that he was approached among the times that an investment was made (*Percent of Times Approached*),[10] a dummy variable equal to 1 if the investor never invested in that environment, the size of the theft X, the difference between the actual payoff and the expected payoff from the previous environment (*Lagged(payoff-Epayoff)*),[11] and the total number of times in a row that an investor was either approached or did not invest (*Streaks*). Correspondingly, the list of explanatory variables in Table 4 includes the absolute amount of money taken from the individual in that environment, the relative amount that would be taken from investors if their theft experience from investment periods was extended to all ten periods (*Extrapolated amount*),[12] the frequency of theft f, and some of the same variables previously defined for Table 3.

Result 1. *Subjective measures of specific, well-defined concepts correlate well with the facts they intend to quantify.*

Table 3 looks at the variables that influence the survey response to Q3 regarding the frequency with which people were approached in an environment. Column 1 and Column 2 of the table show that the reported frequency of being approached for money is strongly correlated with the absolute number of times that the individual was actually approached. Those who were approached more often reported higher numerical responses to Q3. In addition to the absolute number of times approached, the answer to Q3 is increasing in the percentage of times an individual was approached. Thus, a person that was approached twice out of three investments (because they did not invest all of the time) reported a higher answer to Q3 than a person who was approached twice out of ten investments.

In turn, Table 4 looks at the variables that influence the survey response to Q4 regarding the total amount of money that was taken from them within an environment. Column 1

[10] A person who invested in only 3 periods and was approached twice would have *Total Number of Times Approached* =2 and *Percent of Times Approached* = $\frac{2}{3}$.

[11] We checked to see if other lagged variables affected responses. In general, this did not seem to be the case.

[12] The Extrapolated Money Taken is the amount that would be taken from investors if their theft experience was extended to all ten periods. It is calculated as $\frac{\text{Actual Money Taken}}{\text{Fraction of Investments Made}}$. For example, if an investor who chose to invest in only half of the periods had \$15 taken, then the extrapolated money taken would be $\frac{\$15}{0.5} = \$30$.

and Column 2 of the table show that the money reported to be taken is strongly correlated with the actual money taken from investors. Those who had more money taken from them reported higher answers to Q4. In addition to the total amount of money taken, the answer to Q4 is increasing in the extrapolated amount of money taken. That is, a person who had $100 taken away from him when they invested only three times reports a much higher response than a person who had $100 taken away from him and who invested every single period.

Result 2. *Subjective measures of specific, well-defined concepts are influenced by non-relevant information and suffer from systematic biases.*

While survey questions correlate heavily with facts, they also suffer from many different biases. Columns 2-6 of Table 3 show how five separate biases individually affect how people answered the frequency by which they were approached. As shown in Column 2, those participants who made zero investments in a given environment reported much higher rates of being approached than investors who made investments and were never approached. Column 3 shows that the bigger is the amount observers are allowed to take in an environment (X), the more people report the *frequency* with which they were approached to be. Column 4 shows that lagged variables from previous environments have an affect on people's responses in future environments. Namely, if a person did better than expected in the previous environment, they are less likely to report being approached in their current environment. Column 5 shows that the intensity of their worst bad streak, the total number of times in a row that the person was either approached or did not invest,[13] factors positively in their report of being approached and Column 6 shows that the number of times approached in the first half are weighted less heavily than the number of times approached in the second half. Column 7 includes all variables in the regression together, and it is apparent that all biases (except streaks) remain significant, even after controlling for the other biases.

The same exercise is repeated for Question 4 regarding the total money taken from the individual. Columns 3-6 of Table 4 show how four separate biases individually affect how people answered the frequency by which they were approached. Column 3 shows that the bigger the frequency of theft in the environment, the more people report the amount of money taken from them to be. Column 4 shows that the lagged variables from previous environments matter in that the luckier you were in previous environments, the less money you report to be taken in the current environment. Column 5 shows that the intensity with which people were approached matters in addition to the total money taken from them, and

[13] We defined a bad outcome as a situation where the Investor was either approached for money OR chose not to invest. The results are robust to alternative definitions such as times being approached only.

Table 3: **Biases in Frequency.** At the end of every environment, participants were asked "How often were you faced with a situation in which money was taken away from you? Please enter an integer value (0 = never; 100 = always)." The answer to this question was used as the dependent variable and regressed against a number of biases.

Dependent Variable = Numerical Response to Frequency Question, Q3 Standard Errors Clustered by Participant							
	model 1 b/se	model 2 b/se	model 3 b/se	model 4 b/se	model 5 b/se	model 6 b/se	model 7 b/se
Total Number of Times Approached	7.849***	5.547***	4.745***	7.560***	7.868***		
	[0.389]	[0.823]	[0.8437]	[0.5370]	[0.5090]		
Number of Times Approached (first 5)						5.604***	3.073**
						[0.9535]	[1.266]
Number of Times Approached (last 5)						10.089***	8.380***
						[0.9128]	[1.285]
Percent of Times Approached		26.39***					17.140**
		[6.153]					[8.536]
Dummy (never invested in 10 periods)		8.836 **					0.1422
		[3.849]					[5.806]
Size of Theft (X)			0.135***				0.2087***
			[0.0261]				[.0622]
Lagged(payoff - Epayoff)				-0.204**			-0.1798**
				[0.0908]			[.0874]
Streaks					1.567***		0.7681
					[0.3446]		[.5970]
constant	10.396***	4.358***	1.466	10.778***	2.410*	10.405***	-0.787
	[1.734]	[1.217]	[1.524]	[1.891]	[1.314]	[1.732]	[1.455]
R^2	0.2877	0.3206	0.3110	0.2844	0.3144	0.2935	0.3439
Observations	1008	1008	1008	872	1008	1008	872

* significance at 10%, ** significance at 5%, *** significance at 1%

Table 4: **Biases in Money Taken.** At the end of every environment, participants were asked "How much money was taken away from you? Please enter an integer value (0 = none at all; 100 = all of it)." The answer to this question was used as the dependent variable and regressed against a number of biases.

Dependent Variable = Numerical Response to Money Taken Question, Q4
Standard Errors Clustered by Participant

	model1	model2	model3	model4	model5	model6	model7
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Total Amount of Money Taken	0.229***	0.151***	0.215***	0.223***	0.229***		
	[0.015]	[0.022]	[0.015]	[0.016]	[0.015]		
Total Money Taken (first 5 periods)						0.152***	0.058*
						[0.025]	[0.035]
Total Money Taken (last 5 periods)						0.299***	0.218***
						[0.028]	[0.036]
Extrapolated Amount Taken		0.051***					0.031*
		[0.011]					[0.017]
Dummy (never invested in 10 periods)		-5.601***					-13.962***
		[1.525]					[4.740]
Frequency of Theft (f)			9.416***				7.483
			[2.992]				[4.988]
Lagged (payoff-Epayoff)				-0.088			-0.086
				[0.083]			[0.078]
Streaks					0.559***		0.810
					[0.191]		[0.588]
constant	8.515***	8.473***	5.141***	8.502***	5.669***	8.664***	4.910***
	[0.984]	[1.061]	[1.121]	[1.072]	[1.119]	[0.991]	[1.148]
R^2	0.308	0.342	0.314	0.303	0.313	0.317	0.363
Observations	1008	1008	1008	872	1008	1008	872

* significance at 10%, ** significance at 5%, *** significance at 1%

Column 6 shows that the amount of money taken in the first half is weighted differently from the amount of money taken in the last half. Column 7 includes all variables in the regression together. Here, the signs of the estimated coefficients remain the same, but some of them become statistically insignificant.

3.2 The meaning of general subjective measures

We now extend our analysis to see how well subjects' responses to general survey questions correspond to the crime level that was exogenously introduced in the lab. We define the overall crime level as fX, or the total amount of money that would be extracted from the environment if all investors were to invest. We will primarily concentrate our analysis to responses of Q1, as we tailored this question to reflect our definition of crime more closely. All of the results also hold when Q2 is used as the dependent variable instead. These results are relegated to the Appendix.

We conduct an OLS regression with the responses for Q1 serving as the dependent variable. We try to keep our analysis as similar as possible to the earlier tables; we include an explanatory variable that is meant to capture what the question is asking and then include variables that are meant to capture a number of biases. To capture total crime, we include f and X separately at first, and then as a product fX.[14] In some cases, the crime level was further split into *explicit* displays and *implicit* displays of crime. Explicit displays are measured by the total amount of money taken from those projects an investor decided to pursue in any given environment, and implicit displays are measured by the amount of money that would have been taken from projects that the investor decided not to pursue.[15] As explained before, implicit displays of theft are practically impossible to observe in practice, but they are easily observed in the lab. In addition, several other explanatory variables were included in order to capture possible biases in people's answers. These variables include the *Returns* (r) that characterized the environment and the maximum *Streak* of bad outcomes that occur in a row. Finally, the recency of crime was also studied by splitting the amounts of explicit and implicit theft that occurred in the first 5 periods of the environment from those that occurred in the last 5 periods.

The results are presented in Table 5. In all of the estimations, the errors are clustered at the participant level as participant's responses may be correlated across environments. We first regressed the subjects' answers to question Q1 on the exogenously determined crime level (fX) and obtain that responses are positively and significantly correlated with fX. As

[14] The appendix also contains additional regression tables for when the alternate definition of crime "Money Lost" is used instead.

[15] On average, explicit theft and implicit theft sum to a multiple of fX.

shown in column 2 of Table 5, a one unit increase in the value of fX is associated with a 1.6-unit increase in the level reported. When f and X are entered as separate explanatory variables (column 1), we find that each variable significantly influences perceptions; but it seems that responses are more sensitive to changes in the frequency than they are to changes in the amount of money taken. This difference is statistically significant at the 1% level. For example, at the mean value of $f \cdot X = 12$, a 1% increase in frequency raises responses by 0.30 points whereas a 1% increase in money taken raises responses by 0.12 points. We further illustrate this point in Figure 7 in the Appendix by looking exclusively at environments with $f \cdot X = 12$. As shown in this figure, crime is rated higher in environments with higher frequency of theft.

Result 3. *General subjective measures capture both explicit and implicit acts.*

One of the most attractive features of general subjective measures is that they can presumably capture some of the implicit events that objective measures alone cannot. Our results offer evidence of this. Column 3 of Table 5 shows the results of the estimation when the explicit and implicit components of crime are entered separately. These results indicate that general subjective measures are positively and significantly associated with both. In fact, participants' responses react significantly more to implicit displays of crime than they do to explicit displays.

Result 4. *General subjective measures are influenced by non-relevant information and suffer from systematic biases.*

At the same time, in line with other previous studies (Donchev and Ujhelyi (2010), Olken (2009), Razafindrakoto and Roubaud (2010)), our results indicate that general subjective measures suffer from systematic biases. We identify three such biases in our data. The first bias is related to the return levels of the environment. As shown in column 5 of Table 5, given the same levels of crime, investors that received a higher return on their projects reported significantly lower crime levels. The effect of the returns is significant at the 1% level and is consistent with the findings of both Donchev and Ulijev (2011) and Seligson (2006) in relation to corruption measures. Both of these papers find that higher economic performance is associated with lower perceptions of corruption.

The other two biases we find are related to how people recall their memories. It is known that people have two major biases in memory recall: (1) they tend to weight recent events more heavily than past events, and (2) they tend to remember extreme events more than average ones. As shown in column 4 of Table 5, people weight the explicit and implicit components of crime that take place in the last five periods of the environment almost twice

Table 5: **Biases in General Subjective Measures.** At the end of every environment, participants were asked "On an absolute scale from 0 to 100, how would you rate this current environment? Please enter an integer value (0 = Best – In this environment, any amount that I invest realizes its full return, with certainty; 100 = Worst – In this environment, any amount that I invest is taken away from me entirely, with certainty)." The answer to this question was used as the dependent variable and regressed against a number of biases.

Dependent Variable = Numerical Response to General Subjective Question (Q1)
Standard Errors Clustered by Participant

	model 1	model 2	model 3	model 4	model 5	model 6	model 7
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Frequency of Theft, (f)	59.654***						
	[4.334]						
Size of Theft, (X)	0.492***						
	[0.066]						
$f \cdot X$		1.638***					
		[0.089]					
Explicit Theft			1.163***		1.756***	0.873***	
			[0.166]		[0.181]	[0.167]	
Implicit Theft			1.827***		1.628***	0.522***	
			[0.101]		[0.099]	[0.131]	
Explicit (first half)				0.601*			0.553
				[0.359]			[0.363]
Explicit (last half)				1.636***			1.676***
				[0.303]			[0.305]
Implicit (first half)				1.353***			0.114
				[0.277]			[0.281]
Implicit (last half)				2.279***			0.976***
				[0.282]			[0.291]
Return					-0.559***		-0.245***
					[0.061]		[0.059]
Streaks						5.076***	4.608***
						[0.367]	[0.395]
Constant	2.881*	20.690***	22.081***	22.329***	33.194***	7.701***	14.137***
	[1.631]	[1.686]	[1.752]	[1.749]	[2.194]	[1.772]	[2.357]
R^2	0.299	0.256	0.268	0.273	0.324	0.422	0.437
Observations	1008	1008	1008	1008	1008	1008	1008

* significance at 10%, ** significance at 5%, *** significance at 1%

as much as they weight similar events that took place in the first five periods. This difference is significant at the 1% level. At the same time, column 6 shows that people who experience longer streaks of bad outcomes report higher levels of perceived crime. This difference is also significant at the 1% level.

The last column of Table 5 provides our results when all variables are included together. It can be seen that both explicit and implicit components of crime influence perceptions significantly, with later events in an environment being weighted more heavily. Returns affect perceptions negatively, reflecting the fact that people who are richer mind the same level of crime less. Finally the intensity of bad outcomes, as measured by *Streaks*, has a strong positive effect on perceived crime. For each unit that an unlucky streak grows by, the environmental level of crime is perceived to be 4.6 units more.

4 Applications

In this final section, we take a closer look at two major arguments frequently made against the use of general subjective measures. The first argument is that subjective measures are disadvantaged relative to objective measures because their ordinal nature makes it difficult to understand changes in the variable of interest. The second argument used is simply that general subjective measures have been shown to move in opposite directions to the factual observations made about the variable of interest. We address each of these arguments in turn.

Result 5. *Neither objective nor subjective measures are necessarily correlated with the overall crime levels.*

Because of their ordinal scale, responses to general survey questions are incapable of quantifying the absolute levels of crime experienced. At best, they may be correlated with crime. In our experiment, responses to the general survey question (Q1) ranges from 15 in environments with no crime to 70 in environments with a lot of crime. But these numbers convey no direct meaning; so that a change from 30 to 35 cannot be compared to a similar change from 50 to 55, or to any other similar change for that matter.

For environments with implicit displays of crime, however, the same drawback also holds true for observed objective measures. That is to say, that even though objective measures are expressed in cardinal scales, they are nevertheless imperfect indicators of the crime levels exogenously introduced in the lab. The reason why this is true is simple: observed objective measures account for explicit displays of theft only; and explicit displays are not necessarily correlated with the overall crime levels. In fact, an increase (decrease) in objective measures

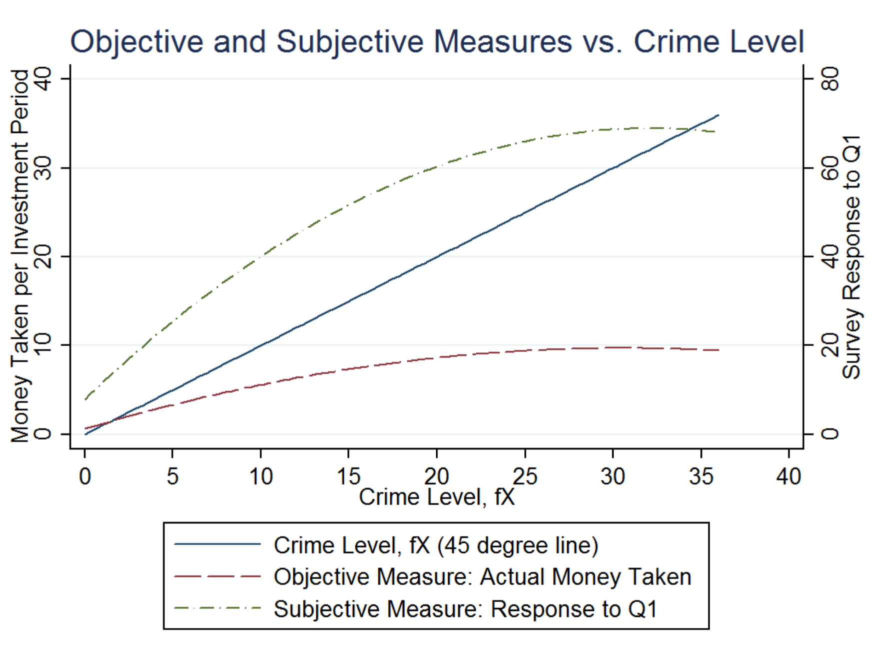


Figure 3: **Crime levels and their correlation with objective and subjective measures.**

such as the total amount of money taken, is perfectly compatible with either an increase or a decrease in the levels of crime that prevailed in any given environment.

To see this point more clearly, refer back to the experiment and imagine an extreme situation in which crime is so pervasive that no one is willing to invest (and thus no money is taken away from investors). Consider what happens when the crime levels *decrease*. Some investors will find it worthwhile to invest and we will find that objective measures will in fact *increase* because some positive amount of money is now taken from those investors. It is clear then that we cannot use the amount of money taken to infer the change in the crime levels. Like general survey questions, objective measures are at best correlated with crime.

The specific correlations between the exogenously determined crime levels and subjective/objective measures for the environments in our experiment are illustrated in Figure 3. In this graph, we arrange our environmental treatments in order of overall crime level, from the lowest crime level ($f \cdot X = 0$) to the highest ($f \cdot X = 36$). General subjective measures are obtained from the mean individual response to Question 1. Objective measures are obtained by computing the average money taken from investors in each period. Each measure is plotted against the exogenous crime level using a quadratic prediction plot. The measures are shown on the same graph using two different scales so to make it easier to see how they trend together. From Figure 3 it is apparent that while both measures seem to correlate positively with the exogenously determined crime levels, neither one tracks its changes exactly.

Result 6. *A negative correlation between objective and subjective measures is not a valid argument against the use of subjective data.*

Previous studies have pointed to the negative correlation that sometimes exists between subjective and objective measures as evidence that subjective measures are wrong or misleading. Olken (2009), for example, cites a negative correlation between the amount of money stolen from a road building project and the perceived amount of corruption involved in that project (derived from a survey questionaire) as evidence that perceptions are misguided or uninformed. As a result, he concludes that perceptions should be used "with considerable caution" when conducting empirical research and that "there is little alternative to continuing to collect more objective measures of corruption, difficult as though that may be." Seligson (2006) expresses a similar conclusion and writes that one should be "careful when estimating corruption on the basis of perception rather than experience since the two may not have a close fit".

In contrast, our findings suggest that the existence of a negative correlation between survey responses and observed data is not sufficient to disqualify subjective measures. The reason why these two measures are uncorrelated may be that some other element independent of the variable being measured (like the returns r in our experiment) affects them both in opposite ways. And as a result, subjective and objective measures can sometimes move counter to each other. Figure 4 shows this is indeed the case in our experimental data. This figure graphs subjective and objective measures of crime, against returns r. By design, the exogenous crime level $f \cdot X$ does not depend on returns r. That crime level is between 12 and 14 regardless of the size of returns. The other two measures, however, each vary significantly with returns. The total money extracted per investment period steadily increases from zero to about ten; while responses to Q1 steadily decrease from 60 to about 30. Thus, as shown in Figure 4, survey responses and available objective measures move in completely opposite directions, but none can be said to measure crime better than the other.

On the one hand, subjective measures wrongly indicate a decrease in crime as returns increase. The reason for this, which was demonstrated earlier, is that perceptions are biased in earnings. For a fixed level of crime, the more money a participant has earned, the lower they report crime to be. As returns increase, more and more people find it profitable to invest in their projects. Higher returns means higher earnings, and higher earnings lead to lower perceived crime. On the other hand, objective measures wrongly indicate an increase in crime as returns increase. The reason for this is that explicit displays are more prevalent in environments with higher returns. For a fixed level of crime, as returns increase and more

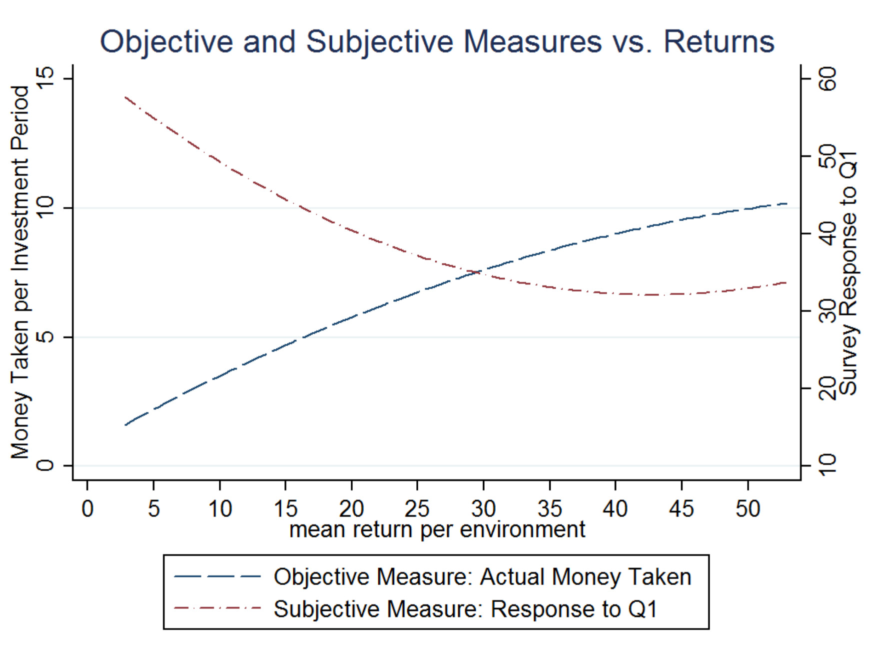


Figure 4: **How objective and subjective measures can be negatively correlated.**

people find it profitable to invest in their projects, more money is taken away from them. But this increase is not justified by an increase in the exogenously determined crime levels that characterize the environments (the level $f \cdot X$ is independent of r); it simply marks an increase in the number of explicit displays observed relative to the number of implicit displays.

5 Conclusion

A spate of recent literature has advocated for a movement away from the use of subjective measures towards objective ones. These studies rely on the observation that subjective measures (1) are known to contain many cognitive and economic biases, (2) are difficult to interpret because of their ordinal scale, and (3) are uncorrelated with facts from the field. We find all these observations are true in our data, but we argue that subjective data can nonetheless be informative. We use an experimental design to examine this issue, which has certain advantages from examining this in the field. First, because we can exogenously affect the crime levels, we can compare subjective measures across treatments without worrying about endogeneity issues. Second, our ability to monitor the environment gives us access to a perfect measure for every incidence of crime, including crimes that did not take place (but would have if an investment was made). This allows us to compare how people responded to survey questions with what actually happened in a more accurate way.

Our findings are intuitive. When objective measures exist for the variable of interest and are free of error, then subjective measures will clearly be worse. However, when perfect information about objective measures for the variable of interest are not available, then subjective measures can perform better because of their ability to measure unobservables. In our experiment, general subjective measures of crime effectively captured both explicit and implicit events, where objective measures were only able to capture the former.

A major implication of our study is that subjective measures can be valuable, independent of their correlation to objective measures. Many other studies have tried to validate subjective measures by showing that they match up well with objective measures. While this is generally a good way to validate subjective measures, we show that subjective measures may be valuable, independent of this relationship. Somewhat ironically, we find that subjective measures are in fact most useful in those situations where they disagree with their objective counterparts. When subjective measures mismatch objective ones, it is likely that one of the measures is not picking up an important aspect of the environment. It would be unwise for researchers to discard either measure beforehand.

Our choice to study crime came partly because of the heavy debate in the field regarding the usefulness of objective and subjective measures and partly because crime is relatively easy to quantify, measure, and compare. For many subjective measures, such as well-being, a true measure is hard to define and nearly impossible to validate. While the study is about crime, we believe that our results generalize to other areas including discrimination, health, worker productivity, well-being, domestic violence, and many other variables. We believe, for example, that asking people about their perceptions of racism can shed light on discrimination that cannot be observed with objective data just as we believe that subjective values of health can inform doctors of issues that their instruments are incapable of finding. The trend away from general subjective measures towards specific measures seems unwarranted. As we have shown, the correct measure will depend on what the researcher is trying to understand. If one is trying to study a broadly-defined concept, such as overall crime, then even perfectly collected field data on every single crime that took place may still be worse than subjective questionnaires.

Certainly, it is possible that our conclusions do not generalize, even outside the lab, where many real-world forces might reduce the accuracy of subjective measures. Survey respondents in the real world do not have the accurate feedback about the level of crime as they do in our experiment. While this is certainly true, we are reluctant to conclude that this makes objective measures more reliable in practice. Just as subjective measures were unrealistically collected in our experiment, so too were objective measures. It is unlikely, for instance, that incidences of crime in the real world can be collected as perfectly as we have done in the lab. Ultimately, data from future empirical studies should inform us as to which is the more reliable measure. Our hope is that future research uses our knowledge of the strengths and weaknesses of subjective data in conjunction with that of objective measures to create more complete overall measures.

References

[1] Banerjee, A., R. Hanna, and S. Mullainathan. “Corruption” NBER Working Paper 17968, April 2012.

[2] Barr, Abigail and Serra, Danila (2010). “Corruption and Culture: An experimental Analysis,” Journal of Public Economics, vol. 94, no. 11-12, pp.862-869.

[3] Bertrand, M. and Mullainathan, S. (2001). “Do People Mean What They Say? Implications for Subjective Survey Data,” *American Economic Review*, vol. 91(2), p. 67-72.

[4] Dominitz, Jeff, and Charles F. Manski. (2004). “How Should We Measure Consumer Confidence?” *Journal of Economic Perspectives*, 18(2): 5166.

[5] Donchev, D., and Ujhelyi, G. (2011). “Do Corruption Indices Measure Corruption?” Working Paper.

[6] Fischbacher, U. (2007). “z-Tree: Zurich Toolbox for Ready-made Economic Experiments” *Experimental Economics*, 10(2), 171-178.

[7] Foster, James E., Horowitz, A., and Mendez, F. (2012). “An Axiomatic Approach to the Measurement of Corruption: Theory and Applications” *World Bank Economic Review*, vol. 26, no. 2, pp. 217-235

[8] Hardoon, K., Derevensky, J., and Gupta R. (2003). “Empirical measures vs. perceived gambling severity among youth. Why adolescent problem gamblers fail to seek treatment.” *Addictive Behaviors*; 28:933-46.

[9] Kahneman, D., and Krueger, A. (2006). “Developments in the Measurement of Subjective Well-Being.” *Journal of Economic Perspectives* 20, no. 1: 3–24

[10] Kaplan D., and Pathania V. (2010). “What influences firms’ perceptions?” *Journal of Comparative Economics*, 38 (4), pp. 419-431.

[11] Manski, Charles. F. (2004), “Measuring Expectations.” *Econometrica*, 72: 13291376.

[12] Olken, B. (2009). “Corruption Perceptions vs. Corruption Reality.” *Journal of Public Economics*, vol. 93, no. 7-8, pp. 950-964.

[13] Oswald, A., and Wu, S. (2010). “Objective confirmation of subjective measures of human well-being: Evidence from the U.S.A.” *Science* 327:576–579.

[14] Podsakoff, Philip M.; MacKenzie, Scott B.; Lee, Jeong-Yeon; Podsakoff, Nathan P. (2003) "Common method biases in behavioral research: A critical review of the literature and recommended remedies." *Journal of Applied Psychology*, Vol 88(5), October, pp. 879-903.

[15] Razafindrakoto, M., and Roubaud, F. (2010). "Are International Databases on Corruption Reliable? A Comparison of Expert Opinion Surveys and Household Surveys in Sub-Saharan Africa," *World Development* 38 (8): 1057-1069.

[16] Redelmeier D.A., Katz, J., & Kahneman, D. (2003). "Memories of colonoscopy: A randomized trial." Pain, 104, 187-194.

[17] Reinikka, R. and Svensson, J. (2006). "Using Micro-Surveys to Measure and Explain Corruption" *World Development*, Vol. 34, No. 2, pp. 359-370.

[18] Rose-Ackerman, Susan, "Corruption and Government." Cambridge University Press, UK (1999).

[19] Seligson, M. A. (2006). "The Measurement and Impact of Corruption Victimization: Survey Evidence from Latin America." *World Development*, Vol. 34, No. 2, pp. 381-404.

A Theoretical Framework

We employ a very simple model where investments are conducted in an environment where money can be taken away. We use this model to outline situations in which objective measures may inaccurately reflect the full extent of crime in that environment.

Consider a population of investors, each of whom start with an endowment, I. For simplicity, assume that all investors have identical, risk-neutral preferences represented by: $U_i(m) = m$,[16] where i denotes the investor and m denotes their income. Each investor must choose whether or not to invest in a project with a known return, r_i, where the returns differ across investors and are randomly selected from the distribution $g(R)$. Without loss of generality, assume that the distribution of returns is uniform with density N on the interval $[a, b]$ where $a \geq 0$. We introduce crime into the environment through the use of two parameters, f and X. Let f denote the probability that an incurred investment has money stolen from it and let X denote the amount which is stolen.[17]

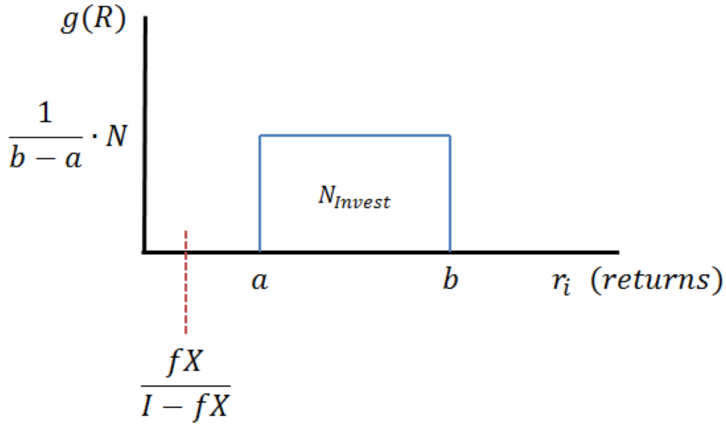
In the case where there is no crime, all investors will find it worthwhile to invest and they will each receive a payoff of $U_i = I(1 + r_i)$ respectively. In the case where theft is prevalent, the expected payoff from investing falls to $U_i = (1-f) \cdot [I(1 + r_i)] + f \cdot [(I - X)(1 + r_i)]$ and only those with high enough returns will find it worthwhile to invest. Namely, investment occurs so long as $U^{Invest} \geq U^{NoInvest}$, or when $r_i \geq \frac{fX}{I - fX}$. In the analysis that follows, we assume that all information regarding the investments that take place are directly observable whereas no information is available on the investments that do not take place.

Let's examine how well observable facts can explain crime in this environment by examining three cases. Consider first the case where average theft fX is small relative to the returns in the environment as in Figure 5a. An example of such small scale theft might be something similar to the fees that banks charge for managing an investor's stock portfolio. As the fee is small relative to the expected return from investment, everybody in society will find it worthwhile to invest. Since full information is known for all incurred investments, N_{Invest}, observable facts can be used to accurately portray parameters in the environment, such as the frequency of theft, as well as more complex measures such as the total level of crime faced by all investors, $\sum_i^N fX$.

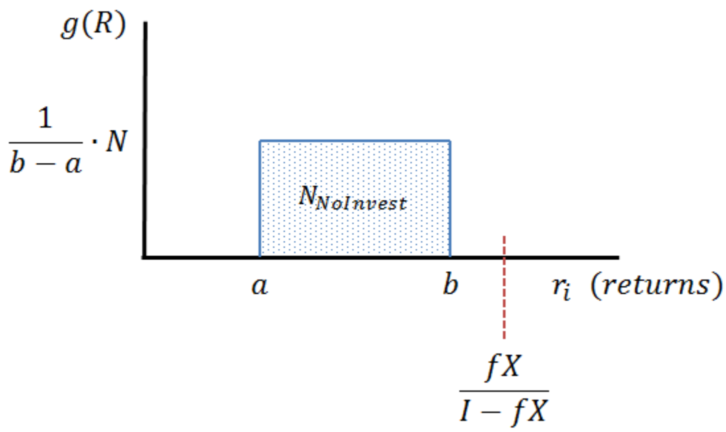
Consider next the case where average crime fX is high relative to returns as depicted in Figure 5b. An example of such a scenario might be the exorbitant fees that some non-profit

[16]None of the conclusions of this section rely on the assumption of risk-neutrality. If agents are risk-averse, they will invest at a slightly higher threshold return level (which will depend on the curvature of the utility function) than that found in the risk-neutral case.

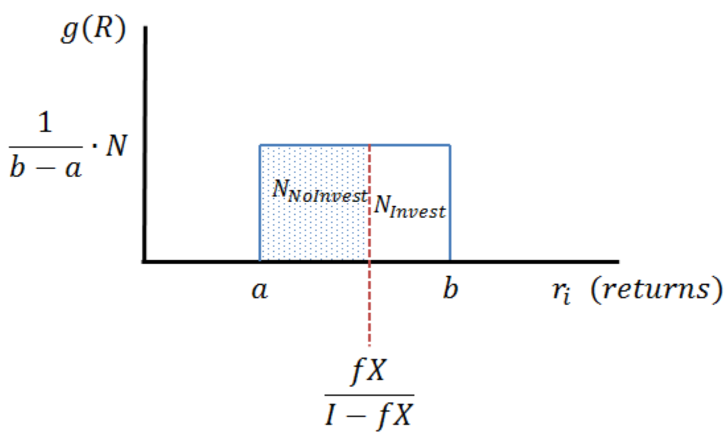
[17]In this analysis, we assume that theft occurs prior to investment and the investor makes returns on the remaining balance. We think of this type of theft as bribes that have to be paid. Alternatively, we could have modeled theft to happen post investment. This would not affect our results.



(a) **Full Investment**



(b) **No Investment**



(c) **Partial Investment**

Figure 5: **Selection into Investment.** The decision to invest will depend on how high returns are relative to the average level of crime. If returns are high relative to theft as in Figure 5a, objective measures will be good. If returns are low relative to theft, then objective measures will be bad.

organizations charge for providing aid to the poor. If the fee is large enough, individuals who would otherwise invest in helping the poor will decide not to invest at all. In a situation where absolutely no investments take place ($N = N_{NoInvest}$), objective measures will simply fail to be useful. Straightforward objective measures, such as the frequency that fees are collected, cannot be computed; obviously, more complex measures are impossible to construct as the required components do not exist. Clearly, with such a strong selection problem, objective measures are at their worst.

Finally consider the situation where there is a moderate level of crime as in Figure 5c. In this case, only those investors who have a return above a certain threshold will invest. The number of people that invest, $N_{Invest} = \left(b - \frac{fX}{I - fX}\right) \cdot \frac{N}{b-a}$, will depend on the overall level of crime, fX, where higher levels of crime will lead to less overall investment. When the goal of the researcher is to measure a variable that is orthogonal to the selection process, objective measures will be accurate. For example, certain environmental parameters such as f and X can be estimated even with truncated data: $f \approx \frac{\text{number approached}}{\text{total investments}}$ and $X \approx$ amount of money taken. On the other hand, when the goal of the researcher is to measure a variable that is being selected into, objective measures will clearly be biased. For example, calculating the average return level for all investors will be upwards biased since only those N_{Invest} individuals with the highest returns will find it worthwhile to invest.

Objective data, even if it can be gathered perfectly with no measurement error, will often be inaccurate even if the tiniest bit of information is missing. In the examples above, the number of people that do not invest is crucial for measuring many important aspects of crime. A researcher that is interested in measuring the expected level of theft felt by all investors, as measured by the sum of money taken if all investors were to participate would be unable to calculate this. With no information on who does not invest, $N_{NoInvest}$, objective measures will fall short in estimating this measure insofar as they cannot shed light on how large the investment market is.

B Additional Tables

This section includes additional figures and tables that are not in the text. Figure 6 breaks down Table 1 into a graphical format. Figure 7 looks only at those environments where the average crime level, fX, was constant at 12. Among those environments, people respond that higher frequency environments are more problematic for investments.

The remaining tables are regressions run for robustness-check purposes. Table 6 is a replication of Table 5 in the text with Q2 used as the dependent variable rather than Q1. Tables 7 and 8 is a replication of Table 5 and Table 6 with a different definition of average crime. Instead of fX, the actual money that is lost to investors is used.

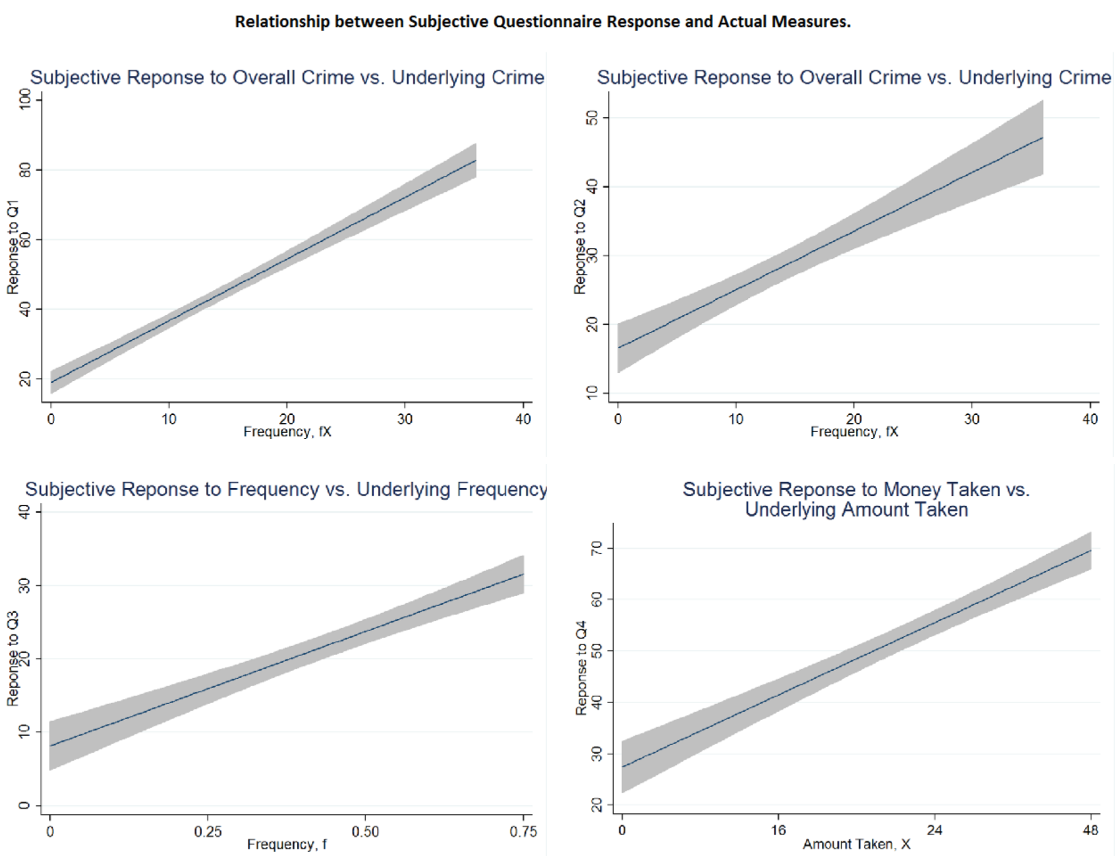


Figure 6: The basic correlation between each dependent variable and the underlying measure is shown here. The shaded area represents the 95% confidence interval from a line of best fit.

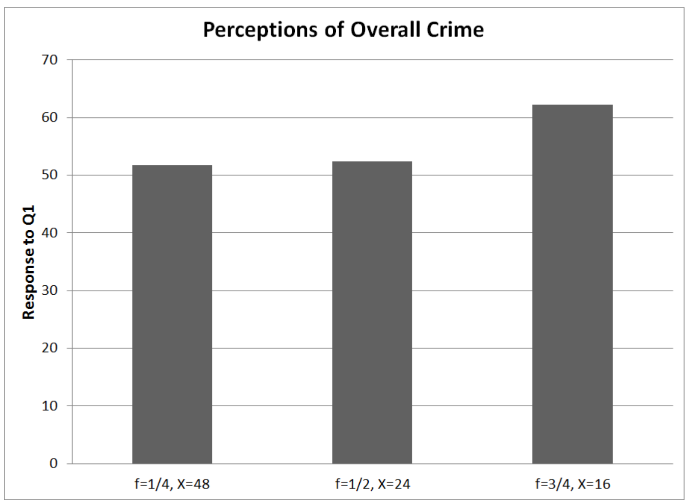


Figure 7: We examine how survey responses differ across a fixed level of crime, where $f \cdot X = 12$: $(\frac{1}{4}, 48)$, $(\frac{1}{2}, 24)$, and $(\frac{3}{4}, 16)$. Investors believe that crime is worse in those environments where the frequency is higher (even when the expected payoff is the same).

Table 6: **Biases in Perceptions of Crime.** At the end of every environment, participants were asked "Please rank the following statement: The fact that this environment allowed observers to take money away from investors was a serious problem for me. Please enter an integer value (0 = completely disagree; 100 = completely agree)." The answer to this question was used as the dependent variable and regressed against a number of biases.

Dependent Variable = Numerical Response to General Subjective Question (Q2)							
Standard Errors Clustered by Participant							
	model 1	model 2	model 3	model 4	model 5	model 6	model 7
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Frequency of Theft, (f)	49.532***						
	[4.946]						
Size of Theft, (X)	0.052						
	[0.071]						
$f \cdot X$		0.937***					
		[0.128]					
Explicit Theft			2.824***		2.917***	2.695***	
			[0.164]		[0.182]	[0.167]	
Implicit Theft			0.182		0.151	-0.398***	
			[0.127]		[0.134]	[0.151]	
Explicit (first half)				2.085***			1.769***
				[0.341]			[0.358]
Explicit (last half)				3.458***			3.303***
				[0.333]			[0.333]
Implicit (first half)				-0.306			-0.898***
				[0.278]			[0.298]
Implicit (last half)				0.647**			0.040
				[0.278]			[0.279]
Return					-0.088		0.086
					[0.069]		[0.075]
Streaks						2.258***	2.451***
						[0.426]	[0.463]
Constant	4.455***	15.325***	9.794***	10.082***	11.540***	3.398***	1.440
	[1.693]	[1.666]	[1.437]	[1.441]	[2.009]	[1.254]	[2.160]
R^2	0.130	0.084	0.278	0.284	0.279	0.308	0.316
Observations	1008	1008	1008	1008	1008	1008	1008

* significance at 10%, ** significance at 5%, *** significance at 1%

Table 7: **Biases in Perceptions of Crime.** At the end of every environment, participants were asked "On an absolute scale from 0 to 100, how would you rate this current environment? Please enter an integer value (0 = Best – In this environment, any amount that I invest realizes its full return, with certainty; 100 = Worst – In this environment, any amount that I invest is taken away from me entirely, with certainty)." The answer to this question was used as the dependent variable and regressed against a number of biases.

Dependent Variable = Numerical Response to General Subjective Question (Q1)
Standard Errors Clustered by Participant

	model 1	model 2	model 3	model 4	model 5	model 6	model 7
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Frequency of Theft, (f)	59.654***						
	[4.334]						
Size of Theft, (X)	0.492***						
	[0.066]						
Money Lost		1.218***					
		[0.117]					
Explicit Theft			0.704***		1.754***	0.749***	
			[0.178]		[0.178]	[0.171]	
Implicit Theft			2.034***		2.453***	0.286*	
			[0.190]		[0.162]	[0.163]	
Explicit (first half)				-0.095			0.484
				[0.385]			[0.374]
Explicit (last half)				1.380***			1.635***
				[0.319]			[0.306]
Implicit (first half)				1.445***			0.328
				[0.487]			[0.411]
Implicit (last half)				2.695***			1.059**
				[0.446]			[0.417]
Return					-0.923***		-0.317***
					[0.056]		[0.063]
Streaks						5.818***	5.035***
						[0.283]	[0.337]
Constant	2.881*	31.217***	31.535***	31.680***	45.648***	7.578***	15.732***
	[1.631]	[1.765]	[1.781]	[1.783]	[2.037]	[1.770]	[2.420]
R^2	0.299	0.087	0.111	0.117	0.267	0.413	0.430
Observations	1008	1008	1008	1008	1008	1008	1008

* significance at 10%, ** significance at 5%, *** significance at 1%

Table 8: **Biases in Perceptions of Crime.** At the end of every environment, participants were asked "Please rank the following statement: The fact that this environment allowed observers to take money away from investors was a serious problem for me. Please enter an integer value (0 = completely disagree; 100 = completely agree)." The answer to this question was used as the dependent variable and regressed against a number of biases.

Dependent Variable = Numerical Response to General Subjective Question (Q2)
Standard Errors Clustered by Participant

	model 1	model 2	model 3	model 4	model 5	model 6	model 7
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Frequency of Theft, (f)	49.532***						
	[4.946]						
Size of Theft, (X)	0.052						
	[0.071]						
Money Lost		1.850***					
		[0.144]					
Explicit Theft			2.776***		2.927***	2.788***	
			[0.170]		[0.182]	[0.168]	
Implicit Theft			0.379		0.439	-0.110	
			[0.278]		[0.280]	[0.293]	
Explicit (first half)				2.013***			1.927***
				[0.336]			[0.356]
Explicit (last half)				3.424***			3.325***
				[0.335]			[0.333]
Implicit (first half)				-0.151			-0.695*
				[0.386]			[0.398]
Implicit (last half)				0.978**			0.328
				[0.450]			[0.466]
Return					-0.133**		0.096
					[0.063]		[0.076]
Streaks						1.625***	1.840***
						[0.345]	[0.406]
Constant	4.455***	10.743***	10.169***	10.301***	12.202***	3.479***	1.249
	[1.693]	[1.566]	[1.476]	[1.480]	[1.951]	[1.260]	[2.153]
R^2	0.130	0.201	0.279	0.284	0.282	0.302	0.308
Observations	1008	1008	1008	1008	1008	1008	1008

* significance at 10%, ** significance at 5%, *** significance at 1%