

Individual Updating of Subjective Probability of Homicide Victimization: a “Natural Experiment” on Risk Communication

José Raimundo Carvalho* Diego de Maria André† Yuri Costa‡

50° ANPEC, 2022: Área 12 - Economia Social e Demografia Econômica

ABSTRACT: We investigate the update of subjective homicide victimization risk after an informational shock by developing two econometric models able to accommodate both optimal decisions of changing or keeping *prior* expectations, which enables us to rationalize *skeptical* bayesian agents with their disregard to new information. We apply our models to a unique household data. It consists of socioeconomic and victimization expectation variables in Brazil, coupled with an informational “natural experiment” brought by the sample design methodology, which randomized interviewers and interviewees. Our results show that our bayesian update approach is reasonable. The higher *priors* about own subjective homicide victimization risk are set, the more likely individuals are to change their initial perceptions. Furthermore, we can rationalize *skeptics* by means of an informational quality/credibility argument. In case of an update, we find that older participants and females are more reluctant to change *priors* and choose the new response level. In addition, even though respondents’ level of education is not significant in our exercise, interviewers’ level of education has a key role in changing and updating decisions.

Keywords: Bayesian Update, Subjective Probabilities, Fear of Crime

JEL Code: D83, D84, C24

RESUMO: Investigamos a atualização do risco de vitimização subjetiva de homicídios após um choque informativo, desenvolvendo dois modelos econométricos capazes de acomodar tanto as decisões ótimas de mudança ou de manutenção de expectativas *priors*, o que nos permite racionalizar agentes Bayesianos *céticos* e suas desconsiderações por novas informações. Aplicamos nossos modelos a um conjunto único de dados domiciliares que consiste em variáveis socio-econômicas e de expectativas de vitimização no Brasil, juntamente com uma “experimento natural” informacional trazido pela metodologia de desenho amostral que aleatorizou entrevistadores e entrevistados. Nossos resultados mostram que a nossa abordagem de atualização bayesiana é razoável. Quanto mais elevados forem as *priors* sobre o próprio risco de vitimização subjetiva de homicídio, mais provável é que os indivíduos mudem suas percepções iniciais. Além disso, podemos racionalizar os *céticos* por meio de um argumento de qualidade/credibilidade informativa. No caso de uma atualização, verificamos que os participantes mais idosos e as mulheres são mais relutantes em mudar suas *priors* e escolher o novo nível de resposta. Além disso, embora o nível de educação dos entrevistados não seja significativo no nosso exercício, o nível de educação dos entrevistadores tem um papel fundamental na mudança e atualização das decisões.

Palavras-chaves: Atualização Bayesiana, Probabilidades Subjetivas, Medo do Crime

*Director CEA/EIDEIA/UFC, Full Professor CAEN/UFC, and LECO/UFC. Author for correspondence (jose-carv@ufc.br). We thank participants of the Workshop on Subjective Expectations, Federal Reserve Bank of New York and the ESRC Research Centre on Micro-social Change, New York, USA - May 12-13, 2016, as well as those from the **AL CAPONE 2017** Meeting, Universidad de Chile, Santiago - Chile. This manuscript circulated before as “Modeling Bayesian Updating with many Non-Updaters: The Case of Own Subjective Homicide Victimization Risk”.

†PPECO/UFRN, NEAQ/UFRN and LECO/UFC.

‡LECO/UFC, Brazil. I would like to acknowledge the financial support from CNPq/Brazil.

1 INTRODUCTION

The last four decades of economic research are characterized by major advances in applied microeconomics. Issues previously restricted to other social sciences have been included into economists' research agenda due to an increasing interest in interdisciplinary topics. This is the context in which the economics of crime emerged with the seminal paper of Becker (1968). In that paper, he models the engagement in illegal activities as the result of rational decisions following a cost/benefit analysis. Based on this idea, discussions on deterrence and punishment strategies were raised assisting the design of public policies for crime control.

Beside criminals, delinquency also involves a victim. In this sense, restricting the analysis of crime on the criminals' perspective is not exhaustive: the victim's side should also be put under scrutiny. Although it is important to stress that becoming a victim is a rare event specifically focused on certain socioeconomic groups, the fear of being victimized, i.e. the subjective expectation of becoming a victim or the "fear of crime", plays a significant role in welfare.

Indeed, the issue of "fear of crime"¹ (sometimes denominated "sensation of (un)safety") and the correlate concept of risk (future probability) of victimization occupy an important space not only in scientists' agendas but also of public safety managers (Warr (2000) and Ferraro (1995a)). As a consequence, policies and actions to manage crime and the fear of crime must be handled as distinct social problems (Ferraro (1995a) and Vanderveen (2006)), although intrinsically related. Thus, given the importance of fear and the weak probabilistic justification for such a severe concern, understanding the factors underlying crime expectations and how information affect those risk perceptions, or "fear of crime", is a gap we want to fill in.

In summary, this paper, much in the spirit of the literature initiated by Viscusi (1979) in a labor market context, assumes that citizens are rational agents with imperfect information on the actual probability of becoming a crime victim. Nevertheless, they might face information about crime, not necessarily credible, in such a manner compatible with a learning process. In light of this understanding, we want to study subjective expectations in the context of victimization, investigating the role of informational shocks on posterior risk perceptions, say, "fear of crime". **More specifically, we aim to look into the factors working on the update of individuals' "fear of crime" after they receive a shock of information regarding official crime rates.**

For this task, we seek to develop an estimable model able to accommodate both optimal decisions on changing or keeping the same initial perception with respect to subjective homicide victimization risk. This enables us to rationalize the lack of use of new information, contributing to the literature by allowing the existence of non-updaters in a bayesian updating context. Furthermore, we search for observable characteristics to explain this two-step decision, i.e. i) Changing or keeping initial perception; ii) Setting an update (if any).

The bayesian approach proposes a link between current and past beliefs based on further evidence. Hence, situations involving initial perceptions followed by informational shocks are ideally suited for such analysis. However, since this approach fundamentally deals with subjective probabilistic beliefs, two relevant concerns arise: i) Are individuals able to think in terms of probabilities? ii) Are those responses reliable? Despite these concerns, there is strong evidence supporting the bayesian approach in situations where respondents face relevant life aspects, such as becoming victimized by homicide.

This is the case, for instance, appearing in Viscusi and O'Connor (1984). They investigate if chemical industry workers learn about the risks they face on the job and how this understanding affects their reservation wages. They conclude that the manner workers update their probabilistic beliefs is in compatibility with a bayesian procedure. This main result is also found in Viscusi (1985). In addition, an application of this technique in health economics can be found in Smith and Johnson

¹For our purposes, we adopted the notion advocated by Warr (2000) who defines the "fear of crime" as an emotion, a feeling of alarm or fear caused by an awareness or expectation of danger. Therefore, fear of crime affects people negatively, many times more than crime itself.

(1988): they analyze agents' ability to process information about the health risks of being exposed to a harmful substance. The main message brought by this paper is that information neither recent nor relevant to individuals has insignificant effects on risk perceptions. On the other hand, information seen as more updated and relevant appears to be positively correlated with the posterior probability response.

The dataset used in our paper also presents strong evidence to back up that last statement. Our data derives from a household survey conducted in the city of Fortaleza, Brazil, between October 01, 2011 and January 19, 2012. Twenty interviewers applied a questionnaire to 4,030 citizens of Fortaleza dealing with subjective expectations about being a victim of homicide on the forthcoming twelve months, among other issues. Apart from its size and detailed information on both interviewees and interviewers, perhaps the most distinguished feature coming out of our sample is that the vast majority of individuals (95%) kept their priors about the risk of being victimized. This is so in spite of the fact that their perceptions were far from official rates and that they received information about the true value. We will refer to this behavior as *skepticism* and to these respondents as *skeptical agents*.

Instead of questioning the rationality of individuals who decided not to update expectations in a learning context (*skeptical agents*), we decided to approach our empirical evidence slightly different: we might be facing skeptical agents whose perceived new informational content lacks the minimum level of relevance and/or trust to (rationally) prompt any change in risk perceptions. Our results support such understanding. For instance, in every model we estimated, interviewers' level of education, a central variable concerning informational quality and credibility, has positive estimated coefficients regarding the changing of subjective expectations decision. The same is implied by a matching variable in gender, whose positive coefficients may tell us that homophily can either facilitate understanding in communication or signal trustworthy information – or both.

Our paper is comprised of five more sections. Section 2 presents a brief discussion on subjective expectations and Bayesian update. Then, it analyses the contributions of Viscusi and O'Connor (1984) seminal paper and its influence on Smith and Johnson (1988) with great detail. Finishing the section, a recent application of subjective expectations in Delavande (2008) is presented. Section 3 presents the data, emphasizing its advantages in comparison with previous papers and our major challenge: to deal with *Skeptical Agents*. Also, it brings an initial exploratory analysis, as well as a brief discussion on selection bias. Econometric models are presented in Section 4. An extension of Viscusi and O'Connor (1984)'s and Smith and Johnson (1988)'s models with a multiple linear regression is found in subsection 4.1. Subsection 4.2 develops two models to overcome the theoretical difficulties raised by our data and Section 5 shows all estimation results. Finally, Section 6 concludes, summarizing major messages and proposing further directions of research.

2 LITERATURE REVIEW

2.1 Subjective Expectations and Bayesian Update

Given the connection between initial perceptions and further evidence, the bayesian appeal for situations where agents are exposed to new informational bits is straightforward. Nevertheless, a previous step for its use is of great resistance in economics and econometrics: traditionally, scholars have been denying the necessary credibility² to information not objectively observed. Advocates of such understanding base their approach on Samuelson (1938)'s point of view that the only information needed and usable is that of behavior³. Proceeding this way, Samuelson (1938) suggests that it is possible to recover initial preferences and the risk of making wrong assumptions about the introspection process

²See Machlup (1946) and Hart et al. (1960).

³An example of this acquaintance in Econometric Theory is found in McFadden (1973).

is avoided. Also, it prevents any inconsistency raised by possible differences between what is said and done by the decision maker.

In this sense, whenever it is necessary to comment on the expectations formation process, the rational expectation assumption is usually taken for granted, or, at least, bounded rationality. However, as pointed out by Pesaran (1987), rational expectations has not proven to be sufficiently convincing to analysis out of macroeconomic long-run even in its weaker version. By the same token, Manski (2004), when looking at returns in a schooling context, argues that students face the same inferential problems that labor econometricians do. Therefore, it is implausible to assume that students share the same objective and correct distribution function.

The problem is that by relaxing the assumption of a unique objective probability distribution, perfectly known and identically processed by all individuals, the standard analysis cannot take place. This is so because any specific choice behavior may arise as a consequence of many different preferences and probabilities combinations: a classical identification problem. Proposing a different perspective, Manski (2004) presents a detailed defense in favor of using expectations self-reported in probabilistic terms, the so called *subjective probabilities*⁴. This data type could then be used to validate or relax hypothesis used in traditional models, i.e. those based on objective probabilities only.

In light of such perspective, several papers have successfully applied the concept of subjective probabilities in many situations where expectations are crucial in their analysis. For example, Manski have co-authored studies in income expectations (Dominitz and Manski (1997)), social security and retirement (Dominitz et al. (2002)) and consumer confidence (Dominitz and Manski (2004)), among others. Also, an application in survival expectations is presented in Hurd and McGarry (1995). Once we have brought the concept of subjective probabilities up to the surface, criticisms about the reliability of that data type started to appear. However, before we arrived at that junction we have to go over the important early contributions of W. Kip Viscusi on applying bayesian update approaches to model updating risk perceptions.

2.2 Early Contributions

Viscusi (1979), interested in the relation between risks for health and physical integrity at work and labor market outcomes, finds that workers' perceived risk has a positive correlation with industry specific risks. Also, he shows that industries whose employees have higher risk perceptions present higher wages as well, supporting the theory of compensating differentials. Therefore, understanding how risk perception evolves and how it is processed emerges as a further step on the study of labor markets dynamics.

Viscusi and O'Connor (1984) investigate whether employees revise their risk perceptions and, if so, how the reservation wage revision under a new informational set would be. The authors applied a questionnaire to 335 employees of three major American chemical companies. Workers were asked to mark their risk perception in a linear scale ranging from very safe to dangerous, with the US private sector accident and diseases rate presented as a reference point. Finally, answers were converted into probabilistic terms giving rise to the variable *RISK*.

The first message of Viscusi and O'Connor (1984) is that the data is consistent with a model in which the choice of employment when facing risks follows a learning process. Reservation wages grow as perceived risks increase, so that there is a risk premium keeping workers at their jobs until such compensation is not enough. From that point on, they decide to leave the job or, at least, not to accept the offer if this choice was put back.

The major contribution of this paper, however, is to investigate the informational shock effect on workers' perceptions, which is exactly what we want to do in a victimization context. The authors present each worker a label containing information regarding a chemical substance that would replace

⁴A first attempt to use subjective probabilities in economics is due to Juster (1966), though.

the old products used in everyday work activities. Then, workers were asked to repeat their responses under this new information set. Regression results show that workers *posteriors* have been influenced by both *priors* and label contents, creating a reviewing process consistent with a bayesian approach. This is the starting point for our initial estimable model presented in subsection (4.1).

Motivated by Viscusi and O'Connor (1984), Smith and Johnson (1988) investigate how a sample of households in Maine, United States, responded to information regarding risks associated with radon concentration in their homes and water supplies. Whereas radon problems in the area had been published widely since the 1970's, the authors seek to address the question more specifically. Besides distributing pamphlets with information about health risks caused by this chemical substance, they informed each household under study the radon concentration in its residence.

The analysis is based only on 117 of 230 observations, those households whose information provided is complete. Although questions may be raised about a possible selection bias, the authors argue that much of the sample loss is due to incomplete information on socioeconomic and behavioral variables not related to one's ability in providing a valid response. The results bring up five central conclusions validating the use of available socioeconomic and environmental information in updating expectations:

- i) The level of exposure to the radon information provided has a positive impact on the *posterior* magnitude;
- ii) Mitigating actions are negatively related to risk perceptions;
- iii) Cancer victims are very likely to increase *posterior* perception and the weight associated with the radon concentration;
- iv) Life expectancy and age exhibit statistically insignificant results;
- v) Individuals perceive new information with a third of their initial expectations;

Somehow extending the discussion presented in Viscusi and O'Connor (1984) to include other information than *priors*, Smith and Johnson (1988) provides an important insight for our study. Since the extent on which information is perceived determines its use when updating expectations, we should be cautious with these aspects in our analysis. Also, it is important to note that, despite the empirical success of these articles, we cannot overlook studies observing some nonconformity between experimental or empirical results and its underlying theory. As pointed out by Tversky and Kahneman (1974), the literature generally recognizes several specific decision rules (*heuristics*) used unconsciously and not necessarily in an optimal way to solve complex problems.

As it is the case for homicide victimization, dealing with rare events is an even greater challenge for most individuals. As an illustration of this, Kunreuther et al. (1978) shows that risk-averse consumers decided not to purchase insurance against floods and earthquakes even though they had been largely subsidized to do so. However, despite the difficulty in modeling choice behavior in this context, papers such as Lichtenstein et al. (1978) indicates a pattern: regarding risk perceptions, agents tend to overestimate low probabilities and underestimate those of high magnitude.

Anticipating part of the data discussion found in Section 3, our empirical evidence is an example of this pattern. There is a huge gap between respondents' initial perceptions and official values. The sample *prior* mean is strikingly 943 times greater than the official rate for homicide. Viscusi (1985) shows that this result is exactly what one might expect from a bayesian learning process as a final argument for using a bayesian update model in this context.

2.3 Recent Developments

More recent papers on subjective expectations focus primarily on the development of probabilistic elicitation mechanisms. More specifically, they aim to assess beliefs in terms of probabilities and to

take the extensive evidence about heuristics on the updating process into account. Naturally, issues concerning expectation revisions add even more challenges than collecting subjective probabilities. In this context, researchers must be aware of how that reviewing process takes place.

In an experimental setting, El-Gamal and Grether (1995) find evidence on the use of certain heuristics – *conservatism* and *representativeness* – in addition to the Bayes’s rule when updating subjective probabilities. In general, results have shown so far that this rule is the most likely approach used by agents, but researchers should be alert to the fact that it is not the only one. This message is also found in Dominitz and Manski (2011). Looking at expectation revisions on equity returns, their paper evidences extensive heterogeneity on the updating process.

In light of these results, Delavande (2008) develops the *ERS* (*Equivalent Random Sample*), a metric for subjective expectation updates, in order to accommodate conflicting results about the specific revision process used by individuals. This metric is developed as follows: Let P_i be the true probability that an individual i experience the occurrence of B , a binary event. Assume that i is unaware of the value of P_i , so that he has only a subjective expectation about it. Furthermore, let $f_{i,1} \in \Gamma$ be the distribution function of this initial subjective expectation, where Γ is the set of all distribution functions defined on the interval $[0, 1]$.

The *ERS* was designed to access posterior subjective probability distributions. Nevertheless, the idea behind it can also be used to discuss informational perceptions and/or credibility. We believe the decision and magnitude of an update is critically dependent on information quality and a no-changing decision perhaps is the best evidence to explore this idea.

3 DATA

Before starting the description of our sample, we want to emphasize some key aspects that make us believe in its uniqueness. Firstly, we are dealing with crime risk perception data, a type of empirical evidence hard to find in the literature of risk and information, despite its relevance for social welfare. Furthermore, our sample relates to a developing country, and it is well known that obtaining microdata in such societies challenges researchers of all fields. Focusing on crime issues, according to a recent study conducted by United Nations (2013), Brazil shows one of the highest homicide rates in the world. Therefore, if one wants to study crime perceptions, our data offers rare information of individuals living in a pertinent environment.

Secondly, as to subjective perceptions, the amount of observations is far greater than what is often presented in some important papers in the field. For example, Viscusi and O’Connor (1984) presents no more than 400 respondents in a seminal paper. Our sample is more than 10 times greater than that, presenting responses of individuals from all of the 116 districts of Fortaleza with an ample range of age, education and income, to name some of the available information.

In addition, unlike Delavande (2008), who conducted herself the survey, we have twenty different interviewers. If we want to model information use and updating decisions, having a unique interviewer might introduce selection bias due to the fact that informational source does not change. More importantly, the random variation coming from the sender’s side, raising differences in gender, age and educational levels, enables us to investigate aspects of matching with no concerns of self selection, since both participants were not able to choose their matches⁵.

Table (1) presents all variables we are going to use in this study. Overall, the big picture we are going to discuss in detail is depicted by a middle-age, non-white and female sample, with low income and educational levels. In other words, it is a relevant representation of Fortaleza’s residents.

Out of 4,030 observations, in which 2,190 are women and 1,840 are men, the average age is slightly more than 39 years old. With respect to marital status, we have 1,928 single individuals, while the portion with a partner sums 2,102. Respondents were also able to identify themselves as White,

⁵The interview was held in a randomized household and interviewers were also randomly allocated.

Table 1: Variable Definitions

Variable	Description	Range
<i>Prior</i>	Initial subjective probability about becoming a victim of homicide on the following 12 months from the interview day onwards.	[0,100]
<i>Change</i>	A dummy for respondent decision to change initial perception.	0: No; 1: Yes.
<i>Post*</i>	Final subjective probability about becoming a victim of homicide on the following 12 months from the interview day onwards.	[0,100]
<i>Gender</i>	A dummy for respondent's Gender.	0: Male; 1: Female;
<i>Age</i>	Respondent's age in years.	[16,94]
<i>Police</i>	A level variable assessing the police work in terms of crime control around the area of respondent residence.	1: Excellent; 2: Good; 3: Fair; 4: Poor; 5: Terrible.
<i>Educ_Int</i>	A dummy for interviewer's educational level.	0: High School; 1: More than High School.
<i>Matching_Gender</i> [†]	A dummy for the existence of a gender matching between respondent and interviewer.	0: No; 1: Yes.

* Note that, depending on *Change*, $Post = Prior$ or $Post \neq Prior$.

† Note that it does not matter what specific gender matching is observed.

Source: Elaborated by the author.

Black, Mestizo, Asian or Indigenous. Grounded on the standard approach, we classified our sample into White and Non-White, finding only 31% as White. In addition, respondents could choose between 8 levels of education and income. Following the Brazilian high income concentration and low educational level, the majority of our sample earns 1 to 2 minimum wage and has an incomplete elementary school degree.

Getting into our major focus, which concerns the probabilities reported, after a preliminary training in basic probability concepts, participants faced an initial question posed as:

“Regarding numerical values, what is the chance (or probability) of you being a victim of the following crimes in Fortaleza in the next 12 months: a) Homicide; b) Robbery”.

Despite the controversy surrounding human ability to think probabilistically, and in accordance with Ferrell and McGoey (1980) and Koriat et al. (1980), only a bit more than 1/4 of our original sample is not able to give a valid response to this question⁶. We consider this result satisfactory and, due to our large data set, we can eliminate those responses and still keep a large number of observations to work with: 2,885. This first response is referred as *Prior*.

The interviewer, after receiving this initial probabilistic perception, gives an informational shock as follows:

“According to the Ministry of Justice, in 2009 the probability of a person in Fortaleza: a) being victim of homicide was 0,037% (37 homicides for each 100.000 inhabitants) and b) being robbed was 0,96%”.

⁶Those cases are due to participants whose answers are either “*Didn't Know*” or “*Didn't Answer*”.

(960 thefts for each 100.000 inhabitants)".

Interviewees are asked if they want to change their initial responses and, if so, to what value. We call this decision to change as *Change* and the posterior perception as *Post*.

Table (2) summarizes the sample statistics. We argue that our interaction outcomes rely not only on the respondent's side, but also on the sender's side and on common features between both of them. Since we are trying to model perceived information, we should consider message quality aspects and a first natural choice is the interviewers' level of education. All of our 20 interviewers completed at least high school studies and 80% of them are undergraduate students or graduates.

Table 2: Descriptive Analysis – Homicide

	Mean	Std.dev	Min	Max	Missing
Prior	34.91	26.49	1.00	98.00	0
Change	0.04	0.21	0	1	0
Post	33.29	26.63	0.10	98.00	0
Gender	0.53	0.50	0	1	0
Age	38.55	16.23	16	94	56
Police	2.95	0.96	1	5	22
Educ_Int	0.83	0.37	0	1	0
Matching_Gender	0.77	0.42	0	1	0

Source: Elaborated by the author.

Equally, sources of information noise are concerns to take into account. Note that differences among interviewers and interviewees possibly raise obstacles on information, and it works against an update. In this sense, an important concept to this analysis is *Homophily*. McPherson et al. (2001) defines it as the contact between similar people occurring at a higher rate than among dissimilar people. It is important to keep in mind that, at first, homophily can be related to race, gender, age, religion, social status or any other similar characteristic relating individuals. For example, whenever a group exhibits one of these heterogeneities at a higher rate than it would be found randomly, there is homophily.

The impact of these differences on spontaneous groups' formation and networks ties is widely documented since the early 1920's. For instance, Bott (1928) noted that school children formed friendships and play groups at higher rates if they were similar on demographic characteristics. The implications of homophily to information diffusion is a research agenda of recent papers such as Jackson (2009) and Golub and Jackson (2012). It is worth stressing that one of the main econometrics concern related to matching/group formation is the fact that those matchings are not exogenous. Nevertheless, we are able to bypass this issue here because of the very design of our survey: it imposes double randomization so that the sender - receiver pair is exogenously determined (see, among others, Graham (2011)).

Using this result in our problem, it is possible that the interaction is affected by a particular heterogeneity between interviewers and interviewees. We conjecture that in Latin cultures gender plays a significant role on information credibility/use. Hence, variable *Matching_Gender* was created to control this issue. In addition, given our interest in victimization risk perception, the way respondents evaluate police forces is a very important piece of information to control for. In a range of five levels, individuals rated the police work around its middle point, which stands for Fare.

Finally, our subjective probabilities, i.e. *Prior* and *Post*, take almost the full range [0, 100] with a huge standard deviation. This initial result piles up to the empirical evidence presented in Manski (2004), refuting usual concerns about subjective probabilities being reported in values around 50%. Also, as in Delavande et al. (2011), we refute the fears that poor, illiterate individuals in developing countries do not understand probability concepts. Equally, the significant difference between respondents's initial perceptions and the official rate could be another source of distrust in

such data type. Nevertheless, it finds reverberation in previous papers such as Dominitz and Manski (1996) and Bursik Jr et al. (1999) also attesting the existence of a crime overestimation.

The fact is: although a high discrepancy between subjective victimization probabilities and official rates is not new, our participants were informed about the “true” probability and still less than 5% of them changed initial perceptions⁷. Therefore, given that initial perceptions are so overrated, what could explain almost no change in responses?

It is important to keep in mind that we have two different steps after individuals receive an informational shock: i) Whether or not to change initial responses; ii) Given that a change is going to happen, to what value it is going to be set. In a sample of 2885 responders, 2758 of them say no to the first question and 127 behave accordingly to previous studies on bayesian updating processes. Since we have a reasonable updating subsample size here, we can restrict our analysis to this subset of observations and look for any noticeable difference between updaters and the entire sample. Table (3) presents the results.

Table 3: Descriptive Analysis – Updating Subsample

	Mean	Std.dev	Min	Max	Missing
Prior	39.55	28.43	1.00	90.00	0
Change	1	0	1	1	0
Post	2.64	2.70	0.10	15.00	0
Sex	0.54	0.50	0	1	0
Age	33.95	14.63	16	78	6
Police	2.99	0.86	1	5	0
Educ_Int	0.91	0.29	0	1	0
Matching_Sex	0.93	0.26	0	1	0

Total of 127 observations.

Source: Elaborated by the author.

Even though we are dealing with a critical question, in which anchoring in expectations is known to be quite frequent, and the amount of information provided might not be ideal, we find that individuals who decide to reevaluate their responses do it in a scathing way. In this case, the subsample set *Post* almost 15 times less than *Prior*. This evidence suggests that our conjecture about the impact of senders’ educational level and gender matching into message quality might be on the right path. When we look back to table (2) and compare it with table (3), we find that *Educ_Int* and *Matching_Gender* means increased considerably from 83% and 77% to 91% and 93%, respectively, in our updating subsample. Furthermore, *Prior* and *Age* are other variables that this crude analysis indicates to be aware of. The results show that our updaters are younger individuals with higher initial perceptions.

In summary, this very simple exercise induces us to believe that a selection bias might be working behind the scenes. Note that the *Post* reported is conditioned to a previous decision and, when we select our sample only for individuals who decide to change initial responses, this subgroup has some different characteristics from the entire sample. Those characteristics might be just what we are looking for.

In short, the data used here also attests the existence of a crime overestimation found in previous works. The novelty is that our respondents faced an informational shock consisting of the official homicide rate, but 95% of them keep the same initial perception. So far, looking at the updating subsample only, we found that these updating interactions were made by more educated interviewers, sharing the same gender with our interviewees. Thus, those two most important results enable us to follow our main line of explanation: informational quality.

In the next section, we start discussing our models. Subsection (4.1) is the first step of our contribution. We extend Viscusi and O’Connor (1984) and Smith and Johnson (1988) by proposing

⁷Actually, following Hoffrage et al. (2000), we presented the same information in numbers (*37 in 100.000*), which brings better results in the participants’s use of probabilities.

a multiple linear regression motivated by victimization expectations. We believe subsection (4.2) is our main achievement. It proposes two alternative models rationalizing the skepticism found in our data and it provides an extension of current models to allow for the existence of skeptical bayesian updaters.

4 ECONOMETRIC MODELS FOR BAYESIAN UPDATING

4.1 Multiple Linear Regression and Victimization Expectation Update

Consider a victimization event, where $y = 1$ if the individual becomes a crime victim and $y = 0$ on the contrary. Let $Y \sim \text{Bern}(\pi)$, *iid*, and assume that the respondent has a probability distribution $\text{Beta}(\alpha, \beta)$ about the true, but unknown, probability π of becoming a homicide victim.

In this context, we propose the following interaction structure: the initial respondent's perception, $\text{Prior}_i \equiv \pi_i^0$, can be modeled as dependent on his observable characteristics, for instance, age, gender, income, education, etc. Formally, we can define a function $f : \mathbb{R}^q \rightarrow [0, 1]$ such that $\pi_i^0 = f(\tilde{\mathbf{X}}_i^r)$, where $\tilde{\mathbf{X}}_i^r$ is the q dimensional heterogeneity vector of individual i , the **receiver (interviewee)**. When asked about his initial perception, he assess $P(y = 1 | \tilde{\mathbf{X}}_i^r) = E(\pi | \tilde{\mathbf{X}}_i^r) = \pi_i^0$ and establishes $\pi_i^0 = \frac{\alpha_i}{\alpha_i + \beta_i}$ as response.

After collecting π_i^0 by means of the questionnaire, **sender (interviewer) j** sends an informational bit consisting of the true population value, π (homicide prevalence rate in Fortaleza). However, from the respondent's perspective, this information may lack credibility and its content may not be fully taken into account. In fact, let π_i^* be the proper value depicted by this information (from the perspective of the receiver) and define $g : \mathbb{R}^k \rightarrow [0, 1]$, with $\pi_i^* = g(\mathbf{X}_i^r, \mathbf{X}_j^s, \mathbf{X}_{i,j}^m, \pi_i^0 | \pi)$, where \mathbf{X}_i^r and \mathbf{X}_j^s are vectors of observable heterogeneity for receivers⁸ and senders, respectively, and $\mathbf{X}_{i,j}^m$ is a vector of matching aspects.

Figure 1: Time Sequence of Sender-Receiver Interaction

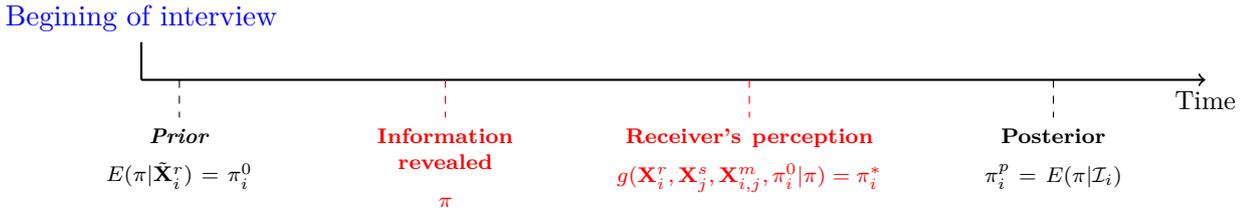


Figure 1 sketches the interaction between sender and receiver so far described. So, we have the following important objects already defined:

1. $\pi_i^0 = f(\tilde{\mathbf{X}}_i^r)$, respondent's prior about his own victimization risk;
2. $\pi_i^* = g(\mathbf{X}_i^r, \mathbf{X}_j^s, \mathbf{X}_{i,j}^m, \pi_i^0 | \pi)$, receiver's perception from the information revealed by sender j when the latter declares that the real victimization risk is π ;

Abstractly, the information was translated into a sequence of victimization observations, generating an informational content $\mathcal{I} = (0, 1, \dots, 0, 1)$ with n_1 values equal 1 and n_0 equal 0, where $N = n_0 + n_1$. In summary, the respondent, whose initial perception was π_i^0 , uses the perceived information, π_i^* , to generate \mathcal{I}_i , creating a set of new evidence which is the foundation to update his response towards

⁸The vector determining π_i^0 is not necessarily the same for π_i^* .

a posterior perception. This new response, $Post \equiv \pi_i^p$, is given by $P_i(y = 1|\mathcal{I}_i) = E(\pi|\mathcal{I}_i) = \pi_i^p$. In this structure:

$$\begin{aligned} p_i(\pi|\mathcal{I}_i) &\propto p_i(\mathcal{I}_i|\pi)p_i(\pi) \\ &\propto \pi^{n_1}(1-\pi)^{n_0} \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} \pi^{\alpha_i-1}(1-\pi)^{\beta_i-1} \end{aligned}$$

Then,

$$\pi|\mathcal{I}_i \sim \text{Beta}(\alpha_i + n_1, \beta_i + n_0)$$

And,

$$\pi_i^p = E(\pi|\mathcal{I}_i) = \frac{\alpha_i + n_1}{\alpha_i + \beta_i + n_1 + n_0}$$

For estimation purposes, π_i^p remains undefined, since it is not possible to identify none of the given parameters. One way to dodge this problem is to realize that $\pi_i^* = \frac{n_1}{N}$ and use the fact that $\pi_i^0 = \frac{\alpha_i}{\alpha_i + \beta_i}$ to establish:

$$\pi_i^p = \frac{N(\frac{n_1}{N}) + (\alpha_i + \beta_i)\frac{\alpha_i}{(\alpha_i + \beta_i)}}{\alpha_i + \beta_i + N}$$

Therefore, we are able to write our structural model as:

$$\pi_i^p = \frac{N}{\alpha_i + \beta_i + N} \pi_i^* + \frac{\alpha_i + \beta_i}{\alpha_i + \beta_i + N} \pi_i^0 \quad (1)$$

Now, if we assume that $g : \mathbb{R}^k \rightarrow [0, 1]$ can be reasonable approximated by a linear form, it allows us to extend Viscusi and O'Connor (1984) and Smith and Johnson (1988) at the same time by taking not only respondent's heterogeneity into account, but also both interviewer and matching information through a coherent conditioning multiple linear regression as follows::

$$\pi_{ij}^p = \lambda + \gamma \pi_i^0 + \mathbf{X}_{ij} \delta + \epsilon_{ij} \quad (2)$$

Where π_{ij}^p is the posterior of receiver i , after receiving information from sender j , \mathbf{X}_{ij} is an extended vector of observed heterogeneity containing \mathbf{X}_i^r , \mathbf{X}_j^s , \mathbf{X}_i^m , and ϵ_{ij} the error term.

4.2 Bayesian Updating with Skepticism

As exposed in Section 3, we have a very different sample in the context of bayesian update. Our data presents an excess of individuals — our skeptical agents — who do not update and keep their posterior responses with the same value as their initial perceptions, i.e. $Prior_i = Post_i$. One might point this issue as an important downside of our study. Also, previous papers present data collection procedures specifically designed to assess subjective probabilities and its revision processes⁹ while our dataset comes from a broad household survey.

Clearly, individuals faced a critical question – homicide expectation – about which initial perceptions were too far from the truth. In addition, the information provided lasted less than a couple of minutes. However, even under such circumstances, if we proceed the estimation of equation (2) considering only those 127 individuals who changed initial responses, we work on a sample almost 40% greater than what is presented in Delavande (2008), for example.

By now, we want to extend the discussion on informational content raised by Viscusi and O'Connor (1984) to move from an extreme, where almost every participant revises initial perceptions, to the other, where the contrary occurs. Is that reasonable to keep considering an economy

⁹Delavande (2008) and Delavande et al. (2011) are good examples of this.

composed by agents who always believe in the information received for every risky decision to be made? Putting in other words, is all information provided about uncertain outcomes informative enough to make almost every individual change their initial perceptions? Our dataset sends a clear message that this is not the case, and restricting attention only to those updaters would lead us, at least, to waste a huge amount of information. Also, these statistics reported might be misleading, due to selection bias.

In this study, it is important not to neglect individuals who did not respond to new information, since they carry important clues about the updating process. Still, how to accommodate in the same model rational individuals, facing the same information, but behaving so differently? **At this point, it is clear for us that our venture and major contribution is to propose a framework compatible to a bayesian update allowing the possibility that no update takes place.** In this sense, we seek to develop an econometric model which is able to address the following issues: a) Agents are bayesian when updating their victimization expectations; b) Observable and non-observable heterogeneity must refer not only to the receiver (interviewee) but also to the sender (interviewer), since matching aspects may influence credibility and information usage; c) The presence of non updating individuals must receive an adequate treatment and be rationalized under a bayesian context.

With that in mind, rewrite (1) as follows:

$$\pi_i^p = \left(1 + \frac{\alpha_i + \beta_i}{N}\right)^{-1} \pi^* + \left(1 + \frac{N}{\alpha_i + \beta_i}\right)^{-1} \pi_i^0 \quad (3)$$

Note that, from π_i^0 , we may interpret $\alpha_i + \beta_i$ as the quantity of samples initially used to form $Prior_i$. Following this reasoning, $\eta_i = \frac{N}{\alpha_i + \beta_i} \in [0, \infty)$ may be interpreted as a measure of the informational quality. In this way, (3) can be rewritten as:

$$\pi_i^p = \left(1 + \frac{1}{\eta_i}\right)^{-1} \pi^* + \left(1 + \eta_i\right)^{-1} \pi_i^0 \quad (4)$$

This structural form will be crucial to skepticism rationalization. We need to explain individuals setting $\pi_i^p = \pi_i^0$, and it is reasonable to assume that the decision to revise perceptions, i.e to make $\pi_i^p \neq \pi_i^0$, is directly related to the credibility and quality of the information obtained. As (4) makes clear, when η_i approaches zero, or the more irrelevant the information is seen by the decision maker, π_i^p approaches π_i^0 , which, in the limit, translates into the complete non-existence of an update. In the other extreme, when η_i approaches the infinity, π_i^p takes exactly the value π_i^* .

In summary, individuals setting $Post_i = Prior_i$ tell us that informational content was either irrelevant or the source lacks credibility, or both, and there was no reason to change initial perceptions. Ignoring these responses introduces a selection bias since establishing no changing in perceptions is a rational decision based on an intrinsic optimization process, which might be dependent on observable and unobservable characteristics.

In this situation, it is hard to believe that the now restricted error term has a zero conditional mean and, even if the structural model is linear, OLS procedure leads to inconsistent parameter estimates. One possible approach to deal with these cases is to think of an unrestricted and unobservable *latent* variable underlining the true observations through a specific structure. That is exactly what we will develop next.

4.2.1 Model 1: A Generalized Tobit for Bayesian Updating

Initially, let Y be a random variable of interest and Y^* its latent counterpart. then,

$$\begin{aligned} Y_i^* &= \mathbf{X}_i \delta + u_i, & u_i | \mathbf{X}_i &\sim Normal(0, \sigma^2) \\ Y_i &= \max(0, Y_i^*) \end{aligned}$$

where \mathbf{X}_i is a vector of conditioning variables.

Adapting this Tobit Model for our purposes, we assume the existence of a latent variable $Post_i^*$ that will govern the censoring mechanism. We conjecture that $Post_i^*$ captures the respondents' latent *posterior* probability and, as such, according to equation (2), we propose it is a function of the *prior* and of a vector of observable variables containing both interviewees' and interviewers' characteristics, as well variables representing homophily. Hence, consider the following Tobit:

$$Post_i^* = \gamma Prior_i + \mathbf{X}_i \delta + u_i, \quad u_i | Prior_i, \mathbf{X}_i \sim Normal(0, \sigma_u^2) \quad (5a)$$

$$Post_i = \min(Prior_i, Post_i^*) \quad (5b)$$

But why the minimum? This is so because, in our sample, all initial perceptions are greater than the official rate. Thus, for the sake of rationality, we must have $Post_i \leq Prior_i \forall i$. Now, note that:

$$Post_i = Prior_i - \max(0, Prior_i - Post_i^*)$$

This means that our Tobit is equivalent to the following "new" Tobit:

$$Prior_i - Post_i^* = (1 - \gamma) Prior_i - \mathbf{X}_i \delta + v_i \quad (6a)$$

$$Prior_i - Post_i = \max(0, Prior_i - Post_i^*) \quad (6b)$$

where, as usual, we assume that $v_i | Prior_i, \mathbf{X}_i \sim Normal(0, \sigma_v^2)$.

Given our data, the way we defined equation (6a) would lead us to work with positive values and, in conjunction with equation (6b), the model just described is a well-known generalization for the original standard Tobit. Note that differently from the traditional approach, the threshold is not constant along observations but vary in a deterministic way, say, $Prior_i = h(\mathbf{X}_i)$. Actually, the form of $h(\mathbf{X}_i)$ does not need to be specified, since we build our model conditional on $Prior_i$.

Finally, we just need to run a Tobit where the dependent variable is $(Prior_i - Post_i) \geq 0$ and the vector of regressors is $(Prior_i, \mathbf{X}_i)$. The estimated parameters are $(1 - \hat{\gamma}, -\hat{\delta})$. However, we need $(\hat{\gamma}, \hat{\delta})$ and interpret these results by means of equation (5a). Now refer to equations (5a) and (5b) and the interpretation is as follows:

i) Interpreting $\hat{\gamma}$:

The fundamental result to obtain with respect to $\hat{\gamma}$, attesting the validity of a bayesian update model, is its statistical significance. In other words, this tells us that *Prior* influences $Post^*$ and, thus, *Post*, i.e. we have a bayesian effect in the sense of Viscusi and O'Connor (1984).

If $\hat{\gamma} > 1$, then agents are **SKEPTICAL**. An incremental change in $Prior_i$ induces a higher incremental change in $Post_i^*$. Hence, since equation (5b) is the minimum between them, we have $Post_i = Prior_i$, i.e., the higher the prior, the less likely to update (skepticism).

If $\hat{\gamma} < 1$, then agents are **UPDATERS**. An increase in $Prior_i$ induces a less than proportional increase in $Post_i^*$. Therefore, we are looking at the updaters, and $Post_i < Prior_i$, i.e., the higher the prior, the more likely to update.

ii) Interpreting $\hat{\delta}$:

If $\hat{\delta} > 0$, then an increase in \mathbf{X}_i implies a higher $Post_i^*$ and we observe the same effect that $\hat{\gamma} > 1$. Analogously, if $\hat{\delta} < 0$, then lower \mathbf{X}_i implies lower $Post_i^*$. Therefore, it is more likely that we observe an update.

In order to build the original Tobit likelihood function, we need the following expressions first:

$$P(Post_i = Prior_i | Prior_i, \mathbf{X}_i) = 1 - \Phi\left(\frac{(\gamma - 1)Prior_i + \mathbf{X}_i\delta}{\sigma_u}\right) \quad (7a)$$

$$P(Post_i = Post_i^* | Prior_i, \mathbf{X}_i) = \Phi\left(\frac{(\gamma - 1)Prior_i + \mathbf{X}_i\delta}{\sigma_u}\right) \quad (7b)$$

As well, we need the density $f(Post_i | Prior_i, \mathbf{X}_i, Post_i^* < Prior_i)$. But this is just:

$$g(Post_i^* | Prior_i, \mathbf{X}_i, Update_i^* > Prior_i)$$

$$f(Post_i | Prior_i, \mathbf{X}_i, Post_i > Prior_i) = \frac{\phi\left(\frac{(Post_i - \gamma Prior_i - \mathbf{X}_i\delta)}{\sigma_u}\right)}{\sigma_u \left(1 - \Phi\left(\frac{(Post_i - \gamma Prior_i - \mathbf{X}_i\delta)}{\sigma_u}\right)\right)} \quad (8)$$

Defining the indicator function by $\mathbf{1}(\cdot)$, the likelihood contribution of a given $Post_i | (Prior_i, \mathbf{X}_i)$ observation is:

$$\left[1 - \Phi\left(\frac{(\gamma - 1)Prior_i + \mathbf{X}_i\delta}{\sigma_u}\right)\right]^{\mathbf{1}(A)} \times \left[\frac{\phi\left(\frac{(Post_i - \gamma Prior_i - \mathbf{X}_i\delta)}{\sigma_u}\right)}{\sigma_u \left(1 - \Phi\left(\frac{(Post_i - \gamma Prior_i - \mathbf{X}_i\delta)}{\sigma_u}\right)\right)}\right]^{\mathbf{1}(B)}$$

where $A : Post_i = Prior_i$ and $B : Post_i > Prior_i$.

4.2.2 Model 2: A Simple Two-Tiered Hurdle Model

An alternative modeling comes from the fact that we have a two-step structure. As mentioned before, individuals decide whether or not change their *priors* and then they give a *posterior* perception. Hence, *Discrete Choice Models* such as *Logit* and *Probit* followed by OLS on the updating subsample emerge as a natural approach. Thus, an alternative to the Tobit model is to dispense with the abstraction of a latent variable $Post_i^*$ and model only the observed empirical evidence directly. If we define $\Delta_i^{update} \equiv -Update = Prior_i - Post_i$ as the actual update to deal only with non-negative numbers¹⁰, we have:

$$P(\Delta_i^{update} = 0 | Prior_i, \mathbf{X}_i) = 1 - \Phi(\gamma Prior_i + \mathbf{X}_i\delta) \quad (9a)$$

$$G(Post_i) | (Prior_i, \mathbf{X}_i, \Delta_i^{update} > 0) \sim Normal(\tilde{\gamma} Prior_i + \mathbf{X}_i\tilde{\delta}, \sigma^2) \quad (9b)$$

Equation (9a) dictates the probability that a real update has not occurred, i.e., $\Delta_i^{update} = 0$ and equation (9b) is a conditional model (conditioned on $\Delta_i^{update} > 0$) for the posterior probability. Equation (9b) just asserts that after a suitable transformation $G(Post_i)$, usually but not necessarily $G(Post_i) = \log(Post_i)$, we achieve normality. There are some features about the model just described:

- i) We could have different explanatory variables for the censoring equation (9a) and equation (9b), as well as estimated parameters for the same explanatory can vary between equations;
- ii) Estimation of $(\gamma, \delta, \tilde{\gamma}, \tilde{\delta}, \sigma^2)$ is quite simple, say: for equation (9a), run a probit; and for equation (9b) run a simple OLS for the subsample where $\Delta_i^{update} > 0$;

¹⁰Remember that in our data set updaters have $Prior_i \neq Post_i, \forall i$.

Estimation of partial effects demands knowledge of $G(\cdot)$. However, if we assume log-normality, $G(Post_i) = \log(Post_i)$, nice results emerge:

$$E(Post_i | Prior_i, \mathbf{X}_i, \Delta_i^{update} > 0) = \exp\left(\tilde{\gamma}Prior_i + \mathbf{X}_i\tilde{\delta} + \frac{\sigma^2}{2}\right)$$

$$E(Post_i | Prior_i, \mathbf{X}_i) = \Phi(\gamma Prior_i + \mathbf{X}_i\delta) \exp\left(\gamma Prior_i + \mathbf{X}_i\delta + \frac{\sigma^2}{2}\right)$$

An import drawback of the hurdle model is the impossibility of testing it against the Tobit specification, at least by simple procedures. On the other hand, a Vuong-type test could be applied to those non-nested hypotheses.

5 RESULTS

5.1 Bayesian Update with Skepticism: Model 1

For convenience, equations (5a), (5b), (6a) and (6b) are shown, again, below:

$$Post_i^* = \gamma Prior_i + \mathbf{X}_i\delta + u_i \quad (5a)$$

$$Post_i = \min(Prior_i, Post_i^*) \quad (5b)$$

$$Prior_i - Post_i^* = (1 - \gamma)Prior_i - \mathbf{X}_i\delta + v_i \quad (6a)$$

$$Prior_i - Post_i = \max(0, Prior_i - Post_i^*) \quad (6b)$$

Table (4) presents the estimation results. As explained in subsection (4.2.1), when referring to the latent equations (5a) and (5b), we see the likelihood of an update. This is the first decision step going on. Keep in mind that, for this part, we are interested in the coefficients of equation (5a) but, in fact, we estimate equation (6a). Thus, we need to proceed an adjustment on coefficients interpretation. For the $Prior_i$ coefficient, our interest relies on γ itself, but we would get an estimative of $(1 - \gamma)$; therefore, it is easy to see that we must subtract it from 1. For all of the other coefficients, represented by δ , simply change estimatives' signs and, then, interpretation follows usual reasoning.

The first result we must obtain is the statistical significance of the $Prior$ coefficient in equation (5a). Table (4) shows that $\hat{\gamma} = (1 - 0,5021) = 0.498$ and we do have a bayesian effect. Also, note that $\hat{\gamma} < 1$, which implies that high $Prior$'s drive an update decision, a quite reasonable result. Remember that the official homicide rate is less than 1%, hence the higher initial responses are set, individuals are making higher mistakes, and the informational shock is strong enough to produce a change in perceptions.

The story behind \mathbf{X} variables is straightforward. *Age* has an estimated coefficient equal to -0.728 , i.e. $\hat{\delta}_{Age} = 0.728$. Thus, respondents' age works undoubtedly against the update: older individuals are less likely to revise initial perceptions. The same direction is found in *Gender*. Being a female drops the likelihood of a change in perceptions. Perhaps both results tell us that vulnerability faced by older citizens and women make them more reluctant to update victimization expectations.

These results are not surprisingly, though. Although fear of crime is a concern for people of all genders, studies consistently find that women around the world tend to have much higher levels of fear of crime than men (see among many, Mark (1984), Ferraro (1995b), Ferraro (1996), Tulloch (2000), and Snedker (2015)). This is so, despite the fact that for most offenses, men's actual victimization rates are higher. As to the age effect, we are in accordance to many studies that found a monotonically increasing or, at least, an "U-shaped" relationship between fear of crime and victimization (see, Mark (1984), and Tulloch (2000)).

Table 4: Tobit Model

	<i>Dependent variable:</i>
	(Prior - Post)
Prior	0.502*** (0.129)
Age	-0.728*** (0.241)
Sex	-18.036** (7.684)
Matching_Sex	63.480*** (13.931)
Educ_Int	19.307* (11.459)
Constant	-232.149*** (39.056)
Observations	2,828
Log Likelihood	-1,017.116
Wald Test	41.278*** (df = 5)

*p<0.1; **p<0.05; ***p<0.01

Source: Elaborated by the author.

Despite much discussion on the importance of individuals' cognitive capability on probabilistic reasoning, our results show that education is not relevant for this task. In fact, controlling for participants' level of education brought us insignificant coefficients. We argue that above a given threshold, ensuring a basic understanding of our probability explanation, the interviewer's capacity to communicate is crucial to persuade an update. This explains a 19.307, i.e., $\hat{\delta}_{Educ_Int} = -19.307$ coefficient for the educational level of the informational sender. On average, graduates or undergraduates conducting our interview increase the likelihood of an update due to, we argue, a higher informational quality inherent to more educated interviewers.

Finally, one of the most important variable in our model is *Matching_Gender*. Controlling for individuals' gender, if the interaction is composed by the same gender, i.e. Male/Male or Female/Female, the update is more likely to occur¹¹. Clearly, due to reasons out of our scope in this study, the informational content depends heavily on gender correspondence.

5.2 Bayesian Update with Skepticism: Model 2

As proposed in subsection (4.2.2), we have a two-tiered model to deal with both updaters and skeptical respondents. The first stage explains the changing decision with a *Logit* and a *Probit* model followed by OLS on the updating subsample in the second stage. Unlike the previous subsection, the estimation results are interpreted directly. Table (5) and Table (6) present our results for the first and second stages, respectively.

Starting for the changing decision (Table (5)), it is important to highlight that our model is robust to the *Logit* or *Probit* choice. All explanatory variables are significant in the same magnitude with the same signs for both of them. Also, the Log Likelihood and the Akaike Information Criterion are almost the same. Although we can not say much on the coefficients magnitude, the signs are informative, and lead us to the conclusions made in the previous section. However, we must test our model adjustment.

So, we performed a simple exercise (results can be obtained under request) to assess the prediction

¹¹Due to the *ceteris paribus* assumption, we can not keep the *Gender* variable in the model and identify at the same time specific matchings such as Male/Female, Female/Male and so on.

power of our models, where the variable *Predicted* is just our *Logit* and *Probit* probabilities estimation. We assume there is a threshold set at 50% such that if $Predicted \geq 50\%$ and $Change = 1$ or if $Predicted < 50\%$ and $Change = 0$, then we have a correct predicted choice. Finally, the variable *Correct* equals one if we had success in predicting and zero on the contrary. Therefore, *Correct* mean equals 0.957 implies that we had a 95.7% success rate.

Table 5: Changing Decision and Marginal Effects

	<i>Coefficients:</i>		<i>Marginal effects:</i>	
	Change (<i>logistic</i>)	Change (<i>probit</i>)	Marginal effects (<i>logistic</i>)	Marginal effects (<i>probit</i>)
Prior	0.007** (0.003)	0.003** (0.002)	0.0003** (0.0001)	0.0003** (0.0001)
Age	-0.020*** (0.007)	-0.009*** (0.003)	-0.001*** (0.0003)	-0.001*** (0.0002)
Gender	-0.498** (0.197)	-0.238** (0.094)	-0.020** (0.008)	-0.021** (0.008)
Matching_Gender	1.876*** (0.410)	0.804*** (0.162)	0.075*** (0.014)	0.070*** (0.013)
Educ_Int	0.654* (0.337)	0.261* (0.141)	0.026** (0.013)	0.023* (0.012)
Constant	-5.879*** (1.056)	-2.813*** (0.436)	-0.236*** (0.040)	-0.244*** (0.037)
Observations	2,829	2,829		
Log Likelihood	-473.495	-473.564		
Akaike Inf. Crit.	958.990	959.129		

*p<0.1; **p<0.05; ***p<0.01

Source: Elaborated by the author.

Given our good adjustment, we proceed to interpret its coefficients. Table (5) presents the marginal effects of our explanatory variables over the changing decision probability. With respect to the *Prior*, in accordance with our previous model, initial perceptions have a positive influence on the probability that an update occurs. Higher initial perceptions imply higher mistakes, hence, once more, this willingness to change can be taken as a rational decision.

Also, for all of the other independent variables, the results lead to the same previous conclusions. In *Age*, as it was the case before, older participants are more reluctant to change and, again, perhaps insecurity drives this results. A stronger evidence of this explanation comes from *Gender*. Being female reduces the changing probability in 2 percentage points.

As expected, educated informational senders had a positive impact on the changing decision. Here, clearness or even social aspects of education in a low-educated society might influence the information quality and credibility. Once again, homophily analysis arises. Being of the same gender as the interviewer is clearly the most important aspect in our data. Male respondents might consider reliable only information provided by male interviewers. On the other hand, female participants might feel more comfortable being interviewed by women. Many explanations are possible but, under this cultural context, it is intuitive that gender matchings reduce information noise and this is key to the informational content.

Once having initial perceptions changed, now we proceed to the updating analysis, see Table 6. We have two key points to emphasize with this exercise: in line with our Tobit model and with Viscusi and O'Connor (1984), the *Prior* coefficient in a linear model is significant. This is a crucial evidence supporting our Bayesian approach both for the general and restricting cases. *Age* and *Gender* have exactly the same role as our previous analysis and there is no need to say anything else. In the same way, the interviewers' level of education has again an important role explaining posterior responses. As expected, since all *Post*'s are less than *Prior*'s, and these initial perceptions are more than 1000

Table 6: Restricted Ols

<i>Dependent variable:</i>	
Post	
Prior	0.023*** (0.008)
Age	0.026* (0.015)
Gender	0.988** (0.455)
Matching_Gender	1.074 (1.179)
Educ_Int	-4.188*** (0.975)
Police	0.494* (0.273)
Constant	1.697 (1.369)
Observations	
	121
Adjusted R ²	
	0.254
Residual Std. Error	
	2.375 (df = 114)
F Statistic	
	7.822*** (df = 6; 114)

*p<0.1; **p<0.05; ***p<0.01

Source: Elaborated by the author.

times greater than the true value, more educated interviewers move participants closer to the truth. This explains a negative value for *Educ_Int* (-4.188).

The insignificance of *Matching_Gender* is new here. However, remember from table (3) that 93% of our updating subsample is composed by interactions of the same gender. Hence, perhaps we do not have variations enough to obtain a significant coefficient which means that this pre-evidence is just what we are looking for.

Also, given the flexibility of the two-tiered modeling, we could introduce another independent variable: *Police*. Its positive coefficient is just what one would expect. Higher levels of *Police* values indeed mean a worse evaluation. Therefore, it leads us to an intuitive result: posterior perceptions are higher or, in other words, the movement towards the truth is less strong, when police work is worse seen by respondents.

6 FINAL CONSIDERATIONS

This paper was built under three major guidances: i) Individuals are rational decision makers when updating subjective perceptions; ii) There is an estimable linear regression model for the bayesian update process, a well-suited framework to deal with the revision of subjective perceptions; iii) There are other variables besides respondents' characteristics influencing this updating process.

With these ideas, we presented an initial sample of 4,030 individuals regarding subjective risk perceptions about becoming a homicide victim for the following 12 months in Fortaleza, Brazil. Besides having a much larger sample than those presented in many papers in the field, our data brought information about interviewers and matching aspects that were used to account for information quality. In addition, we had a very different data sample: 95% of respondents did not want to change their *Prior*'s, setting $Prior_i = Post_i$. This made the non-updaters, or skeptical agents, our protagonists.

We showed that, under simple assumptions, Viscusi and O'Connor (1984) and Smith and Johnson

(1988) could be extended at the same time. We proposed a multiple linear regression model in the context of a bayesian update approach using a vector of independent variables for respondents, interviewers and matching in gender. However, our empirical evidence imposed the development of a model able to accommodate both updaters and non-updaters.

Following this guidance, we proposed two alternative estimable models suited to a more general context than what is found in the literature. Indeed, we believe this generalization to be a simple, yet important, achievement of this study. A modified Tobit, easily implemented in any statistical package, was developed as the first approach. The second model was a well-known two-tiered Hurdle model, allowing the possibility to use a different set of variables to explain the changing decision and the update.

Firstly, our results showed that we could proceed with a bayesian update approach, since our *Prior* coefficient was significant for every model used. More specifically, referring to our Tobit model, the initial response coefficient was $\hat{\gamma} = (1 - 0.5021) = 0.498$, and it was statistically significant. It permitted us to conclude that *Prior's* induce an update through a latent variable. In other words, higher initial perceptions imply higher misleading perceptions and this leads to a change in responses. Despite the fact that we are using a different framework, our results stay in accordance with that of Viscusi and O'Connor (1984), i.e. there is a statistically significant linear equation relating *priors* to *posteriors*.

Also, fundamentally, we could rationalize a non-updating decision following a perceived informational quality/credibility argument. This was made through respondents' age, gender and initial perceptions, as well as interviewers' level of education and matching in gender between members of the interaction.

We found that older participants and females are more reluctant not only to change initial responses, but also to choose the level of the new response, in case of an update. We argued that insecurity aspects might be used to explain these findings. Also, respondents' level of education was insignificant in our exercise. In fact, interviewers' level of education plays a key role in both the changing and updating magnitude decisions. This is a consistent piece of evidence supporting our main line of explanation: the perceived informational content. The relation between sender's education and information quality is straightforward.

Finally, our results also raised important evidence on homophily aspects. In almost every regression, *Matching_Gender* had a major impact on the decision to change and in the magnitude of the update. The only insignificant one was in a simple OLS on the updating subsample. However, in this group, 93% of the interactions happened under the same gender. It suggests that Latin cultures, as the Brazilian, do not pose weight on race questions, as it is the case for other cultures. However, it has an analogue restriction due, perhaps, to its characteristic machismo: gender.

Further developments should be done in many aspects of this study. Firstly, the design of our interaction was clearly not ideal and the informational shock could be better customized to adjust official rates to different socioeconomic characteristics. Although it raised a remarkable opportunity to propose a generalization, our estimation procedure can be improved in several ways. One of particular importance is to allow different revision directions, since our data and modeling present just the downwards update. Also, we could have more information regarding interviewers. Further study on homophily is another gap to be filled in.

References

- Becker, C. (1968). Punishment: An economic approach, 76j. *Pol. Econ*, 169(10.2307):1830482169.
- Bott, H. (1928). Observation of play activities in a nursery school. *Genetic Psychology Monographs*, 4(1):44-88.

- Bursik Jr, R. J., Grasmick, H. G., et al. (1999). *Neighborhoods & crime*. Lexington Books.
- Delavande, A. (2008). Measuring revisions to subjective expectations. *Journal of Risk and Uncertainty*, 36(1):43–82.
- Delavande, A., Giné, X., and McKenzie, D. (2011). Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics*, 94(2):151–163.
- Dominitz, J. and Manski, C. F. (1996). Perceptions of economic insecurity: Evidence from the survey of economic expectations. Technical report, National bureau of economic research.
- Dominitz, J. and Manski, C. F. (1997). Using expectations data to study subjective income expectations. *Journal of the American Statistical Association*, 92(439):855–867.
- Dominitz, J. and Manski, C. F. (2004). How should we measure consumer confidence? *Journal of Economic Perspectives*, pages 51–66.
- Dominitz, J. and Manski, C. F. (2011). Measuring and interpreting expectations of equity returns. *Journal of Applied Econometrics*, 26(3):352–370.
- Dominitz, J., Manski, C. F., and Heinz, J. (2002). Social security expectations and retirement savings decisions. Technical report, National Bureau of Economic Research.
- El-Gamal, M. A. and Grether, D. M. (1995). Are people bayesian? uncovering behavioral strategies. *Journal of the American statistical Association*, 90(432):1137–1145.
- Ferraro, K. F. (1995a). *Fear of Crime: Interpreting Victimization Risk*. State University of New York Press.
- Ferraro, K. F. (1995b). *Fear of crime: Interpreting victimization risk*. SUNY press.
- Ferraro, K. F. (1996). Women’s fear of victimization: Shadow of sexual assault? *Social forces*, pages 667–690.
- Ferrell, W. R. and McGoey, P. J. (1980). A model of calibration for subjective probabilities. *Organizational Behavior and Human Performance*, 26(1):32–53.
- Golub, B. and Jackson, M. O. (2012). Network structure and the speed of learning measuring homophily based on its consequences. *Annals of Economics and Statistics/ANNALES D’ÉCONOMIE ET DE STATISTIQUE*, pages 33–48.
- Graham, B. (2011). Econometric methods for the analysis of assignment problems in the presence of complementarity and social spillovers. *Handbook of social economics*, 1:965–1052.
- Hart, A. G., Modigliani, F., and Orcutt, G. H. (1960). The quality and economic significance of anticipations data.
- Hoffrage, U., Lindsey, S., Hertwig, R., and Gigerenzer, G. (2000). Communicating statistical information. *Science*, 290(5500):2261–2262.
- Hurd, M. D. and McGarry, K. (1995). Evaluation of the subjective probabilities of survival in the health and retirement study. *Journal of Human resources*, pages S268–S292.
- Jackson, M. O. (2009). Social structure, segregation, and economic behavior. *Segregation, and Economic Behavior (February 5, 2009)*.

- Juster, F. T. (1966). Consumer buying intentions and purchase probability: An experiment in survey design. *Journal of the American Statistical Association*, 61(315):658–696.
- Koriat, A., Lichtenstein, S., and Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human learning and memory*, 6(2):107.
- Kunreuther, H., Ginsberg, R., Miller, L., Sagi, P., Slovic, P., Borkan, B., and Katz, N. (1978). *Disaster insurance protection: Public policy lessons*. Wiley New York.
- Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., and Combs, B. (1978). Judged frequency of lethal events. *Journal of experimental psychology: Human learning and memory*, 4(6):551.
- Machlup, F. (1946). Marginal analysis and empirical research. *The American Economic Review*, pages 519–554.
- Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5):1329–1376.
- Mark, W. (1984). Fear of victimization: Why are women and the elderly more afraid? *Social science quarterly*, 65(3):681.
- McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444.
- Pesaran, M. H. (1987). *The limits to rational expectations*. Blackwell Oxford.
- Samuelson, P. A. (1938). A note on the pure theory of consumer’s behaviour. *Economica*, pages 61–71.
- Smith, V. K. and Johnson, F. R. (1988). How do risk perceptions respond to information? the case of radon. *The Review of Economics and Statistics*, 70(1):1–8.
- Snedker, K. A. (2015). Neighborhood conditions and fear of crime a reconsideration of sex differences. *Crime & Delinquency*, 61(1):45–70.
- Tulloch, M. (2000). The meaning of age differences in the fear of crime: Combining quantitative and qualitative approaches. *The British Journal of Criminology*, pages 451–467.
- Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157):1124–1131.
- United Nations (2013). *Global Study on Homicide 2013: Trends, Contexts, Data*.
- Vanderveen, G. (2006). *Interpreting Fear, Crime, Risk and Unsafety: Conceptualisation and Measurement*. Boom Juridische Uitgevers.
- Viscusi, W. K. (1979). *Employment hazards: An investigation of market performance*. Number 148. Harvard University Press.
- Viscusi, W. K. (1985). Are individuals bayesian decision makers? *The American Economic Review*, pages 381–385.
- Viscusi, W. K. and O’Connor, C. J. (1984). Adaptive responses to chemical labeling: Are workers bayesian decision makers? *American Economic Review*, 74(5):942–956.
- Warr, M. (2000). Fear of crime in the united states: Avenues for research and policy. Technical report, National Institute of Justice/NCJRS, Rockville, MD - USA.