Essays on (small) crime

Douhou, S.

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Salima Douhou

Essays on (Small) Crime: Perception, Social Norms, Happiness, and Prevention

Essays on (Small) Crime: Perception, Social Norms, Happiness, and Prevention

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University, op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in zaal DZ1 van de Universiteit op vrijdag 8 juni 2012 om 14.15 uur door

Salima Douhou

geboren op 23 december 1983 te IJsselstein.

Promotiecommissie

Promotores:	prof. dr. J.R. Magnus
	prof. dr. A.H.O. van Soest
Overige Leden:	prof. dr. J.W.M. Das
	dr. O. Marie
	dr. B.A. Vollaard
	prof. dr. J.K. Winter

Essays on (Small) Crime: Perception, Social Norms, Happiness, and Prevention

ISBN 978 90 5668 316 0

To Rachida

Preface

> If it wasn't hard, everyone would do it.
> It's the hard that makes it great.
>
> Tom Hanks

I would like to express my gratitude to the people and organization that made it possible to write this thesis. In 2007 I was awarded a Ph.D. grant by the Netherlands Scientific Organization for Research (NWO), which means that this dissertation is supported by NWO (under project number 017.004.018).

I would like to thank Jan Magnus, one of my supervisors, for his continuous effort, patience, and critical remarks that learned me to always check, test, and carefully consider any output produced. I am also grateful to Arthur van Soest, my other supervisor. Although he is a very modest person, his knowledge of (applied) econometrics is phenomenal and I feel privileged that I had the chance to benefit from this.

This is a natural moment to thank my committee members as well: Marcel Das, Olivier Marie, Ben Vollaard, and Joachim Winter. I am privileged to have them in my committee and I would like to thank them for their extensive feedback during my pre-defense and for taking the time to read the chapters. During my last year as a Ph.D. student I got the opportunity to work at CentERdata for several months, where I met the director and member of my committee, Marcel. It was a pleasure to work together with Jan Nelissen en Peter Fontein during these months. Since the topics of my thesis are more related to the Economics department I joined their presentations and workshops on several occasions. Here I met Olivier Marie and Ben Vollaard, who are two of the few people that work on the economics of crime. I was working on the basis of Chapter 3 and 4 in the period I met Olivier and Ben and both have had an important influence on how to approach and present the topics I was working on. In January 2011 I visited the Ludwig-Maximilians-Universität in Munich for a month. I would like to thank my host and committee member, Joachim Winter, for his hospitality, time, and great lunches we had in Munich.

The majority of the chapters in this thesis are the result of a survey conducted by CentERdata. I am grateful to Corrie Vis for her time and effort to answer my questions and provide me with some of the data that is used in this thesis.

I can fairly say that I spent most of my time as a Ph.D. student in room K514, which I shared with John(ny). I would like to thank him for being a good audience for my monologues, complaints, and bad jokes. Your apologies for humming, tapping, clapping, and singing along with the music (while you were wearing ear plugs) are accepted. During my stay in Munich I had Gregor Tannhof as my roommate, who made me feel at home. Gregor has this list of 21 advices that apply for any occasion. Some examples are (in German): Sich im Internet über Alternativen informieren; den Cache leeren; keine Angst ziegen (können sie riechen!); in kreisende Bewegungen arbeiten. Thanks Gregor, they proved to be very useful! Although Mohammed was not officially my roommate he always seems to produce enough noise be considered one. Thank you for the funny anecdotes, your performances in the corridor, and being a good friend.

When I was not too busy I always tried to join the subdepartmental lunches in the mensa. Although the composition and the turnout of the

lunch group differed quite a lot it was always nice to be part of this group and sometimes be involved in discussions about nothing. Members of this group include(d) a.o.: Edwin D., Edwin L., Elleke, Gerwald, Gijs, Hans, Henk, Herbert, Jacob, John, Marloes, Marieke, Mirjam, Peter, René, Ruud, Soesja, and Willem. I had the pleasure to work with Marieke, Henk, and Hans on several courses that I have taught during my years in Tilburg. A special thanks goes to Marieke for investing many hours in me to improve my teaching skills. Marloes and Elleke made sporting an even more fun activity: we have spent many hours at yoga and in the swimming pool, which helped to relax a bit. Soesja and Mirjam were always there for me when I felt like chatting, complaining, or just drinking tea. Once in a while Edwin L. loudly joined/interfered in these conversations.

I learned that secretaries, especially those at the department of Econometrics and Operations Research, can make life much easier. I would like to thank Anja, Annemiek, Heidi, and Korine for their effort.

Some people with which I shared an occasional tea, cookie, chat, sports, dinner, etc. are: Barış, Beatrice, Bettina, Chris, Cristian, Gaia, Jan S., Jarda, Jiehui, Marta, Martin K., Martin S., Michele, Miguel, Nathanael, Patrick, Pedro, Ramon, Rasa, Renata, Romeo, Sara, Stefan, Ting, Tunga, Unnati, Verena, the Vonkovi, and Ying. A special mention for the several dinners and drinks I had with Peter O., Gerard, and Jenny. Thanks!

Nabil and Nasima, I would like to thank you for agreeing to be my paranymphs. It is comforting to know that a mechanical engineer/mathematician and an economist will be standing behind me during my official defense. Special thanks go to my mother, Rachida, for being a great support. This list is not complete without Adil, Karim, Gerda, and Robert. Finally, I am greatly indebted to the support and stability my husband, Rob, offered me during my life as a Ph.D. student.

Salima Douhou
Tilburg/Sint-Oedenrode, April 2012

Contents

Chapter 1

Introduction

Let's look at crime from an economic perspective. Crime is a good (or, a bad as some may argue) and there exists a market for this good. On the supply side are the producers of crime (called offenders) and on the demand side are the consumers of illegal goods and services (see, for example, Freeman, 1999). The activities of this market have strong spillover effects to the market for anti-crime goods and services (or, protection) since a higher supply of crime induces a higher demand for protection.

The supply side of the market is modeled in a path-breaking article of Becker (1968). How much offenses an offender commits is a function of incentives that result from (implicit) pricing of punishment via severity and probability of crime and 'other activities', which include income from legal work. This approach assumes that 'a person commits an offense if the expected utility to him exceeds the utility he could get by using his time and other resources at other activities' (Becker, 1968, p. 176). Ehrlich (1973) refines this by explicitly modeling both costs and benefits from criminal and non-criminal activities and analyzes the interaction between these types of activities. This research has spurred more interest in the economic analysis of crime. Further attempts to model the supply side of the market have mainly focussed on dealing with heterogeneity across offenders in terms of risk or earning opportunities with respect to legitimate and illegitimate activities. The relation between crime rate and different crime types and rates of urbanization have also been important contributions in the post-Becker era (see Glaeser et al., 1996; Sah, 1991).

The demand side of the market for anti-crime goods shows both a public and private demand. Potential victims have an incentive to protect themselves against victimization or its consequences, for example by means of locks, security systems, and market insurance. A government that represents the interest of the public protects her people via law enforcement. This includes arresting, prosecuting, convicting offenders, and managing penalties. Philipson and Posner (1996) show that potential victims take more protective measures in response to higher crime rates and Becker discusses a socially optimal level of enforcement by a government in his theoretical model.

Empirical research on the economics of crime started out by estimating the cost of crime. In the early days this was done by means of data from government sources only (Czabański, 2008). This restricted the estimation to specific regions and crime categories and reported crimes. Nowadays, statistics on criminal activities rely more on general surveys that include modules of questions on crimes and victimization surveys. This is considered to give more reliable data than registered crime records.[1] The empirical literature on the (social) cost of crime aims at quantifying the size of the total market (or, for submarkets if one considers studies on specific types of crime). Empirical studies on the supply side of the market, in very general terms, look at incentives for criminal behavior such as the effect of probability and severity of punishment on crime and benefits and cost of legal versus illegal activities. More specifically, incentives from the environment that influence the decision making of potential offenders are important inputs for empirical studies on the economics of crime. How preferences for criminal activities may vary between groups seems to have less priority (Eide, 2001).

[1] The reliability of reported crime as a basis for criminal statistics is still a point of discussion. Myers (1982), for example, does not find much difference between the use of recorded crimes and victimization data.

Research on the demand for protection is mainly concerned with the effectiveness of preventive actions by both the (local) government and individuals in reducing (the likelihood of) victimization. Studies typically focus on a specific preventive action regarding (a specific type of) crime. However, research on the economics of crime is not limited to the investigation of the market of crime. A related issue concerns perception of seriousness of crime, which includes studies on the relation between crime and fear of crime, mental health, and relative importance of different crime types. As Becker notes, crime is a major business in terms of costs it puts on society. If we were able to get a better understanding of this issue it would allow (local) governments to better allocate resources for deterrence and punishment to those instruments of deterrence and punishment that lead to less crime for less money. The general understanding is that exposure to crime has a detrimental influence on perception. Perception implies the perception on the probability of being victimized (fear of crime) and perception on the seriousness of crimes. This, in turn, may have important consequences for mental health and preventive measures one may take. As a result, individuals may exert more pressure on governmental bodies that have budgetary power to allocate more resources to protecting citizens and they may take private action by spending more of their disposable income to preventive measures.

This thesis consists of four chapters that deal with different topics within the domain of crime. In broad terms, it deals with the following how's: how to prevent crime? (Chapter 2), how does one perceive crime? (Chapter 3), and how does it influence decision making and well-being? (Chapters 4 and 5). Hence, we will only look at the demand side of the anti-crime market. Part of this thesis contributes to the discussion on the relevance of social norms for economic issues (see chapters 3, 4, and 5). Normative concerns play an important role in defining crime but also contribute to the understanding of people's preferences, which in turn guide their courses of action as they serve as a 'motivational mechanism' (Elster, 1989, p. 102). Admittedly, socio-demographic characteristics such as age and gender also measure differences in norms but there is a need to make social norms an integral part of economic analysis, considering its great potential for the understanding of (ir)rational decision making. The interested reader is referred to the article by Jon Elster (1989) for a nice discussion on the relevance of social norms in economics. Social norms are present but not widespread in the economic literature. Some recent examples are studies on social capital (Bjørnskov, 2006; Helliwell, 2006), conditional norms (Traxler and Winter, 2012), and norm enforcement (Kube and Traxler, 2011).

Overview of the four essays

This thesis starts out in Chapter 2 with a study on the use of human typing behavior, called *keystroke dynamics*, as a means of authentication. The underlying idea is that every human being has a unique pattern or rhythm in which she types a known combination of letters or numbers, or both. This idea has been around since World War II during which telegraph operators could recognize a sending operator by means of her unique typing behavior, called 'fist'. Typing behavior is a branch of behavioral biometrics, which entails that a person can be identified by means of characteristic traits. Other examples of behavioral biometrics include handwriting, voice recognition, and gestures. In essence, keystroke dynamics is a crime prevention mechanism as it is an instrument to distinguish an authorized user from a non-authorized user of, for example, an e-mail address. Although it is a time-consuming operation to implement it has many advantages as a security instrument for network access. Two advantages are mentioned. First, it is user-friendly since no extra action is necessary than what a user is used to. Second, measurements can be adapted to changing behavior, which allows for constant refinement of keystroke dynamics related to a user. We use statistical analysis to see whether keystroke dynamics are sufficiently reliable as a security measure. More specifically, we develop a statistical test and use it to calculate the power —the probability that access to a non-authorized user is denied— and the size —the probability that access to an authorized user is denied— when we use keystroke dynamics as a security instrument. We contribute to the literature on keystroke dynamics by proposing a new statistical test and drawing on a dataset with more than 1000 participants, while most studies base their results on a much smaller number of observations. This data come from an experiment conducted by a group of students of the Systems and Network Engineering Group of the University of Amsterdam in 2007. We use two instruments to measure typing behavior: (i) *dwell time* records the time a key is held pressed and (ii) *flight time* is the time it takes a user to move between two consecutive keys. The experimental design is such that every participant is asked to login twenty times into a fictive network environment, using the same username and password. We find that dwell times are more powerful in distinguishing a user from a hacker but that typing behavior is only sufficient as a verification tool, not for identification.The chapters 3, 4, and 5 are based on a survey —entitled *Incorrect behavior in everyday life*— we have fielded in 2008 among participants of CentERpanel, which is managed by CentERdata. The survey participants form a representative sample of the Dutch population aged 16 years and older. As participants in CentERpanel typically take part in multiple surveys we have access to detailed respondent characteristics. The survey consists of three blocks. In the first block, respondents are asked to rate their perceived severity and justifiability on an ordered scale of 24 activities that may be considered to be more or less incorrect. Some examples are taking a bundle of printing paper from the office for private use, littering in a public place, and accepting a bribe. For the second block respondents were asked to answer vignette questions on several incorrect behaviors —called small crimes— that were selected from the first block. The respondents were asked to rate their perceived justifiability of the crimes on an ordered scale. We allowed for variation in offender (e.g., gender and age) and offense (e.g., behavior of superior and probability of getting caught) characteristics in the vignettes. Additional respondent characteristics such as past victimization were gathered in the last block of questions.

A vignette is a short story about hypothetical characters in specified circumstances to which an individual is asked to respond (Finch, 1987). The use of vignettes has the advantage that everyone can respond to it; a person does not need to have been in a particular situation to be able to respond. It also allows a researcher to extract information from a respondent for phenomena for which it is very difficult to find the exact question wording such as beliefs, norms, and perception. More generally, we make use of subjective measures (e.g., self-assessed health and subjective well-being) to reveal individual preferences.

In Chapter 3 we study whether and how a person's perception of incorrect behavior is influenced by offense and offender characteristics. Small crimes are usually considered to have less impact on an individual or situation, however the frequency of occurrences is much higher than for more serious crimes. This makes the topic very policy relevant in terms of possible deterrence mechanisms and the seriousness of this crime type, which may help to prioritize public action against it. We define 'perception' to measure perception on severity or justifiability of an action. The criminology literature has a long history regarding the study of crime perception although the focus has mainly been on more serious and property crimes. It is only in the last decade that studies include white-collar crimes, which have less serious direct consequences for victims. We even take a step further and look at incorrect behaviors that are not necessarily considered criminal by the judicial system. The apparent variation in the perception of the severity of different crimes suggests that social norms appear to vary across crimes and socio-economic groups. We analyze the perception of small crime using ordered response models. Firstly, we consider a very brief description of an incorrect behavior and secondly, we analyze the same incorrect behaviors but now we add offender and offense information, which results in vignette questions. The results show that a person's judgement of an incorrect behavior usually changes when more information about the offender and offense is available and that socio-economic groups have different perceptions of the justifiability of the small crimes under consideration. This implies that social norms not only depend on the (small) crime but also on the context in which it is committed.

In Chapter 4 we again look at respondents' perceptions but now only for a single small crime and in relation to the willingness to report this behavior if a respondent were to witness such act. Since reporting in the current set up occurs only within an organization —as it concerns incorrect behavior at the workplace— it is referred to as *peer reporting*. We look at how fairness perception interacts with decision making regarding reporting incorrect behavior after controlling for social norms. We measure fairness perception by means of the degree of justifiability regarding the situation of a fictive person that takes a bundle of printing paper from the office for private use. If an employee finds herself in a situation where a colleague displays incorrect behavior and considers this as unfair she can decide to report this (not necessarily to her supervisor) or not. The relation between peer reporting and fairness perception is, however, a more complex one as reporter characteristics and characteristics of the 'offense' and 'offender' are related to both. We find that internal attitude towards incorrect behavior is important for the decision making: Is the act considered fair? Did she herself take material from work home (self-justification)? Does she have a high social norm? Furthermore, we link a new aspect to peer reporting namely victimization. We conclude that past victimization, especially of incorrect behavior, increases the probability of peer reporting. This paper highlights the relevance of social norms for decision making and offers firms more insight on triggers for peer reporting. As employee theft can be very costly, firms benefit hugely from a social norm to peer report.

In the last chapter we explain subjective well-being (or, 'happiness') with a focus on attitude measures such as trust and social norms. We highlight the complexity of the relation between crime and well-being. A substantial part of the literature on well-being finds a negative effect from crime-related measures for both victims and non-victims, which is bad news as it implies that the cost of crime is not limited to direct costs such as medical expenses incurred by victims and loss of property. With this paper we add to this discussion by taking a broader view on the impact crime can have on our lives by arguing that there are also interrelations between victimization on one side and trust, health, and social norms on the other side. The implication of this is that victimization may have a much broader impact on a person, and hence is more costly (both financially and emotionally) than what is generally thought. Crime is measured locally (victimization and fear of crime rate in the region) and individually (small and serious crime victimization). We find evidence that personal victimization has a negative but weak relation with subjective well-being. In addition, local rates of victimization and fear of crime tell us that living in an area with many victims is negatively related to happiness but that fear of crime is positively related to well-being. Furthermore, we find that bad health, low social norms, and low trust are associated with lower subjective well-being. Furthermore, personal victimization implies lower trust and lower perceived health, while the relation with social norms is ambiguous. The empirical analysis relies on data that is a combination of information from several surveys conducted in 2008 (amongst others our survey on incorrect behavior).

Chapter 2

The Reliability of User Authentication through Keystroke Dynamics[1]

2.1 Introduction

People can be authenticated by something they know (password), something they have (credit card), or by something they are (finger prints). When typing on a keyboard a user can be authenticated through *what* he/she types (username, password), but also through *how* he/she types, that is, through keystroke dynamics. The purpose of this chapter is to investigate whether authentication through keystroke dynamics is sufficiently reliable as a security instrument to be used together with the more standard instruments.

The study of personal typing behavior (keystroke dynamics) is part of biometrics, where the underlying idea is that certain physical characteristics are (almost) unique and can therefore be used for authentication. Well-known examples are finger prints, voice recognition, and the iris scan.

[1] This chapter is based on Douhou and Magnus (2009).

The fact that people can be identified through their typing behavior, already known in the early days of the telegraph (Bryan and Harter, 1899), became important during the Second World War. Morse code is made up of dots and dashes, each of which has its described length. But no one replicates those prescribed lengths perfectly. The variation of spacing and the stretching out of the dots and dashes defines a 'rhythm' specific to the operator. This rhythm is called the operator's *fist*. In the Second World War, thousands of British so-called interceptors listened to German military radio broadcasts. These broadcasts were in code, so they could not be understood, but after a short while the interceptors could identify the fists of the German operators, just by listening to the rhythm of the transmission. Since the British were also able to locate the radio signals, they could follow the German radio operators around Europe, a very useful piece of war information (see Gladwell, 2005). The war experience has proved that a fist emerges naturally and unconsciously, that it reveals itself in even the smallest sample of Morse code, and that it is stable.

A sizable literature on keystroke dynamics has developed since Gaines et al. (1980) reported on an experiment where seven professional typists were each given a paragraph of prose to type, and the times between successive keystrokes were recorded. Since then, various authors have proposed different approaches, more specifically:

Statistical: Joyce and Gupta (1990), Bleha et al. (1990), Song et al. (1997)), Monrose and Rubin (1997; 2000), Bergadano et al. (2002), Guven and Sogukpinar (2003), Kacholia and Pandit (2003);

Data Mining: Brown and Rogers (1993), Obaidat and Sadoun (1997), Cho et al. (2000), Gutiérrez et al. (2002), Yu and Cho (2004).

The basic idea of the *statistical* approach is to compare a reference set of typing characteristics of a certain user with a test set of typing characteristics of the same user or a test set of a hacker. The distance between these two sets (reference and test) should be below a certain threshold or else the user is recognized as a hacker. Data mining is a collection of techniques from the field of Artificial Intelligence and Machine Learning, and includes also neural networks. A data mining process typically first builds a prediction model from historical data, and then uses this model to predict the outcome of a new trial (or to classify a new observation). In contrast to statistics, data mining makes no assumption about the data. The key difference between the statistical method and the data mining method is therefore the information that is used. For example, in a data mining approach, not only the similarities between the patterns of the same user are considered but also the differences of this pattern with all the other patterns observed in building the model. Thus, Lee and Cho (2007) develop a retraining framework by employing not only the user's but also hackers' characteristics. Our approach will be statistical.

Leggett et al. (1991) and Hoquet et al. (2005) propose *dynamic* authentication, where the system continuously monitors a user's typing pattern. If the pattern doesn't match the profile of the logged-on user the computer shuts down or asks the user or hacker to type a password. With this method one continuously updates and monitors a logged-on user's profile.

An excellent review on statistics and fraud is given by Bolton and Hand (2002). More specialized reviews on keystroke dynamics can be found, *inter alia*, in Lipton and Wong (1985) and Peacock et al. (2004). Distinguishing between real users and hackers can also be viewed as a one-class classification problem where one tries to distinguish one class of objects (real users) from all other possible objects (hackers) by learning from a training set containing only the objects of that class; see Duin and Tax (2005), Loog and Duin (2004), Zeng et al. (2006), and Kwak and Oh (2009) for discussion and examples of one-class classification problems.

One problem with the empirical applications is the lack of data. Gaines et al. (1980) have seven participants, and the studies by Umphress and Williams (1985), Obaidat and Sadoun (1997), Gutiérrez et al. (2002), Monrose et al. (2002), Hoquet et al. (2005), and Kang et al. (2008) employ between fifteen and twenty-five participants. Monrose and Rubin (1997) and Clarke and Furnell (2007) — in a study on mobile devices — employ around thirty participants. Somewhat larger are the studies by Schonlau et al. (2001), Bergadano et al. (2002), and Bartlow and Cukic (2006) — studying shift-key patterns — who employ around fifty participants.

In contrast, our data set consists of 1254 participants who typed *the same* username and password, twenty times each. Of course, mistakes were made and not all participants completed the full session of twenty logins. Nevertheless, the data set is large enough to be informative. The fact that each participant has the same username and password is important, because this allows us to consider each as a possible hacker to the other.

In Section 2.2 we describe the data. In Section 2.3 we develop a test statistic and obtain *theoretical* critical values for this test statistic. In Section 2.4 we obtain *empirical* critical values, which lead to better sizes and are therefore preferable in our study. In Section 2.5 we study the power of the tests, and Section 2.6 concludes.

2.2 The data

The data were collected in May and June 2007 by three students of the Systems and Network Engineering Group of the Faculty of Science at the University of Amsterdam; see Van Abswoude, Tavenier, and Van der Schee (Van Abswoude et al.). The students created a website (no longer in existence), which they advertised through the website of the weekly magazine of the University of Amsterdam (*http://www.folia.nl*), the principal Dutch website read by those with an interest in security systems (*http://www.security.nl*), and other channels.

When a potential participant hits the website, a 'session' is started. In total, 3476 sessions were started in this way. The first step for the participant is to click the relevant link and download a *flash applet* (developed by the students) to his/her own computer. The purpose of the flash applet is to record the necessary timings during the session, based on the clock of the participant's computer. The main activity thus takes place on the participant's computer and not on the website's server, and therefore technical problems such as network latency or overloading the server are avoided. Understandably, many potential participants did not download the flash applet or logged off immediately afterwards, without recording any timings. This happened in 64% of the sessions. This leaves us with 1254 sessions where timings have been recorded.

The participants were given a username (*patrick*) and a password (*water83*), the same for all participants. They were then asked to type username and password twenty times. For each of the twenty login attempts, the press (P) and release (R) clock times of each of the fourteen characters were recorded. This gives (P_i, R_i) for $i = 1, \dots, 14$. From these data we can calculate dwell times (D) and flight times (F) as

$$D_i := R_i - P_i, \qquad F_i := P_i - P_{i-1}.$$

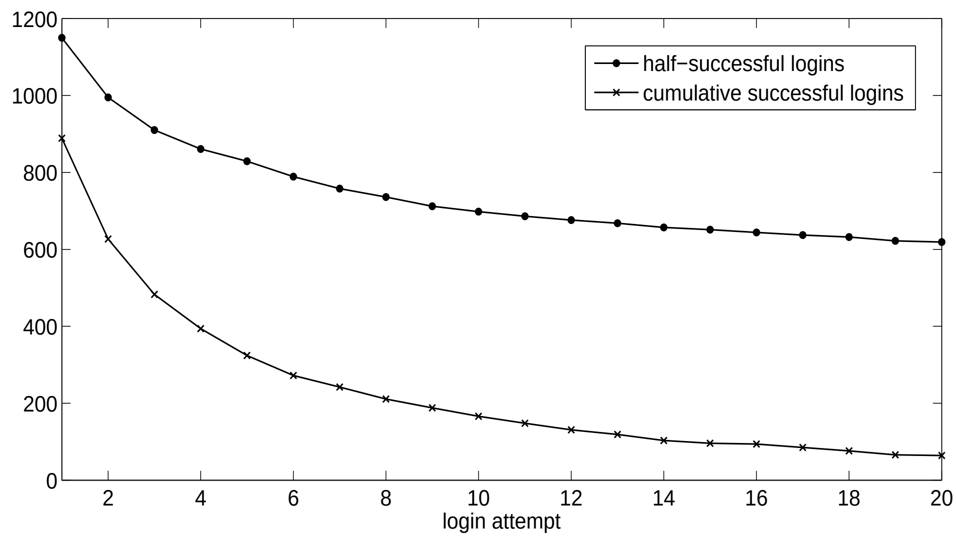
Hence, the dwell time records the time that each key is held pressed, and the flight time records the time between two consecutive press times. Clearly F_1 has no meaning. We also disregard F_8, because we attach no significance to the time elapsed between the last letter of the username and the first letter of the password. This gives us fourteen dwell times and twelve flight times per login attempt.

It might seem more natural to define flight time as $F_i^* = P_i - R_{i-1}$, so that the login duration is broken up in 'independent' non-overlapping pieces. This is not, however, a good idea, because F^* can be (and often is) negative. While the flash applet records both press and release times, characters registered by the computer are controlled only by the moment the key is pressed, not by the moment the key is released, and one may (and often will) press the next key when the previous key is not yet released.

If all participants would complete their session (twenty logins) and would make no typing errors, then we would have $26 \times 20 \times 1254 = 652,080$ data points. In fact, some participants quitted voluntarily (they closed their browser) or involuntarily (their computer crashed), so that they did not complete all twenty logins. In addition, participants made typing errors. If a typing error is made in a username (or password), then all dwell and flight times for that attempted username (password) are deleted. Errors can not be corrected using *backspace*, since this would confuse the interpretation of the dwell and flight times. If an error is made in the username but not in the password (or vice versa), then the correctly typed password (or username) data are *not* deleted.

Some information about early exits and error rates is provided in Figure 2.1. Of the 1254 participants who started, 104 made a mistake in both

Figure 2.1: Number of 'half-successful' and cumulative 'successful' login attempts



username and password in the first login; 1150 participants 'half-successfully' completed the first login (by being error-free in either username or password or both). This is the first point on the upper graph. Then, 995 participants half-successfully completed the second login: the second point. Finally, 619 participants half-successfully completed the twentieth login: the last point.

Hence, at least 619 participants completed the whole session — at least, because some made an error in both username and password in the final login. The curve is decreasing because some participants drop out during the session. Of the 1150 participants who were half-successful (error-free in either username or password or both) in their first login, 889 were 'successful' (error-free in both username and password). Of those, 627 were successful in the first two logins, and only 64 were successful in all twenty logins. Hence, in contrast to the upper graph, the lower graph in Figure 2.1 provides cumulative information.

The participants are taken from a small group, consisting primarily of Dutch students, university employees, and those interested in security systems. We do not claim that this is a representative sample. It is possible that the typing behavior of the people in our sample differs from that of individuals with less computer experience. If there is a difference, then the people in our sample are expected to be more homogeneous than the average population, making it more difficult to detect differences in their typing patterns. Hence if we find that we can detect differences in typing patterns in our sample, then it should be easier in a less homogeneous group.

The username and password were chosen to reflect common practice. Both username and password have seven characters. They are simple and easy to remember. The addition of two digits (83) in the password is also quite common (typically, year of birth). We note that there are no repeated characters within username or password, which has practical importance in our experiment because it means that difficult issues of identification are avoided. All letters are lowercase symbols.

All participants were given the same username and password. As we shall see, this is of great practical use in our analysis, because we can consider each participant as a possible 'hacker' to everybody else.

To gain some insight into the dwell and flight times and their variation, we consider all participants with at least six error-free username attempts (898 people) and all participants with at least six error-free password attempts

(897 people). For each person we calculate the average dwell and flight times: seven average dwell times and six average flight times per person for username and password separately. These averages define an empirical distribution from which we can calculate quantiles. The 10%, 50% (median), and 90% quantiles are given in Figures 2.2 and 2.3. We see from Figure 2.2 that the median dwell times fluctuate around 90 ms for the username and around 100 ms for the password, and that there is not much difference between the fourteen characters.[2] The 10% and 90% quantile lines reveal, however, considerable variation among participants.

Figure 2.2: Median dwell times with 80% bounds for username (upper panel) and password (lower panel)

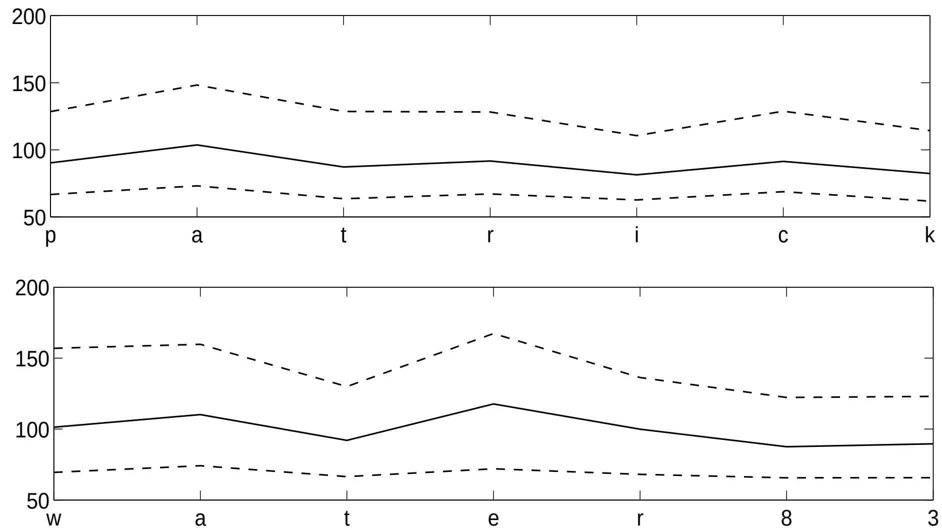
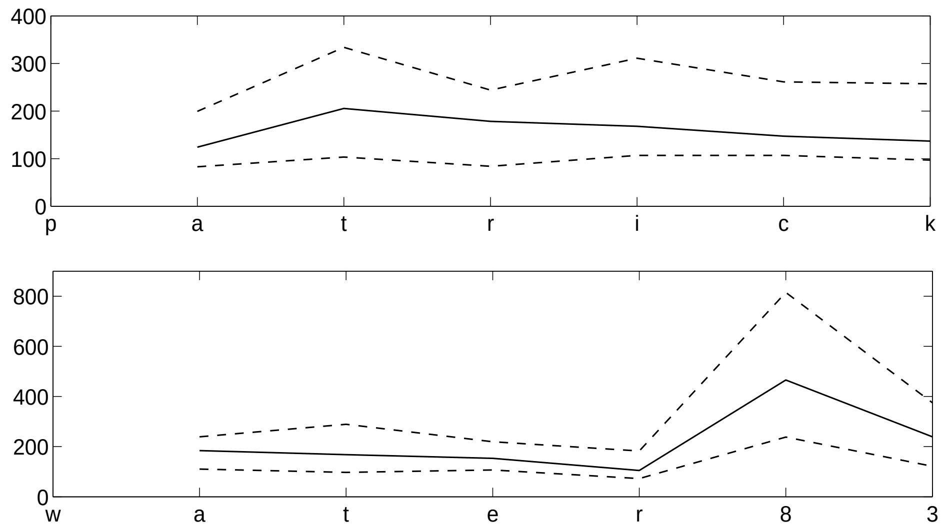


Figure 2.3 shows that the median flight times fluctuate around 160 ms for the username and around 219 ms for the password. The large difference between average flight times in username and password can be contributed to the time it takes to move from r to 8 in the password *water83*, namely 465 ms. Apart from this, there is not much difference between the average

[2] One millisecond (ms) is one thousandth of a second.

Figure 2.3: Median flight times with 80% bounds for username (upper panel) and password (lower panel)



flight times. The first four flight times of the password (only letters) fluctuate around 152 ms. Again, there is considerable variation among participants. In fact, there is more variation in flight times than in dwell times, because of individual differences in keyboard-control: a person who uses only two fingers will have a larger flight time on average than a person who uses ten fingers.

Finally we comment briefly on the within-person variance. We compare participants from the group where the first login is deleted and exactly fifteen of the remaining nineteen logins are correct (96 participants, later called group 15(1)) with the group of all participants who have at least six error-free attempts. We then calculate for each of the 96 participants and for each character the standard deviation of the dwell times and compare this with the average over one thousand random draws of fifteen attempts on the same character from the entire population. The within-person standard deviation is about 47% for the username and 44% for the password compared to the standard deviation in the whole population. We repeat the experiment for a second group where the first five logins are deleted, and all fifteen remaining logins are correct (136 participants, later called group 15(5)). Then the within-person standard deviation drops to about 42% for the username and 38% for the password compared to the standard deviation in the whole population. The percentages in the second experiment are lower because these participants make fewer errors and are therefore likely to be more consistent typists. A drop in standard deviation of 50-60% may not seem much to develop a powerful test. Nevertheless we shall see that considerable power can be achieved.

2.3 The test statistic and theoretical critical values

For a given participant we have n observations on each of m characteristics, for example $n = 20$ (number of logins) and $m = 26$ (number of characteristics: 14 dwell times and 12 flight times). Let x_{ij} denote the i-th observation on the j-th characteristic. If we assume that the $x_i := (x_{i1}, \ldots x_{im})'$ are independently and identically distributed as

$$x_i \sim \mathrm{N}(\mu, \Sigma), \qquad \Sigma := \mathrm{diag}(\sigma_1^2, \ldots, \sigma_m^2), \tag{2.1}$$

so that the characteristics are independent of each other, then the maximum likelihood estimators of μ_j and σ_j^2 are given by

$$\hat{\mu}_j = \bar{x}_j := \frac{1}{n}\sum_{i=1}^{n} x_{ij}, \qquad \hat{\sigma}_j^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2.$$

Note that the means μ_j and variances σ_j^2 are assumed to be individual-specific.

We have made two strong independence assumptions: the characteristics are independent of each other, and consecutive logins are independent. The first assumption means that if, for example, within one login the first flight time is smaller than expected, this has no impact on the next flight time. This is certainly not entirely true but it seems a reasonable simplification.[3] The second assumption is more difficult to defend and to repair, and dependence between consecutive logins could very well be the reason why the critical values are unsatisfactory, as we shall see.

For the moment we adopt these two independence assumptions. Now assume that, in addition to $x_1, \dots, x_n$, we have one other $m \times 1$ vector y, independent of the $\{x_i\}$. Under the null hypothesis that y is generated by the same distribution as the $\{x_i\}$, we have

$$\bar{x}_j \sim \mathrm{N}(\mu_j, \frac{\sigma_j^2}{n}), \qquad y_j \sim \mathrm{N}(\mu_j, \sigma_j^2),$$

so that

$$\frac{n}{n+1} \sum_{j=1}^{m} \frac{(\bar{x}_j - y_j)^2}{\sigma_j^2} \sim \chi^2(m). \tag{2.2}$$

As our test statistic we propose

$$T_{m,n} := \frac{n}{m(n+1)} \sum_{j=1}^{m} \frac{(\bar{x}_j - y_j)^2}{\hat{\sigma}_j^2}, \tag{2.3}$$

whose distribution depends only on m and n. Since

$$t_j := \sqrt{\frac{n-1}{n+1}} \cdot \frac{\bar{x}_j - y_j}{\hat{\sigma}_j} \sim \mathrm{Student}(n-1),$$

we can write

$$T_{m,n} = \frac{n}{m(n-1)} \sum_{j=1}^{m} t_j^2. \tag{2.4}$$

For large n, the statistic $T_{m,n}$ can be approximated by a $\chi^2(m)/m$-distribution. For large m, it can be approximated by a normal distribution using the exact moments

$$\mathrm{E}(T_{m,n}) = \frac{n}{n-3}, \qquad \mathrm{var}(T_{m,n}) = \frac{2}{m} \cdot \frac{n^2(n-2)}{(n-3)^2(n-5)}.$$

[3] We discuss a possible extension to dependence in the Conclusion.

However, for values like $n = 20$ and $m = 26$, the asymptotic behavior is of little use, and we have to resort to simulation.

For given values of m and n and for given significance levels α, the distribution of $T_{m,n}$ can be simulated and quantiles k_α satisfying

$$\Pr(T_{m,n} > k_\alpha) = \alpha$$

can be estimated. As shown in Shorack and Wellner (1986, Example 1, p. 639), the sample quantiles $\hat{k}_\alpha$ are consistent and asymptotically normal, and

$$\widehat{\mathrm{var}}(\hat{k}_\alpha) \approx \frac{\alpha(1-\alpha)}{r\left(f_r(\hat{k}_\alpha)\right)^2}, \tag{2.5}$$

provides a consistent estimate of the variance of $\hat{k}_\alpha$, where $f_r(\hat{k}_\alpha)$ denotes an estimate of the density of $T_{m,n}$ at k_α after r replications. Since we want our estimates for k_α to be accurate to two decimal places, we could use (2.5) to determine the number of replications r. In practice, it is more efficient to take N independent batches of $100,000$ replications each for every combination of m and n, and calculate the mean and variance over these N batches. For $N = 1000$ we obtain a standard deviation of $\hat{k}_\alpha$ of about 0.0003 for $\alpha = 0.05$ and 0.0004 for $\alpha = 0.01$, which secures the required accuracy. Thus we

Table 2.1: Theoretical critical values of the $T_{m,n}$ test

	α	0.01			0.05			0.10		
n	m	12	14	26	12	14	26	12	14	26
13		3.29	3.11	2.55	2.45	2.37	2.07	2.10	2.05	1.85
15		3.07	2.92	2.41	2.33	2.25	1.97	2.01	1.95	1.77
17		2.94	2.79	2.30	2.25	2.16	1.90	1.94	1.89	1.71
19		2.83	2.69	2.23	2.18	2.10	1.85	1.89	1.84	1.67
∞		2.18	2.08	1.76	1.75	1.69	1.50	1.55	1.50	1.37

obtain Table 2.1 with the relevant quantiles (critical values) k_α of the $T_{m,n}$ test statistic for fifteen (m, n) combinations and three commonly used values of α.

2.4 Empirical critical values

We will see shortly that the critical values of Table 2.1, dictated by statistical theory under the simplest assumptions, are not accurate enough to make predictions about the power of the test.

Let us distinguish between the $m_1 = 12$ flight times and $m_2 = 14$ dwell times in our sample, and consider two test statistics, using (2.3),

$$T_1 = \frac{n}{m_1(n+1)} \sum_{j=1}^{m_1} \frac{(\bar{x}_j - y_j)^2}{\hat{\sigma}_j^2}, \quad T_2 = \frac{n}{m_2(n+1)} \sum_{j=1}^{m_2} \frac{(\bar{x}_j - y_j)^2}{\hat{\sigma}_j^2},$$

for flight times and dwell times separately, together with the combined statistic

$$T = \frac{m_1}{m} T_1 + \frac{m_2}{m} T_2.$$

We shall consider four subsets of our data. Since the participants are unfamiliar with their username and password, they need some time to practice. In the first three subsets we therefore delete the first of the logins, as follows:

Group 19(1): first login is deleted, all nineteen remaining logins are correct (78 participants);

Group 17(1): first login is deleted, exactly seventeen of the remaining nineteen logins are correct (161 participants);

Group 15(1): first login is deleted, exactly fifteen of the remaining nineteen logins are correct (96 participants).

These three groups are mutually exclusive. In addition, we consider one further subset where the first five logins have been deleted.

Group 15(5): first five logins are deleted, all fifteen remaining logins are correct (136 participants).

Notice that Group 19(1) is a subset of Group 15(5), and that Groups 17(1) and 15(1) intersect with Group 15(5), but are no subsets.

For each of these four groups we perform a small experiment, as follows. Suppose our group is 19(1). For each of the 78 people in this group we select two login attempts, labeled $y_{(1)}$ and $y_{(2)}$, which can be done in $\binom{19}{2} = 171$ ways. From the remaining $n = 17$ login attempts we calculate $\bar{x}_j$ and $\hat{\sigma}_j^2$ for each j. For both $y_{(1)}$ and $y_{(2)}$ separately we then calculate T_1, T_2, and T. If we do this for each of the 78 people in the group, we obtain 156 values for T_1, T_2, and T. Repeating the experiment for each person and each combination provides us with $78 \times 171 \times 2 = 26,676$ values for T_1, T_2, and T. Each test outcome is then confronted with the appropriate theoretical critical value in Table 2.1 for $n = 17$ and $m_1 = 12$ (T_1), $m_2 = 14$ (T_2), and $m = 26$ (T), respectively, and the proportion of times that the test rejects (the size) is calculated. In Table 2.2 we report the empirical sizes for two of the four

Table 2.2: Size of the $T_{m,n}$ test based on theoretical critical values

		α	0.01			0.05			0.10		
n	G	m	12	14	26	12	14	26	12	14	26
13	15(5)		0.15	0.04	0.15	0.18	0.08	0.19	0.21	0.12	0.22
17	19(1)		0.15	0.04	0.15	0.18	0.09	0.19	0.20	0.13	0.22

subsets, namely 15(5) and 19(1); the other two subsets behave similarly. We see that the empirical sizes are about 15 (T_1) to 4 (T_2) times as large as predicted when $\alpha = 0.01$, about 3.6 (T_1) to 1.7 (T_2) times as large when $\alpha = 0.05$, and about 2.1 (T_1) to 1.3 (T_2) times as large when $\alpha = 0.10$. The larger is α, the better the empirical size is approximated by the theoretical size. Also, the approximation works better for T_2 (dwell times) than for T_1 (flight times).

Although the theoretical sizes are possibly acceptable for $\alpha = 0.10$, they are not for $\alpha \leq 0.05$. Hence we shall obtain better results for the values of interest when we use empirical critical values, instead of the theoretical critical values of Table 2.1.

The empirical critical values are obtained as follows. Suppose again that the group of interest is 19(1). The calculations are the same as for Table 2.2 leading to $78 \times 171 \times 2 = 26,676$ values for T_1, T_2, and T. For α equal to 0.01, 0.05, and 0.10 we then estimate the critical value k_α satisfying $\Pr(T^* > k_\alpha) = \alpha$, where T^* takes the values T_1, T_2, and T, respectively. We repeat these calculations for each of seven groups:

19(1): $n = 17$, $n = 15$, $n = 13$;

17(1): $n = 15$, $n = 13$;

15(1): $n = 13$;

15(5): $n = 13$.

For example: group 17(1) contains all participants where, ignoring the first login, precisely 17 of the remaining 19 logins are correct. From these 17 correct logins we select n (15 or 13) at random. The results are given in Table 2.3, which confirms that these (empirical) critical values are quite different from the theoretical values in Table 2.1. We notice that, within each group, we have selected *two* login attempts, $y_{(1)}$ and $y_{(2)}$, and for each separately we have calculated T_1, T_2, and T. Let us denote these statistics

Table 2.3: Empirical critical values: one draw

		α	0.01			0.05			0.10		
n	G	m	12	14	26	12	14	26	12	14	26
13	15(5)		86.44	5.53	57.66	14.82	2.81	8.35	5.21	2.21	3.79
13	19(1)		103.35	5.99	54.41	17.58	2.93	9.88	5.81	2.27	3.97
13	17(1)		71.77	7.32	42.42	15.62	2.86	9.39	5.74	2.26	3.91
13	15(1)		90.06	6.58	54.36	18.27	2.70	10.68	6.38	2.20	4.35
15	19(1)		82.41	5.20	41.26	15.47	2.69	8.54	5.24	2.14	3.56
15	17(1)		66.22	5.36	34.49	14.51	2.72	8.25	5.38	2.15	3.65
17	19(1)		85.41	4.77	42.17	15.48	2.60	8.27	4.90	2.06	3.36

as $T_1^{(1)}$, $T_2^{(1)}$, and $T^{(1)}$ for $y_{(1)}$, and $T_1^{(2)}$, $T_2^{(2)}$, and $T^{(2)}$ for $y_{(2)}$. Defining

$$T_1^{min} := \min(T_1^{(1)}, T_1^{(2)}), \quad T_2^{min} := \min(T_2^{(1)}, T_2^{(2)}), \quad T^{min} := \min(T^{(1)}, T^{(2)}),$$

we obtain $78 \times 171 = 13,338$ values for T_1^{min}, T_2^{min}, and T^{min}. For α equal to 0.0001, 0.0025, and 0.0100 (the squares of the previous α-values), we then estimate the critical values k_α, and we repeat these calculations for each of seven groups. This leads to Table 2.4. Since

Table 2.4: Empirical critical values: two draws

		α	0.0001			0.0025			0.0100		
n	G	m	12	14	26	12	14	26	12	14	26
13	15(5)		106.20	—	—	14.53	4.61	14.12	6.02	2.40	4.43
13	19(1)		137.90	131.66	92.99	25.99	3.57	18.93	7.49	2.52	4.81
13	17(1)		117.22	—	—	19.95	5.51	17.77	6.21	2.52	4.75
13	15(1)		—	—	—	25.93	4.01	20.88	8.26	2.42	5.35
15	19(1)		106.79	71.73	49.71	15.39	3.14	9.36	6.42	2.31	4.09
15	17(1)		120.13	—	—	17.27	3.48	10.25	5.79	2.31	3.92
17	19(1)		121.46	17.81	56.44	15.59	2.90	8.29	6.07	2.24	3.75

$$\Pr(T^{(1)} > k_\alpha \text{ } and \text{ } T^{(2)} > k_\alpha) = \Pr(\min(T^{(1)}, T^{(2)}) > k_\alpha) = \Pr(T^{min} > k_\alpha),$$

we see that Table 2.4 contains the required critical values for two consecutive draws. If these two draws were independent (which they might not be), then we would have

$$\Pr(T^{(1)} > k_1 \text{ } and \text{ } T^{(2)} > k_2) = \Pr(T^{(1)} > k_1) \Pr(T^{(2)} > k_2)$$

for all k_1 and k_2, and the numbers reported in Tables 2.3 and 2.4 would be identical. The results in Tables 2.3 and 2.4 suggest that in fact

$$\Pr(T^{(1)} > k_1 \text{ } and \text{ } T^{(2)} > k_2) \geq \Pr(T^{(1)} > k_1) \Pr(T^{(2)} > k_2)$$

or, what amounts to the same, that

$$\Pr(T^{(2)} > k_2 \,|\, T^{(1)} > k_1) \geq \Pr(T^{(2)} > k_2),$$

implying that $T^{(1)}$ and $T^{(2)}$ are 'positively quadrant dependent' (Lehmann, 1966). Although the two draws are clearly not independent, the deviation from independence does not appear to be large.

It is already clear from Table 2.3 that the critical values for $\alpha = 0.01$ are less stable than those for $\alpha = 0.05$ and $\alpha = 0.10$. This effect is even stronger in Table 2.4: the critical values for $\alpha = 0.0001$ are very unstable. In fact, certain critical values become infinite, due to the fact that one or more of the $\hat{\sigma}_j^2$ in Equation (2.3) become zero. This can only happen if a participant has n identical logins. If a critical value is infinitely large then we will find an empirical power close to zero, and hence a hacker is always treated as an authorized user. This requires further investigation. We find the following:

Group 19(1): No infinite values were found, but for $n = 13$ and $n = 15$ (not for $n = 17$) some may in fact be there because not all possibilities have been examined.

Group 17(1): two participants recorded fourteen (out of seventeen) identical dwell times on the letter i in the username *patrick*, and one recorded fifteen identical dwell times on the letter a in *patrick*. It is possible that there are more infinite values for $n = 13$ (but not for $n = 15$) because not all possibilities have been examined.

Group 15(1): one participant recorded thirteen (out of fifteen) identical dwell times on the letter t in the password *water83*, while another participant recorded thirteen identical dwell times on the letter w in *water83*, and also thirteen identical flight times on the passage t–e in *water83*.

Group 15(5): one participant recorded fourteen (out of fifteen) identical dwell times on the letter i in *patrick*, and also thirteen identical dwell times on the letter r in *water83*.

This is rather surprising, at least it was surprising to us. Apparently some participants display a very high degree of regularity in typing behavior, which underlines the potential for using keystroke dynamics for user authentication.

2.5 Power of the test

Now that we have computed the empirical critical values for three given sizes α, we can consider the power of our test, that is, the probability that a 'hacker' is recognized as a hacker. Suppose one of the other people in the same group 'breaks in'. What is the probability that he/she is found out?

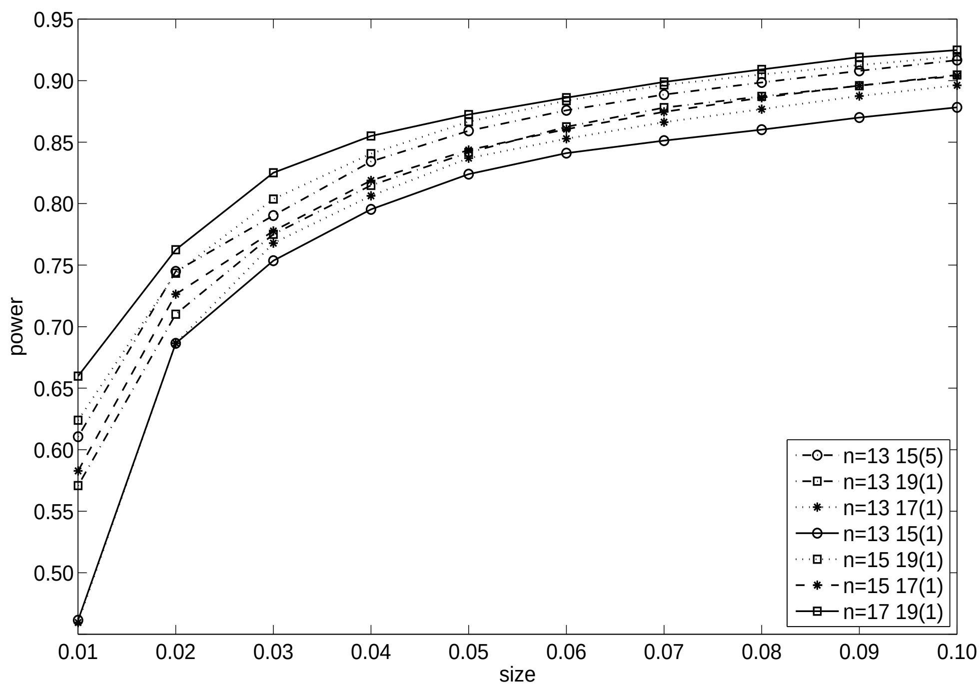
Based on the empirical critical values of Tables 2.3 and 2.4, we perform the following experiment. Choose one group, say 19(1) with $n = 17$. Choose two people (ordered) in this group, say (i, j), where person i is the potential victim and person j is the hacker. This can be done in $78 \times 77 = 6006$ ways. Draw randomly n observations from i and two observations from j, and calculate the test statistics. Thus we obtain $6006 \times 2 = 12,012$ values for each of the three test statistics. We then confront these values with the appropriate critical values in Table 2.3. This will give us the probability that our tests will label a person as a hacker when the login was indeed performed by a hacker, that is, the power of our tests. Table 2.5 shows that the power

Table 2.5: Empirical power of the $T_{m,n}$ test: one draw

		α	0.01			0.05			0.10		
n	G	m	12	14	26	12	14	26	12	14	26
13	15(5)		0.07	0.61	0.07	0.42	0.86	0.63	0.76	0.92	0.90
13	19(1)		0.05	0.57	0.06	0.31	0.84	0.53	0.68	0.90	0.88
13	17(1)		0.06	0.46	0.07	0.30	0.84	0.48	0.64	0.90	0.86
13	15(1)		0.04	0.46	0.04	0.22	0.82	0.38	0.55	0.88	0.78
15	19(1)		0.06	0.62	0.09	0.31	0.87	0.56	0.68	0.92	0.90
15	17(1)		0.06	0.58	0.08	0.30	0.84	0.52	0.64	0.90	0.86
17	19(1)		0.05	0.66	0.08	0.31	0.87	0.57	0.69	0.92	0.91

for $\alpha = 0.01$ is not good. But for $\alpha = 0.05$ the test based on dwell times (T_2) gives a power of about 85%, and for $\alpha = 0.10$ the test based on dwell times (T_2) gives a power of about 90% and the overall test (T) gives a power of about 87%. This suggests that dwell times are more discriminatory and therefore more powerful than flight times. For the dwell times ($m = 14$) we

Figure 2.4: Empirical power versus size, dwell times: one draw



visualize the trade-off between size and power in Figure 2.4. All data sets behave in the same way, but obviously more information (larger number of successful logins) leads to higher power.

The results in Table 2.5 assume that a person is labeled as a hacker when the test fails once. In many situations one is allowed a second chance, and a person is only labeled as a hacker when he/she fails twice. The power of test based on two attempts is similarly calculated, now using the critical values in Table 2.4. The results in Table 2.6 confirm the results in Table 2.5. For $\alpha = 0.0025$ the test based on dwell times (T_2) gives a power of about 67%

Table 2.6: Empirical power of the $T_{m,n}$ test: two draws

		α	0.0001			0.0025			0.0100		
n	G	m	12	14	26	12	14	26	12	14	26
13	15(5)		0.03	—	—	0.35	0.62	0.33	0.64	0.86	0.82
13	19(1)		0.01	0.01	0.01	0.14	0.71	0.19	0.50	0.83	0.76
13	17(1)		0.01	—	—	0.16	0.51	0.17	0.52	0.82	0.73
13	15(1)		—	—	—	0.08	0.60	0.10	0.34	0.80	0.61
15	19(1)		0.02	0.02	0.04	0.22	0.76	0.44	0.52	0.86	0.81
15	17(1)		0.01	—	—	0.17	0.69	0.33	0.52	0.84	0.79
17	19(1)		0.01	0.14	0.03	0.22	0.78	0.48	0.53	0.87	0.83

and for $\alpha = 0.01$ the test based on dwell times (T_2) gives a power of about 84%. The overall test (T) gives a power of about 76%.

Recall that 19(1) denotes the group where the first login has been deleted and all remaining nineteen logins are correct, and that 17(1) and 15(1) denote the groups where again the first login has been deleted and where exactly seventeen or fifteen of the remaining nineteen logins are correct. For $n = 13$ we see that the power increases when we move from 15(1) to 19(1), and for $n = 15$ we see that the power increases when we move from 17(1) to 19(1). This confirms that if a user exhibits more regularity, it will be easier to establish his/her pattern, and it will be more difficult for a potential hacker to break in.

2.6 Conclusion

Based on our experiments we conclude that keystroke dynamics can be a reliable security instrument for authentication. It appears that dwell times (how long a key is held pressed) are more discriminatory and therefore more powerful than flight times (time between consecutive press times), confirming a similar finding by Obaidat and Sadoun (1997). Our T_2-test based on dwell times tells us that:

if we reject a person if the T_2-test fails once, then it will reject the true owner 5% of the time and recognize a hacker 85% of the time (Table 2.5, $\alpha = 0.05$, power $= 0.85$);

if we reject a person if the T_2-test fails twice, then it will reject the true owner 1% of the time and recognize a hacker 84% of the time (Table 2.6, $\alpha = 0.01$, power $= 0.84$).

In practice, a biometric test will be used in combination with another test or perhaps several other tests. In such, more realistic, cases we have:

$$\begin{aligned}\Pr(\text{hacker successful}) &= \Pr(\text{our test fails } \textit{and} \text{ current tests fail})\\ &= \Pr(\text{our test fails} \,|\, \text{current tests fail}) \times \Pr(\text{current tests fail}).\end{aligned}$$

Suppose that the hacker is recognized with the current tests in about 99% of the attempts, so that $\Pr(\text{current tests fail}) = 0.01$. Suppose also that, if the current tests do not recognize the hacker, our test does recognize the hacker in about 85% of the attempts, so that $\Pr(\text{our test fails} \,|\, \text{current tests fail}) = 0.15$. Then we find that $\Pr(\text{hacker successful}) = 0.0015$, so that the hacker will be unmasked in 99.85% of the attempts. The biometric test thus improved the power from 99.00% to 99.85%, and the cost caused by hackers will be reduced by 85% if the biometric test is added to the authentication procedure.

It is difficult to compare our results to the literature, because every paper has a different number of participants, a different set-up, and a different statistical method. As a tentative guide we summarize below the type-I errors (α) and type-II errors (β) as reported in the literature:

Umphress and Williams (1985): $\alpha = 0.12$, $\beta = 0.06$;
Bleha et al. (1990): $\alpha = 0.08$, $\beta = 0.03$;
Leggett et al. (1991): $\alpha = 0.06$, $\beta = 0.05$;
Monrose and Rubin (1997): $0.09 < \alpha < 0.37$;
Bergadano et al. (2002): $0.02 < \alpha < 0.06$, $\beta < 0.01$;

Gutiérrez et al. (2002): $\alpha = 0.20$, $\beta = 0.04$;
Kacholia and Pandit (2003): $0.01 < \alpha < 0.08$, $0.01 < \beta < 0.08$;
Gunetti and Picardi (2005): $\alpha = 0.01$, $\beta = 0.05$.

The high power (around 95%) obtained in these studies is a little puzzling given the typically very small number of participants. We have more participants and obtain lower power. Nevertheless, the main conclusion from our analysis is that especially dwell times (how long a key is pressed) can be used to create a powerful test. At a size of 1% the power of our best-performing two-draw test is 84% ($\beta = 0.16$).

Some caution is required in applying our results to different situations. First, our data may not be representative. Mostly students and people interested in security systems take part in our experiment. We do not know whether this affects our analysis, but if it does then the people in our sample are expected to be more homogeneous than the average population, making it more difficult to detect differences in their typing patterns. Since we can detect differences in typing patterns in our sample, it should be easier to detect such differences in a less homogeneous group. The reported power can thus be viewed as a lower bound. Second, we only consider the username-password combination, which together contains fourteen characters. In an environment where fewer (sometimes only four) characters are required from the user, it is doubtful that the user can be authenticated with sufficient accuracy. Third, the fact that our set-up has no repeated characters may influence our results.

We have developed the test statistic under the assumption that the characteristics are independent. This is probably unrealistic and more power can be obtained by allowing for some dependence, perhaps using Markov models (Jiang et al., 2007). Suppose, as before, that for a given participant we have n observations on each of m characteristics, and let x_{ij} denote the i-th observation on the j-th characteristic. Assume again that the $x_i := (x_{i1}, \dots x_{im})'$

are independently and identically distributed, but now as

$$x_i \sim \mathrm{N}(\mu, \Sigma), \qquad \Sigma := \Sigma(\theta),$$

where θ is a $k \times 1$ vector of unknown parameters. In Equation (2.1) we assumed that $\theta = (\sigma_1^2, \ldots, \sigma_m^2)'$ and $k = m$, implying that the characteristics are independent of each other. If we drop this assumption, then the maximum likelihood estimator of μ is again given by $\hat{\mu} = \bar{x} := (1/n) \sum_i x_i$, and the maximum likelihood estimator of θ is obtained by

$$\min_{\theta} \left(\log |\Sigma(\theta)| + \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})' \Sigma^{-1}(\theta)(x_i - \bar{x}) \right).$$

More explicitly, letting

$$S := \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})',$$

the $\hat{\theta}$'s are found by solving the k equations

$$\operatorname{tr}\left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_h} \right) = \operatorname{tr}\left(\Sigma^{-1} S \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_h} \right) \qquad (h = 1, \ldots, k),$$

from which we see that the maximum likelihood estimator $\hat{\theta}$ depends on the observations only through S. Instead of (2.2) we then have

$$\frac{n}{n+1} (\bar{x} - y)' \Sigma^{-1} (\bar{x} - y) \sim \chi^2(m),$$

and the test statistic becomes

$$T_{m,n} := \frac{n}{m(n+1)} (\bar{x} - y)' \Sigma(\hat{\theta})^{-1} (\bar{x} - y).$$

Since we have to estimate more parameters, we would require more data in this case. We recommend this extension only if the number of observations is large, because otherwise the additional noise generated by having to estimate more parameters might outweigh the additional power of the test.

We have taken account of the fact that username and password are unfamiliar to our participants, by deleting either the first or the first five logins. When comparing the power for $n = 13$ and $G = 15(1)$ and $15(5)$, respectively, we see in Table 2.5 at $\alpha = 0.05$ that the power of the T_2 statistic increases from 0.82 when only the first login is deleted to 0.86 when the first five observations are deleted. Similarly, in Table 2.6 at $\alpha = 0.01$, the power of the T_2 statistic increases from 0.80 to 0.86.

We also note that in practice the number of observations on a specific user (n in our analysis) will be larger than what we use in our experiment (maximum 17) and hence will increase the power of our tests. For example, in Table 2.5 at $\alpha = 0.05$ and $G = 19(1)$, the power of the T_2 statistic increases from 0.84 when $n = 13$ to 0.87 when $n = 17$, and, similarly, in Table 2.6 at $\alpha = 0.01$ and $G = 19(1)$, the power of the T_2 statistic increases from 0.83 to 0.87. This confirms that a larger value of n will increase the power of our test.

In practical applications, the user will be familiar with his/her username and password, and also the number of observations will be larger than 17. It seems therefore reasonable to believe that our power estimates are lower bounds, and that the power of our tests will be higher in practice.

Finally, the balance between type-I error and type-II error can be controlled by the company. In a period when many hackers are active, the company may choose to increase α, thus increasing the power. Users may be annoyed because they may be denied access to their own accounts, but hackers will find it more difficult to break in.

In conclusion, keystroke dynamics can be a reliable and flexible security instrument for authentication, if used in addition with other instruments. It seems more suitable for authentication (verification) than for identification.

Chapter 3

The Perception of Small Crime[1]

3.1 Introduction

Living together in a society is guided by formal and informal rules. Violations of these rules can be costly to society and they are, in the case of large crimes, followed by prosecution. Minor misbehaviors — small crimes — do not usually result in legal proceedings, because the cost of enforcing compensation of small crimes would be too high or because the law does not permit prosecution. Although the economic consequences of a single small crime will be low, such crimes are often quite common and can, in the aggregate, generate substantial losses. For example, in the year 2000, surfing the Internet at work for private use may have cost society worldwide \$50 billion per year and employee theft around \$200 billion (Greenberg and Scott, 1996).

In standard economic models of criminal behavior Becker (1968), individuals who undertake illegal actions evaluate the probabilities and consequences of being punished, and commit a crime only if the expected value of doing so exceeds the utility of the status quo. Thus, an individual would commit

[1] This chapter is based on Douhou et al. (2011b).

a (small) crime if the risk-sanction trade-off is favorable. The legal sanction acts as a market price, and the individual treats the sanction as an *external* constraint. Alternatively, the individual may *internalize* the obligation associated with the sanction. When many people in a community do this, it becomes a social norm (Cooter, 1998). Since the Becker model is at odds with the data, an extension of this model with social norms seems appropriate.

Balestrino (2008) uses the lack of social norms as an explanation why digital piracy (downloading and copying films or music illegally) is much more common than other types of small crimes. Orviska and Hudson (2002) use survey data on attitudes towards tax evasion and the tendency to evade to show that social norms. Traxler (2010) introduces a formal model for tax evasion in which the utility of evading taxes depends negatively on the social norm, which in turn depends on how many others evade taxes. Kube and Traxler (2011) emphasize the relevance of social norms for public policy on legal enforcement, since higher penalties not only have a direct effect on the expected gains of non-compliance, but also an indirect effect by changing the social norm. The variation in the perception of the severity of small crimes in society, rather than (or, in addition to) the (low) probability of being caught or the punishment in case of being caught, shows how social norms vary across crimes and across socio-economic groups. Social norms are increasingly important in theoretical and empirical work in economics, and the value of our study mainly relies on the link to and the relevance for the social norms literature.

Measuring the perception of crime can be useful to evaluate how sentencing guidelines correspond to public sentiment and to the allocation of police resources (Miethe, 1982). The perception of larger crimes has been studied extensively in the criminology literature (see, e.g., the survey of Stylianou, 2003). The first to study the perceptions of 'crime seriousness' were Sellin and Wolfgang (1964) who developed a new method to measure seriousness, thus providing new insights on public consensus and relative ordering of criminal acts. The existing literature in criminology focuses on serious crimes or property crimes (O'Connell and Whelan, 1996; Rosenmerkel, 2001; Rossi et al., 1974), and white-collar crimes (Isenring, 2008; Piquero et al., 2008; Rosenmerkel, 2001). This literature ranks crimes in terms of seriousness, and finds that there is relative consensus in the sense that different groups usually give the same ranking, but not absolute consensus in the sense that the seriousness scores are approximately equal. *Harmfulness* and *wrongfulness* are found to be the key dimensions driving perceived seriousness (Rosenmerkel, 2001; Warr, 1989). While harmfulness refers to the perceived consequences for the victims, wrongfulness refers to morality and the social norms in society or a socio-economic group. Differences were sometimes found between groups of different gender, age, education, degree of urbanization, etc., but in many cases the differences were statistically insignificant, so that no clear systematic picture emerges. There is also evidence that the perceived seriousness of crimes may depend on characteristics (such as age or gender) of who commits the crime (Rossi et al., 1997), on whether the crime is committed once or repeatedly (Herzog and Oreg, 2008), and on other circumstances under which the crime is committed. For example, the justifiability of employee theft depends on behavior of superiors and the peer group of co-workers (Jones and Kavanagh, 1996). All this shows that the perceived seriousness not only depends on the consequences for victims but also on social norms in society or an organization. There is more variation in the perceived seriousness of victimless crimes and less serious behaviors (Stylianou, 2003), probably because of larger differences in social norms towards such crimes than towards more violent and more serious crimes. This makes it particularly interesting to study less serious crimes.

Our study differs from the existing literature because we look at incorrect behaviors ('small crimes') that are not always condemned by the general public. These small crimes go beyond white-collar crimes committed by individuals within an organization. Our analysis is related to Halman and Luijkx (2008) who examined the public's opinion on small crimes from a social values point of view. Some of our small crimes are the same as the short descriptions used by Halman and Luijkx (2008), taken from the 1999 and 2008 waves of the European Values Study (EVS). Our approach is different in that it includes both short descriptions and hypothetical settings of specific small crimes (vignettes). This allows us to investigate the influence of offender and offense characteristics on a respondent's perception in a systematic way.

In this chapter we measure perceptions of small crime and relate these to information on crimes committed, based on a questionnaire developed by us and administered to participants of the CentERpanel, a large representative sample from the Dutch population. In the questionnaire we ask the respondents to subjectively rate the severity or justifiability of a number of small crimes. We also ask them to evaluate six small crimes presented in a setting with more (hypothetical) context. In such 'vignette' questions, several characteristics of a fictitious person committing the small crime and other factors related to the situation are included in the description.

Using survey questions to measure perceived seriousness of crime is quite common in the criminology literature (see, e.g., Rosenmerkel, 2001, or Herzog and Oreg, 2008, and the references in these studies). In the literature on the economics of crime, some studies use survey questions but many others use actual data or experimental data. The use of survey data has both advantages and disadvantages. The main advantages for our purpose are that our survey is representative for a broad population and that many background variables on the respondents are available, such as various indicators of socio-economic status (education, income, wealth). A potential disadvantage is that the respondents do not get any incentives to reveal their true opinions. On the other hand, there is no reason why they would give strategically biased answers, and the temptation to give socially desirable answers is likely to be small since the interview is an Internet survey with no personal contacts with an interviewer (see Chang and Krosnick, 2009). Moreover, there is evidence in the experimental economics literature that for relatively simple questions, respondents do not need real incentives to reveal their true preferences (see, e.g., Beattie and Loomes, 1997; Camerer and Hogarth, 1999).

The plan of the chapter is as follows. In Section 3.2 we describe the set-up and framework of the questionnaire and present descriptive statistics, including an ordering of the small crimes by their mean perceived severity. The statistical analysis of the short questions and the vignette questions is presented in Sections 3.3 and 3.4. Section 3.5 discusses some policy implications and concludes. Section 3.A provides more details on the vignette questions.

3.2 Questionnaire and descriptive statistics

The results in this chapter are based on an survey conducted in the Summer of 2008 through CentERdata at Tilburg University. CentERdata manages a panel of over two thousand 'respondents' (the CentERpanel, hereafter CP), forming a representative sample of the adult Dutch population. The sample is based on a probability sample of the non-institutionalized Dutch population of ages 16 years and older. Selected households without Internet access or without a personal computer are provided with the necessary equipment so that the sample also covers the non-Internet part of the population. Every week a questionnaire is sent out (through the Internet) to all respondents, each week on a different topic. The response rate is generally above 70%. Since respondents have typically participated in previous surveys, detailed background information is available, including gender, age, income, education, role in the household, and area of residence.

Respondents who did not respond to the survey in the first weekend were asked again a few weeks later. The combined response rate was 83% (1932 respondents). The average completion time was about thirty minutes. It seems reasonable to assume that participating and completing the questionnaire is independent of the variables of interest, conditional on several background variables (gender, age, education) that are used to construct survey weights. CentERdata constructs these weights by comparing the sample with a larger household survey administered by Statistics Netherlands. These weights will be used below in computing some of the descriptive statistics.

3.2.1 Short questions

Our survey consists of three parts. First, the respondents were asked to rate the severity of 18 offenses and the justifiability of 6 other offenses. The offenses range from taking a ballpoint from the office for private use to accepting a bribe. The wording of the questions for the first 18 offenses is:

> Below we list examples of situations that might occur in daily life. Please evaluate the severity of these actions as you perceive them on a scale from 1 (very severe) to 10 (not severe).

The other six offenses are taken from EVS; their wording is comparable but uses 'justifiability' instead of 'severity' (exactly as in EVS). Some of the types of small crime included in the survey were also used by Traxler and Winter (2012), but our list of small crimes is much longer.

Table 3.1: European Values Study (EVS) 1999 and 2008 versus CentERpanel (CP) 2008

Offense	EVS 1999 mean (std)	CP mean (std)	EVS 2008 mean (std)
Claiming government benefits to which one is not entitled	1.52 (1.28)	1.44 (1.04)	1.52 (1.33)
Accepting a bribe at work	1.60 (1.31)	1.65 (1.26)	1.55 (1.23)
Throwing away litter in a public place	1.74 (1.30)	1.98 (1.42)	
Avoiding a fare on public transport	2.79 (2.21)	2.47 (1.81)	2.58 (2.10)
Cheating on taxes if one has a chance	2.74 (2.22)	2.92 (2.14)	2.28 (1.96)
Smoking in a public building	3.81 (2.65)	2.98 (2.16)	

Answers are on a scale from 1: never justifiable to 10: always justifiable. All statistics are weighted. The number of observations N varies over studies and also (slightly) over offenses. We have 1001–1003 observations for the EVS 1999, 1929 for the CP, and 1542–1549 for the EVS 2008.

In Table 3.1 we present the means and standard deviations for the answers to the six short questions that appear in both the European Values Study and our CentERpanel survey. Two questions from EVS 1999 were not asked in EVS 2008. Applying for social benefits to which one is not entitled is considered the least justifiable of all offenses considered, followed by accepting a bribe in the course of duty. Remarkably, throwing away litter in public places also ranks quite high.

There seems to be general agreement between the CentERpanel and the EVS data for most questions. An exception is smoking in a public place, which is seen as less justifiable in the CentERpanel than in EVS 1999. This is explained by the nine-year gap between the two data sets. The perception of smoking in The Netherlands has changed in those nine years, because smoking was banned from governmental organizations in 1990 and from the private sector (including restaurants and bars) in July 2008, just after the first weekend that our survey was fielded. A widely publicized event like the introduction of a smoking ban may well lead to a (possibly temporary) change of the social norm (Ramchand et al., 2009). Comparing the two EVS waves, it appears that people consider most offenses less justifiable in 2008 than in 1999. This particularly applies to cheating on taxes. Surprisingly, the CentERpanel mean for the perceived justifiability of cheating on taxes is much closer to EVS 1999 than to EVS 2008, even though EVS and CP were conducted in the same year. Three of the six offenses in Table 3.1 (littering, fare dodging, and evading taxes) were also considered by Traxler and Winter (2012), and their ordering corresponds to what we find.

Table 3.2 describes the 18 short questions on small crimes which were not included in EVS. They are ordered according to their mean severity, from most severe to least severe. The two most severe offenses are harmful to other individuals, stressing the importance of 'harmfulness' for another private person (Rosenmerkel, 2001). Not cleaning up the dog's pooh (ranked 3) also ranks quite high, in line with the high ranking of throwing away litter in public places, the offense related to polluting the environment in Table 3.1.

Traffic violations like driving 170 km/h on a highway where the speed limit is 120 km/h, are not considered as very severe, suggesting perhaps that many people see the maximum speed rules as unnecessarily strict.

Table 3.2: Ordering of small crimes in terms of perceived severity

Offense	mean (std)
Damaging a car by accident and not informing the owner	2.10 (1.36)
Turning up the volume of music late in the evening	2.15 (1.40)
Walking the dog and not cleaning up the dog's pooh	2.71 (1.73)
Pretending to be sick and staying at home for two days	2.84 (1.90)
Driving 170 km/h on a highway (maximum is 120 km/h)	3.09 (2.13)
Leaving a barking dog alone at home	3.19 (1.78)
Taking cutlery from a canteen	3.21 (1.91)
Taking a bundle of printing paper and 5 ballpoints from the office	3.30 (2.01)
Practicing the piano in an apartment building from 7:00–10:00 am	3.47 (1.96)
Taking software from the office to install it at home illegally	3.94 (2.31)
Taking a bundle of printing paper from the office	4.09 (2.28)
Breaking a coffee mug in a store and not informing the owner	4.13 (2.10)
Making daily private phone calls from the office	4.49 (2.33)
Working two evenings per week without paying income tax	4.51 (2.34)
Driving 60 km/h within town (maximum is 50 km/h)	5.19 (2.56)
Downloading music illegally from time to time	5.98 (2.53)
Taking a ballpoint from the office	6.27 (2.70)
Taking soap and shampoo from a posh hotel room	7.03 (2.66)

Answers are on a scale from 1: very severe to 10: not severe. All statistics are weighted. The formulation of some offenses is shortened to fit the table. The full survey is available upon request.

As expected, taking away soap and shampoo from a hotel room is considered the least severe of small crimes. Most respondents do not consider this as a small crime at all, but see the soap and shampoo as a gift from the hotel. Taking a ballpoint home from the office is also one of the least severe small crimes. It is an example of 'internal fraud' and, according to Greenberg (2002), this occurs more frequently when employees feel underpaid or when employees consider the decision-making criteria as unfair. In general, offenses at the cost of the employer seem to be perceived as less severe than offenses at the cost of another individual. Downloading music illegally also appears in the bottom three of the ranking; downloading music is not illegal in The Netherlands as long as it is for private use and from a legal source, but the majority of music offered at peer-to-peer networks comes from illegal sources. Apparently, there is no strong social condemnation of digital piracy as this has no perceived social cost. This is in line with the theoretical arguments of Balestrino (2008).

3.2.2 Vignettes

In the second part of the survey we asked our respondents in 12 questions to rate the perceived justifiability of six offenses, this time described in short stories (so-called 'vignettes') concerning hypothetical persons in a hypothetical setting. The vignette questions were asked after the short questions to avoid framing effects on the short questions, which would hamper comparing the answers to the short questions with other studies. The six offenses are: (a) not having a valid (train) ticket; (b) accepting a bribe; (c) reporting a lower income than the actual income to the tax authorities; (d) breaking a coffee mug and not reporting it; (e) taking a bundle of printing paper; and (f) driving too fast on a highway.

Each of the six offenses was described in two vignettes with varying characteristics of the hypothetical person (the 'vignette person') committing the offense and of the hypothetical setting. A typical example, concerning offense (a), is:

> [Jack] is [27] years old and earns [€2500] per month before tax, a [low] wage for the type of work he does. Each day he takes the train to work, a trip of about [5] minutes. Today he is in a hurry since he does not want to arrive late at work. He jumps on the train without a valid ticket. It has [not] happened before that he knowingly did not have a valid ticket. The probability that someone will check the tickets on this route is [very small]. Do

> you think [Jack]'s behavior is absolutely not justifiable (1),. . . , always justifiable (10)?

The parts in square brackets vary across vignettes. For each situation and each respondent the offender's income is lower in the first variant than in the second, guaranteeing that the two vignettes on the same offense are always different.[2] The other parts in square brackets are randomized (independently of each other). In the example, the name of the offender is either Jack or Diana, both with probability 0.5; the offender's age is randomly set to 27, 43, or 55 years (with equal probabilities); and the absolute income level (€2500 or €3500) can be low or usual for the type of work the offender does (both with probability 0.5). The other randomizations do not concern the offender but the context in which the crime is committed: how long does the trip take (5 or 15 minutes); is the offense committed repeatedly or only once; what are the chances of getting caught (low or 50%)? Similar randomizations are used for the other vignettes. A full description of the vignette questions and the randomizations is provided in Section 3.A. The dummy variables that capture the characteristics of the offender and the circumstances in the vignettes are listed in Table 3.3. These are used as explanatory variables in our models for the vignette justifiability evaluations.

In Table 3.4 we compare the means and standard deviations of the vignette evaluations for the six offenses with the answers to the corresponding short questions. Accepting a bribe in the course of duty is considered least justified, both in the short questions and in the vignette questions. Avoiding a fare on public transport is considered less justifiable than cheating on taxes in the short questions, but this reverses in the vignette questions, where avoiding a fare is evaluated as the least serious offense of all. The opposite difference between short questions and vignette questions is found for the justifiability (severity) of breaking a coffee mug, taking a bundle of printing paper, or driving too fast on a highway. There are substantial differences between the answers to the short questions and the vignette questions.

[2] In the example, €2500 in the first variant and €3500 in the second.

Table 3.3: Binary vignette variables with explanation

vign_wage	1 if vignette person (vp) has a high wage
vign_female	1 if vp is a woman
vign_27y	1 if vp is 27 years old
vign_43y	1 if vp is 43 years old
vign_55y	1 if vp is 55 years old
vign_freq	1 if small crime has been committed more often before
vign_catch	1 if the probability of getting caught is 50% (0 if very small)
vign_distance	1 if the travel distance is 20 minutes (0 if 5 minutes)
vign_boss	1 if the boss of the vp behaves correctly
vign_entrepr	1 if the vp is an independent entrepreneur
vign_wage_hi	1 if vp receives substantial wage for type of work, given vign_wage = 1
vign_wage_us	1 if vp receives usual wage for type of work, given vign_wage = 0

Table 3.4: Mean (standard deviation) of dependent variables

	Short	Vignette
EVS questions	Justifiability	Justifiability
(a) Avoiding a fare	2.47 (1.81)	3.88 (2.33)
(b) Accepting a bribe	1.65 (1.26)	2.10 (1.59)
(c) Cheating on taxes	2.92 (2.14)	2.81 (1.96)
Our own questions	Severity	Justifiability
(d) Breaking a coffee mug	4.13 (2.10)	3.47 (2.08)
(e) Taking a bundle of printing paper	4.09 (2.28)	3.19 (2.05)
(f) Driving too fast on a highway	3.09 (2.13)	2.73 (1.96)

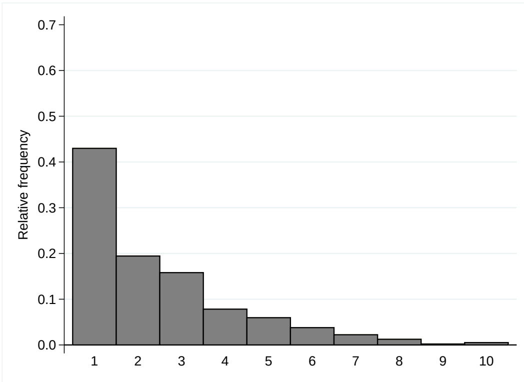
Answers are on a scale from 1: very severe/never justifiable to 10: not severe at all/always justifiable. All statistics are weighted. N varies between 1929 and 1932 for short questions and between 3840 and 3846 for vignette questions.

There may be several reasons for this. Since the vignette questions provide more information about the context in which the offense is committed, one explanation is that context matters. This is in line with Riedel (1975) who asked respondents to rate the importance of offense and offender characteristics for judging the seriousness of a described offense. He concluded that respondents need external factors to make a judgement. On the other hand, Rossi et al. (1997) found that the offender's background only has a small impact on sentencing preferences. How context matters will be studied in detail in Section 3.4. An alternative explanation for differences between the ratings of short questions and vignettes might be framing effects: there are many small crimes in the short questions, and it is likely that respondents try to rank these with their ratings. On the other hand, there are only six small crimes in the vignettes. This may explain differences in the absolute ratings, but it seems implausible that it explains the observed reversal of some of the average ratings.

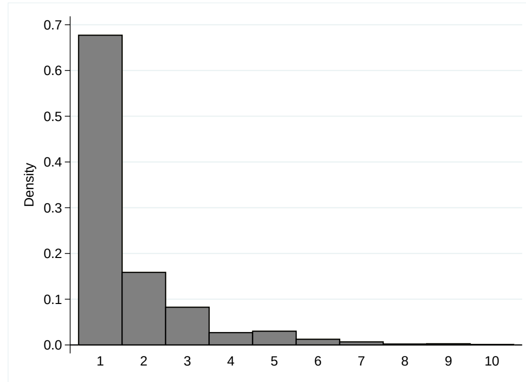
The sample standard deviations in the answers to the short questions and the vignette questions are of similar size; two of the six standard deviations are larger for the vignette questions; the other four are larger for the short questions. Herzog (2003) argued that when judgements are based on less information regarding the circumstances of the crime (e.g. offender characteristics) respondents will make quick judgements based on shared norms in a society, which would suggest that the dispersion in the answers to the short questions would be smaller than for the vignette questions. We do not find any such evidence in Table 3.4.

Figures 3.1 and 3.2 provide the complete distributions of the answers. We have almost twice as many observations (3840) for the vignette questions as for the short questions (1930), because the respondents evaluated two vignette questions for each type of offense. As explained above, the income of the person committing each offense is always lower in the first vignette than in the second vignette (while other characteristics are randomized).

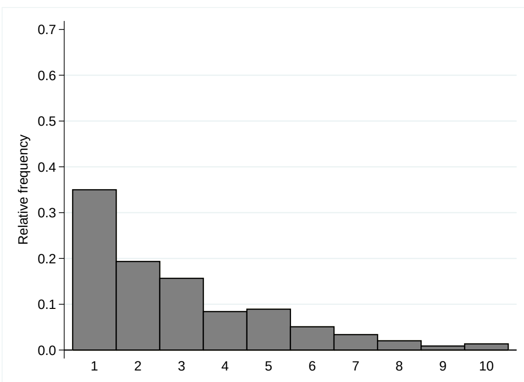
Figure 3.1: Answers to selected short questions (items (a)–(c) refer to justifiability; items (d)–(f) to severity)



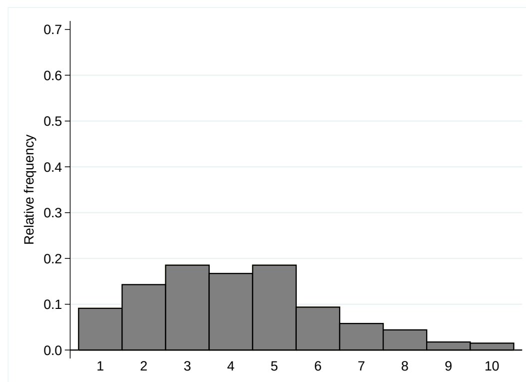
(a) Avoiding a fare



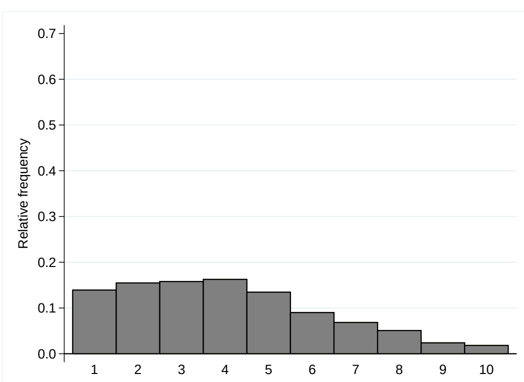
(b) Accepting a bribe



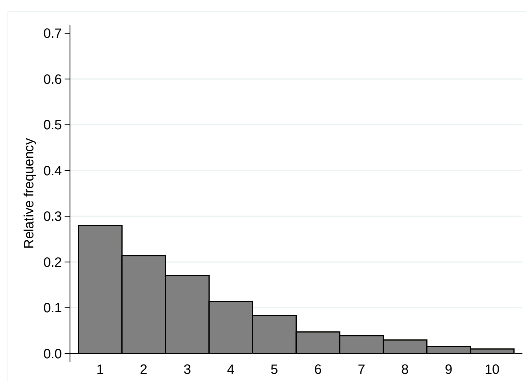
(c) Cheating on taxes



(d) Breaking a coffee mug

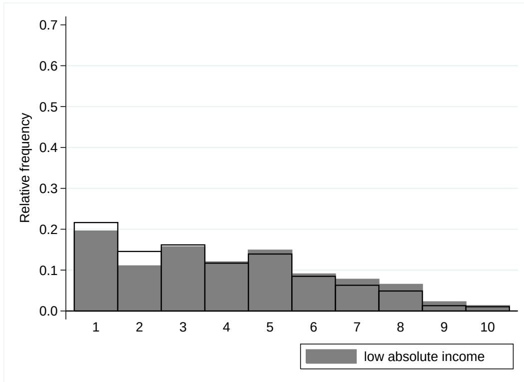


(e) Taking a bundle of printing paper

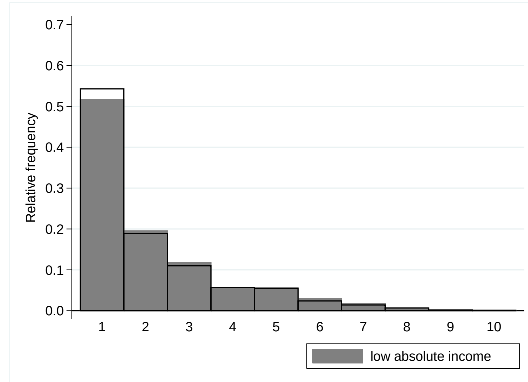


(f) Driving too fast on a highway

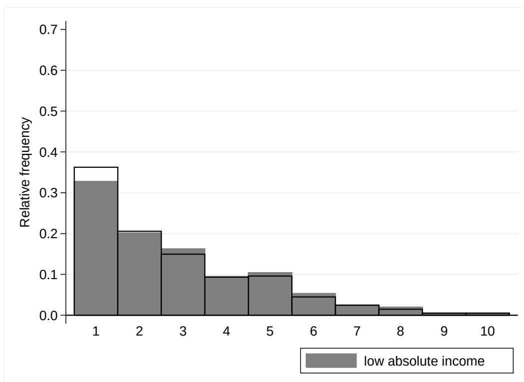
Figure 3.2: Answers to vignette questions (all items refer to justifiability)



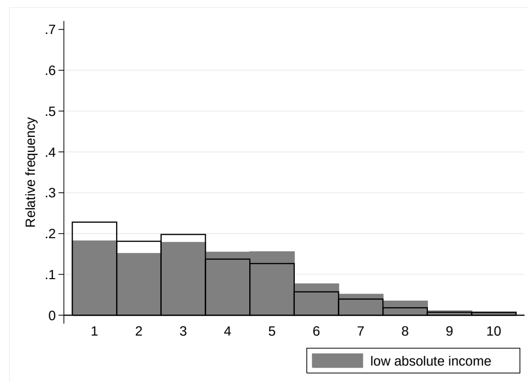
(a) Avoiding a fare



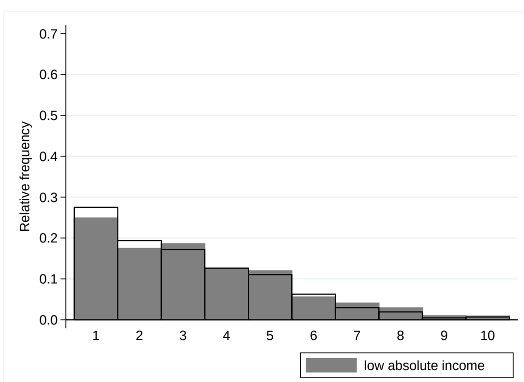
(b) Accepting a bribe



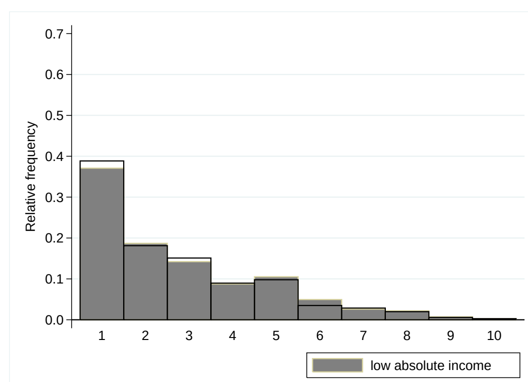
(c) Cheating on taxes



(d) Breaking a coffee mug



(e) Taking a bundle of printing paper



(f) Driving too fast on a highway

Figure 3.2 shows separate histograms for the answers to these two questions, clearly illustrating that respondents tend to perceive an offense as more severe if the income of the person committing the offense is higher.

3.2.3 Respondent characteristics

The respondent characteristics used as explanatory variables are presented in Table 3.5 (definitions and descriptive statistics). Roughly 47% of the sample is female. The age of the respondents ranges from 15 to 93 with a mean of 51. Highly-educated respondents are overrepresented: 36% completed higher vocational school or has a university degree in our sample as compared to 25% in the population in 2006 (Statistics Netherlands, 2008). This is because the higher educated have a larger probability to participate in the CentERpanel. We use sample weights to correct for this.

To capture the effect of how familiar respondents are with crime, we include *crime_rate* (the number of registered crimes per capita) at the provincial level, which varies from 4.6% to 9.0%. Within a given province, crimes are more common in cities than in rural areas. Hence, we also include the degree of urbanization. About 41% of our respondents live in cities, 20% in larger towns, and 39% in small towns or villages.

It is likely that one's occupational status influences one's perception of crime. For example, employees may be more sympathetic than others to someone taking a bundle of printing paper from the office for private use, because they are more familiar with this kind of situation. We distinguish between four types of occupations. The largest group (48%) contains those in paid employment (*occup_empl*).

The majority of the respondents (62%) are head of a household. In about 67% of all cases, household heads live together with a partner (married or unmarried). Being head of a household or the partner of the household head may imply that one's behavior is an example to the rest of the household, which may lead to a different attitude to (small) crimes. About four out of five respondents reported that they support a national political party; the

Table 3.5: Respondent variables with explanation

Variable	Mean (std)	Explanation
Non-binary variables		
age	50.7 (16.1)	age of respondent (in years)
hh_lincome	7.93 (1.43)	log of gross monthly household income
crime_rate	7.31 (1.22)	% number of crimes in total population per province
Binary variables		
female	0.47 (0.50)	1 if respondent is a woman
edu_prim	0.07 (0.25)	1 if respondent's highest education is primary school
edu_secon1	0.26 (0.44)	1 if — lower general secondary school
edu_secon2	0.12 (0.33)	1 if — higher general secondary school
edu_vocat1	0.19 (0.39)	1 if — intermediate vocational school
edu_vocat2	0.24 (0.43)	1 if — higher vocational school
edu_univer	0.12 (0.32)	1 if — university
urban_low	0.39 (0.49)	1 if respondent lives in a less urbanized area
urban_high	0.41 (0.49)	1 if — more urbanized area
urban_middle	0.20 (0.40)	1 if — an intermediate urban character
occup_empl	0.48 (0.50)	1 is respondent has an (unpaid) job
occup_pension	0.23 (0.42)	1 if — is retired or $\geq$ 65 years
occup_indep	0.05 (0.21)	1 if — works as independent entrepreneur or in a family firm
occup_nowork	0.24 (0.43)	1 if — has no occupation (incl. students)
position_head	0.62 (0.49)	1 if respondent is head of the household[a]
partner	0.79 (0.41)	1 if head of household has a partner (married or unmarried)
party_nochr	0.59 (0.49)	1 if respondent votes for non-Christian national political party
party_christ	0.22 (0.42)	1 if — Christian national political party
party_other	0.19 (0.39)	1 if — local party or does not vote

[a]The 'head' is the person who owns the house or signed the rental contract; if this applies to more than one person, then the one with the highest personal income is the head.
Statistics are not weighted. N varies between 1918 and 1931.

others support a local party or do not feel affiliated with any political party. Of those supporting a national party, about one-quarter supports a Christian party. We included a dummy for supporting a Christian party as a proxy for ethical norms and values that may possibly affect attitudes towards (small) crime.

Finally, we asked some questions about the respondent's past victimization incidence and exposure to crimes in daily life. These questions are not analyzed in the current chapter. The complete survey is available upon request.

3.3 Models

We analyze the determinants of the perceived justifiability (and severity in the case of some of the short questions) of small crimes using econometric models. We focus on explaining the answers to the vignette questions, from respondent characteristics, offender characteristics, and other variables describing the context of the offense. In addition, we consider models explaining the answers to the six short questions on the same types of offenses described in the vignettes from respondent characteristics only. This is in order to investigate to which extent providing the context changes the conclusions about the association between perceived seriousness and respondent characteristics. We model the answers to the short questions and the vignette answers to each of the six offenses separately. Since the response scale is discrete and ordered, ranging from 1: never justifiable to 10: always justifiable — or from 1: very severe to 10: not severe at all — we use ordered probit models: a standard ordered probit model for each of the short questions, and a panel-data version of this model for the vignettes.

3.3.1 Model for short questions

The model for each of the short questions describes the reported evaluation as the category containing the value of an unobserved (latent) continuous variable y_i^*, which is driven by a vector of explanatory variables x_i (respondent characteristics, in our case) and an error term ϵ_i:

$$
\begin{aligned}
y_i^* &= x_i'\beta + \epsilon_i, \\
y_i &= j \quad \text{if} \quad \alpha_{j-1} < y_i^* \leq \alpha_j,
\end{aligned}
$$

where

$$\epsilon_i \sim \mathrm{N}(0,1) \text{ independent of } x_i,$$

and $i = 1, \ldots, N$ denote the respondents, and $j = 1, \ldots, m$ are the possible values that y_i can have. We set $m = 10$ and let $\alpha_0 = -\infty$ and $\alpha_m = \infty$.

3.3.2 Model for vignette questions

The fact that each respondent answers two vignette questions on each offense (with different values of the randomized vignette variables; see Section 3.2) allows us to use a random-effects panel-data ordered probit model with $T = 2$ 'time periods':

$$\begin{aligned} y_{it}^* &= x_{it}^{'}\beta + \epsilon_{it}, \quad i = 1, \ldots, N, \quad t = 1, 2, \\ y_{it} &= j \quad \text{if} \quad \alpha_{j-1} < y_{it}^* \leq \alpha_j, \quad j = 1, \ldots, m, \end{aligned}$$

where

$$\begin{aligned} \epsilon_{it} &= u_i + \upsilon_{it}, \\ u_i &\sim \mathrm{N}(0, \sigma_u^2), \text{ independent of } x_{i1}, x_{i2}, \upsilon_{i1}, \upsilon_{i2}, \\ \upsilon_{it} &\sim_{iid} \mathrm{N}(0, \sigma_\upsilon^2), \text{ independent of } x_{i1}, x_{i2}. \end{aligned}$$

Again, we set $m = 10$ and let $\alpha_0 = -\infty$ and $\alpha_m = \infty$. Without loss of generality we normalize σ_ϵ^2 $(= \sigma_u^2 + \sigma_\upsilon^2)$ to 1. For the explanatory variables in x_{it}, we distinguish between respondent characteristics (income, age, gender, education, occupational status), characteristics of the vignette person committing the crime, and variables describing the context in which the crime is committed. This allows us to disentangle the effects of respondent characteristics and characteristics of the offender on the perceived severity of each offense. Note that vignette characteristics vary with i and t, while respondent characteristics vary with i only.

The model is estimated by maximum likelihood, integrating out the random effects. The random effects capture the correlation between the unobservable components in the two vignette questions for each individual, and this correlation is automatically taken into account in computing the standard errors (so that accounting for clustering is not needed).

3.4 Results

In the baseline model for the vignette questions, x_{it} includes the respondent characteristics that are also used for the short questions (gender, age, household income, education, the crime rate in the province of residence, and the urbanization rate), as well as the vignette characteristics. Because of the design, there is some variation in vignette characteristics across the six situations. An example is *vign_boss*, capturing the effect on perceived justifiability if the boss of the vignette person behaves correctly under the same circumstances. This variable is only included in two of the six situations.

We also estimated models with interactions. For example, it might be that the difference between perceived justifiability of a young and an older person committing an offense varies with the age of the respondent, or it could be the case that the effect of income of the offender on the seriousness perception is different for respondents with lower or higher income. Such interactions, however, were hardly ever significant and adding them did not lead to additional insights. Since the interactions also make it harder to interpret the results, we decided to only present the results of the models without interactions.

The estimation results for the short questions are presented in Table 3.6, and the results for the baseline model of the vignettes are in Tables 3.7a and 3.7b. We focus on the results for the vignettes and the differences between the effects (of respondent characteristics) according to the vignette evaluations and the short questions.

3.4.1 Respondent characteristics

We first consider the respondent characteristics. Some of the earlier studies focus on measuring the degree of consensus between different demographic

Table 3.6: Ordered probit on short questions

	Situation					
Variable	1	2	3	4	5	6
female	−0.1887***	−0.2170***	−0.1430***	−0.3282***	−0.1572***	−0.2653***
	(0.0509)	(0.0477)	(0.0481)	(0.0496)	(0.0575)	(0.0498)
age	−0.0153***	−0.0111***	−0.0216***	−0.0192***	−0.0123***	−0.0012
	(0.0017)	(0.0016)	(0.0016)	(0.0016)	(0.0018)	(0.0015)
hh_lincome	−0.0101	−0.0059	−0.0273*	−0.0034	0.0111	−0.0191
	(0.0186)	(0.0185)	(0.0162)	(0.0167)	(0.0212)	(0.0184)
crime_rate	0.0155	0.0254	0.0234	0.0242	0.0026	0.0557***
	(0.0217)	(0.0205)	(0.0200)	(0.0204)	(0.0244)	(0.0216)
edu_secon1	−0.0330	−0.0662	−0.0549	0.1592	−0.1906*	−0.1030
	(0.1085)	(0.1048)	(0.1063)	(0.1019)	(0.1120)	(0.1012)
edu_secon2	−0.0693	−0.0520	−0.1256	0.1782	−0.3303***	−0.1198
	(0.1198)	(0.1121)	(0.1179)	(0.1117)	(0.1243)	(0.1137)
edu_vocat1	−0.0718	−0.1210	−0.0518	0.3209***	−0.1598	−0.2684***
	(0.1090)	(0.1070)	(0.1075)	(0.1025)	(0.1128)	(0.1041)
edu_vocat2	−0.2870***	−0.1863*	−0.1638	0.1592	−0.6597***	−0.4281***
	(0.1074)	(0.1032)	(0.1054)	(0.0993)	(0.1133)	(0.1032)
edu_univer	−0.1151	−0.1348	−0.0707	0.2409**	−0.5960***	−0.2650*
	(0.1186)	(0.1115)	(0.1130)	(0.1130)	(0.1305)	(0.1112)
urban_high	0.0307	0.1348**	0.1006*	−0.2780***	−0.1212*	−0.1982***
	(0.0589)	(0.0546)	(0.0557)	(0.0582)	(0.0673)	(0.0583)
urban_middle	0.0133	0.0865	0.0788	−0.1731***	−0.0500	0.0248
	(0.0703)	(0.0668)	(0.0658)	(0.0655)	(0.0758)	(0.0684)
N	1914	1917	1917	1917	1914	1914

$^{***} = \{p < 0.01\}$; $^{**} = \{0.01 \leq p < 0.05\}$; $^{*} = \{0.05 \leq p < 0.10\}$
Standard errors in parentheses.
Situations: 1 = not having a valid (train) ticket; 2 = breaking a coffee mug; 3 = taking a bundle of printing paper; 4 = driving too fast on a highway; 5 = accepting a bribe; 6 = reporting a lower income to the tax authorities.

groups (Kwan et al., 2002; O'Connell and Whelan, 1996; Rossi et al., 1974; Sellin and Wolfgang, 1964), since public consensus is required to develop a generally supported seriousness scale of criminal activities. Differences

Table 3.7: Random effects ordered probit

(a) Respondent characteristics

	Situation					
Variable	1	2	3	4	5	6
female	−0.1775	−0.4754***	−0.3359***	−1.4000***	−1.2307***	−0.9644***
	(0.1105)	(0.1142)	(0.1135)	(0.1350)	(0.1142)	(0.1259)
age	−0.0087***	−0.0082**	−0.0430***	−0.0509***	−0.0416***	−0.0237***
	(0.0032)	(0.0039)	(0.0036)	(0.0042)	(0.0034)	(0.0037)
hh_lincome	−0.0067	−0.0476	−0.0272	−0.1745***	−0.0924***	−0.1762***
	(0.0378)	(0.0477)	(0.0298)	(0.0364)	(0.0340)	(0.0269)
crime_rate	0.0880**	0.2405***	0.1691***	0.3897***	0.2232***	0.3639***
	(0.0446)	(0.0432)	(0.0642)	(0.0570)	(0.0450)	(0.0487)
edu_secon1	−0.1647	−0.8979***	−0.1432	−0.0977	−0.6907***	−0.2853
	(0.3054)	(0.2229)	(0.2766)	(0.2764)	(0.1839)	(0.2572)
edu_secon2	0.0167	−0.4125*	−0.3468	0.4387	−1.3166***	−0.3959
	(0.3286)	(0.2436)	(0.3007)	(0.2875)	(0.2252)	(0.2761)
edu_vocat1	−0.1069	−0.8834***	−0.0484	0.2829	−1.3366***	−0.8246***
	(0.3065)	(0.2167)	(0.3069)	(0.2517)	(0.1862)	(0.2629)
edu_vocat2	−0.2527	−0.8612***	−0.4725*	−0.6029**	−2.2939***	−1.3780***
	(0.3208)	(0.2196)	(0.2762)	(0.2403)	(0.2173)	(0.2868)
edu_univer	0.3684	−1.4330***	−0.2069	−0.1554	−2.6853***	−1.9023***
	(0.3327)	(0.2530)	(0.2965)	(0.2738)	(0.2152)	(0.2703)
urban_high	−0.1307	0.0079	0.1733	−1.8075***	−0.0109	−0.8207***
	(0.1227)	(0.1219)	(0.1274)	(0.1769)	(0.1156)	(0.1429)
urban_middle	0.2656*	−0.0118	0.0984	−1.1946***	0.0501	0.1816
	(0.1424)	(0.1356)	(0.2282)	(0.1780)	(0.1604)	(0.1557)
N	3816	3812	3810	3810	3810	3810
ρ	0.8382	0.8813	0.8564	0.9258	0.9166	0.9176

$^{***} = \{p < 0.01\}$; $^{**} = \{0.01 \leq p < 0.05\}$; $^{*} = \{0.05 \leq p < 0.10\}$
Standard errors in parentheses.
Situations: 1 = not having a valid (train) ticket; 2 = breaking a coffee mug;
3 = taking a bundle of printing paper; 4 = driving too fast on a highway;
5 = accepting a bribe; 6 = reporting a lower income to the tax authorities.

Table 3.7: Random effects ordered probit (cont.)

(b) Vignette characteristics

Variable	Situation 1	2	3	4	5	6
vign_wage	−0.3496***	−0.5564***	−0.3154***	−0.2604***	−0.3403***	−0.3813***
	(0.0594)	(0.0637)	(0.0632)	(0.0700)	(0.0798)	(0.0688)
vign_female	0.0014	−0.0937*	0.0095	0.1205**	−0.0878	−0.1146**
	(0.0499)	(0.0527)	(0.0520)	(0.0590)	(0.0610)	(0.0568)
vign_43y	0.0683	0.1153*	0.0875	0.1039	−0.0330	0.0030
	(0.0594)	(0.0631)	(0.0625)	(0.0717)	(0.0759)	(0.0733)
vign_55y	0.0937	0.1683***	0.0584	0.0176	−0.0189	0.0246
	(0.0599)	(0.0625)	(0.0622)	(0.0761)	(0.0743)	(0.0699)
vign_freq	−1.2838***		−0.4162***	−0.5922***	−0.3840***	−0.2644***
	(0.0531)		(0.0518)	(0.0624)	(0.0616)	(0.0570)
vign_catch	−0.0390		−0.1340***			−0.2631***
	(0.0484)		(0.0508)			(0.0556)
vign_distance	−0.0719					
	(0.0482)					
vign_boss			−0.5716***		−0.3429***	
			(0.0519)		(0.0608)	
vign_entrepr				0.0138		
				(0.0622)		
vign_wage_us	−0.0941	−0.0117	−0.1123	−0.1171	−0.1473*	−0.1626**
	(0.0684)	(0.0726)	(0.0730)	(0.0879)	(0.0873)	(0.0787)
vign_wage_hi	−0.0520	−0.1416*	−0.0974	−0.0830	0.0658	−0.1025
	(0.0685)	(0.0728)	(0.0730)	(0.0878)	(0.0885)	(0.0798)

*** = $\{p < 0.01\}$; ** = $\{0.01 \leq p < 0.05\}$; * = $\{0.05 \leq p < 0.10\}$
Standard errors in parentheses.
Situations: 1 = not having a valid (train) ticket; 2 = breaking a coffee mug; 3 = taking a bundle of printing paper; 4 = driving too fast on a highway; 5 = accepting a bribe; 6 = reporting a lower income to the tax authorities.

between groups were studied by Rosenmerkel (2001), who also looks at a larger set of respondent characteristics, including detailed indexes of socio-economic status. We interpret our results as follows.

Gender: Women consider the offenses less justifiable than men, especially regarding driving too fast on a highway. According to the short questions as well as the vignette questions, women perceive all six small crimes as more serious than men with the same characteristics (that is, the same age, education, household income, urbanization rate, and provincial crime rate). This is in line with the results reported by Herzog and Oreg (2008), O'Connell and Whelan (1996), and Rossi et al. (1985), and may be due to the fact that women are more vulnerable and have a stronger fear of being victimized (Warr, 1984). On the other hand, Kwan et al. (2002) find a gender effect only for crimes that disproportionately affect women, and Isenring (2008) finds no gender effect on the perceived seriousness of white-collar crimes. Kwan et al. (2002) find that bribery (similar to our situation 5) is rated as more serious by men than by women. Orviska and Hudson (2002) find that women are more likely to approve tax evasion (specifically, value-added tax), which is in contrast to our result for situation 6 (reporting a lower income to the tax authorities). The large differences in magnitude across offenses in Table 3.7a, much larger than in Table 3.6, suggest a violation of relative consensus. For example, speeding on the highway will be higher in the seriousness ranking for women than for men.

Age: The signs and significance levels for the short and vignette questions largely correspond; older respondents always give significantly more severe ratings in all situations. For tax evasion, the negative age effect is significant and larger in magnitude than for some of the other small crimes in the vignette questions, while it was insignificant in the short questions. The negative age effects are in line with Orviska and Hudson (2002) and O'Connell and Whelan (1996); older people may have stricter social norms than younger people, perhaps due to different behavior of their peer group (Traxler, 2010).

Income: In the short questions, we find no significant income effects. But in the vignette questions, household income has a negative and significant effect in three of the six situations: respondents with a higher household income perceive driving too fast, accepting a bribe, and tax evasion as more serious than low-income respondents. This is in contrast to the findings of Rossi et al. (1985), who report that higher income is associated with *more* tolerance towards white-collar crimes. On the other hand, Rosenmerkel (2001) found no income effect on white-collar crime, and reports that respondents with higher income considered violent crimes as less serious than lower income respondents. Again, the most likely reason for the income effect seems differences in social norms, probably in relation to differences in peer groups.

Education: In the vignette questions, educational dummies are jointly significant in five of the six situations. More education leads to harsher evaluations. These effects are quite different from those in the short questions, where no clear pattern can be found, although educational dummies are jointly significant in four out of six cases. The strongest effect is found for bribery followed by tax evasion, particularly according to the vignette questions: higher-educated respondents rate tax evasion as much more severe than the lower educated. This is in line with Orviska and Hudson (2002), who also find that a higher education level increases disapproval of tax evasion. This suggests that the social norm to disapprove tax evasion is stronger for the higher educated. Our results for the short questions are closer to Rossi et al. (1985), who also find an inconsistent pattern of the effect of education on the perception of different types of crime. That a higher education would lead to *less* harsh judgements is found by Rossi et al. (1974), Isenring (2008), Payne et al. (2004), O'Connell and Whelan (1996), and, for white-collar crime, Schrager and Short (1980). We find this only for the short question on situation 4 (driving too fast).

Crime rate: Respondents in provinces with higher crime rates judge less harshly than respondents in provinces with lower crime rates. The effect is significant in all six situations for the vignette regressions, but only in one situation for the short questions. The size of the effect varies. According to the vignette questions, the effect is highest for driving too fast and for tax evasion, and lowest for using public transport without a valid ticket. The significant effect of the crime rate may seem surprising. Respondents who live in areas with a higher crime rate are expected to be more familiar with serious crime, and this may, indirectly, also affect their social norm concerning small crime. On the other hand, the provincial crime rate might also proxy other differences in social norms across provinces, particularly between the more densely populated North-West of the country (where the crime rate is higher) and the rest of the country.

Urbanization: Living in an urbanized area may have an effect on the perception of crime through social norms. Moreover, crime rates are higher in large cities than in smaller towns or rural areas (Glaeser and Sacerdote, 1999). Since we include the crime rate by province but not by municipality (since we do not have the data on the latter), this implies that the degree of urbanization can be seen as a proxy for within-province variation in the exposure to crime. Important is also that people in cities tend to more tolerant than people in the country, not only on crime but also on many other issues. Offenses 4 and 6 (speeding and tax evasion, which are among the more serious of the small crimes considered), are considered less serious by respondents living in a (highly) urbanized area, both in the short and in the vignette questions. This is in line with Rose and Prell (1955) who discuss the effect of urbanization on 'punitiveness' and find that respondents who do not live in an urban area think that punishments should be harsher than respondents in urban areas. Stylianou (2003) also cites several studies that find an effect of the degree of urbanization on other social norms, such as abortion. On the other hand, no significant effects (at the 5% level) are found for the other vignettes and in the short question on traveling without a ticket, we even find an unexpected effect in the opposite direction. Apparently, if social norms concerning small crime vary with degree of urbanization, this does not apply to all small crimes in the same way.

We also considered extensions of the baseline model for the vignette evaluation with more respondent characteristics (respondent's occupation, position within a household, and preference for a Christian political party). The latter two were only included for the offenses where they played a significant role. Adding these additional characteristics leaves the effects of the respondent and vignette characteristics in the baseline model virtually unchanged, and we therefore only present and discuss the effects of the additional respondent characteristics in the extended model (Table 3.8).

Table 3.8: Random effects ordered probit: extended specification

	Situation					
Variable	1	2	3	4	5	6
occup_pension	0.0522	0.0410	0.1931	0.0990	0.4226**	−0.2838
	(0.1641)	(0.1724)	(0.1947)	(0.1654)	(0.1989)	(0.1791)
occup_indep	0.7616***	0.8261	0.7554***	1.2896***	1.6483***	0.9740***
	(0.2042)	(0.5943)	(0.2578)	(0.2014)	(0.2008)	(0.2355)
occup_nowork	−0.2611**	−0.0872	0.0908	−0.6818***	0.1566	0.6460***
	(0.1328)	(0.1608)	(0.1726)	(0.1613)	(0.1731)	(0.1610)
position_head					0.0537	0.1817
					(0.2110)	(0.1937)
partner					0.3436	0.8216***
					(0.2175)	(0.2191)
party_christ						−0.4256***
						(0.1283)
party_other						0.4221***
						(0.1288)
N	3816	3812	3810	3810	3810	3806
ρ	0.8406	0.8825	0.8559	0.9319	0.9139	0.9214

*** = $\{p < 0.01\}$; ** = $\{0.01 \leq p < 0.05\}$ Standard errors in parentheses.
The extended specifications include the same respondent and vignette characteristics as in Tables 3.7a and 3.7b.
Situations: 1 = not having a valid (train) ticket; 2 = breaking a coffee mug; 3 = taking a bundle of printing paper; 4 = driving too fast on a highway; 5 = accepting a bribe; 6 = reporting a lower income to the tax authorities.

Occupational status: Self-employed respondents are significantly less harsh on five types of small crime than employees, while pensioners are less harsh in only one situation. The latter result is not in line with Herzog and Oreg

(2008) who find that part-time employees consider crimes relatively less justifiable than full-timers. Wärneryd and Walerud (1982) find no effect of self-employment or occupation on the attitude towards tax evasion.

Political party: The final additional variable is affiliation with a Christian political party. The literature is ambiguous on this issue. Herzog and Oreg (2008)) found that individuals who lead a conservative life also have more conservative views towards crime. Similarly, Payne et al. (2004) reported that conservativeness is positively related to the tendency to punish harder. On the other hand, Isenring (2008) did not find a significant effect of political preferences on crime seriousness ratings. We find no significant effect either, with one exception: respondents who feel attached to a Christian party rate tax fraud as a more serious offense than other respondents.

3.4.2 Vignette characteristics

In 1996 the Catholic Dutch Bishop Tiny Muskens declared that the poor have a right to steal bread when they are hungry and see no other way to survive. This statement caused some turmoil, especially in the bakery industry, but was also applauded, and some years later Bishop Muskens was appointed Honorary Citizen of Breda. We find that the most salient effect of the vignette characteristics is the effect of the vignette person's earnings level (*vign_wage*). For all situations, respondents consider the offense less justifiable if the person who commits it earns more. The explanation is probably that the respondents feel that people with higher income can better afford to be honest. The coefficients for this variable are of approximately the same size, except for situation 2 (breaking a coffee mug in a shop), for which the effect is by far the largest, and situation 4 (speeding) for which the effect is lowest.

In addition to the absolute earnings level, each vignette situation also provides information on how earnings compare to those of others with a similar job. This information depends on the earnings level: if the earnings level is high, then the vignette states either 'this income is usual for this type of work' or 'this income is high for this type of work' ($vign\_wage\_hi = 1$). If absolute earnings are low, the vignette states either 'this income is low for this type of work' or 'this income is usual for this type of work' ($vign\_wage\_us = 1$). A negative sign on both $vign\_wage\_us$ and $vign\_wage\_hi$ implies that respondents are harsher if earnings of the offender are relatively high, given the type of work. It seems that relative income matters more if the offender's absolute income is low than if it is high: the coefficient of $vign\_wage\_us$ is significant in two situations; that of $vign\_wage\_hi$ in only one situation. These effects are much smaller than those of absolute earnings. Perhaps surprisingly, the relative wage level plays no significant role for the only work-related situation (taking a bundle of printing paper home).

As expected, if a vignette person has committed the same crime before ($vign\_freq = 1$), it is considered less justifiable than if the crime is committed for the first time. The effect is significant in all five situations where this information is provided. This finding that people are generally harsher if the offense is repeated corresponds to the results of Herzog and Oreg (2008) and Rossi et al. (1985).

Important is also the probability that the offender gets caught. A larger probability to get caught ($vign\_catch = 1$) leads to a harsher judgement, and the effect is significant in two out of three cases. An explanation could be that a small probability to get caught (for example in evading taxes) suggests that the offense is taken less seriously by society, so that the respondent interprets it as a proxy for the social norm. According to the theory of expressive law, the expression of social values is an important, perhaps the most important, function of the courts (Cooter, 1998). See also Kube and Traxler (2011) who focus on the interaction of formal (legal) and informal (social) enforcement of compliance with the law.

The behavior of the offender's superior also matters. The superior sets an example to the employees and influences the norms within the organization. If the superior behaves correctly (e.g. does not take printing paper home for private use), then the respondents think it is less justifiable for the employees to behave incorrectly and consider the offense significantly more severe. This type of behavior is referred to as 'parallel deviance', where unethical behavior on the part of a superior sends a message to an employee that deviant behavior is legitimate or even the standard within an organization (Greenberg and Scott, 1996; Jones and Kavanagh, 1996). Jones and Kavanagh (1996) find that unethical behavior of the superior significantly raises intentions to behave unethically in one of their two experiments.

The effects of other vignette characteristics are specific to the situation. Older offenders are judged significantly less harshly than others when breaking a coffee mug in a shop and not reporting it (situation 2). Differences between ratings of small crimes committed by male and female offenders are insignificant in four situations, and marginally significant with opposite signs in the other two situations. These results are not in line with those of Rossi et al. (1985) who find, in the case of property crimes, that older offenders are judged more severely than young offenders, and females are judged more mildly than males.

3.5 Concluding remarks

There are many studies on the perception of crime. The studies typically consider serious crimes such as murder and armed robbery and sometimes also white-collar crimes. The literature on the perceived justifiability of small crime or incorrect behavior is, however, small. This chapter tries to fill this gap. An analysis of the perception of small crime at the individual level is of interest because it tells us something about the social norms held by different socio-economic groups in society, and social norms play a crucial role in many recent models of economic and social behaviors.

In this chapter we have tried to disentangle the factors that drive perceptions of small crime using data on subject, offender, and offense characteristics. One of the strengths of the chapter is the quality and quantity of the data. We had access to an excellent panel, representative for a broad population and with a high response rate, and we were able to ask almost 2000 respondents many questions on incorrect behavior of which some activities are forbidden by law while other activities are not forbidden but can be perceived as morally wrong.

A methodological novelty of our approach is that we use vignette questions to incorporate characteristics of the offender and the context in which the offense is committed. Our results comparing vignettes and short questions (Tables 3.6 and 3.7a) confirm that respondents evaluate a given (small) crime differently if they know more about the offender and the circumstances. From a methodological point of view, this means that the analysis through vignettes is useful, even if we are only interested in how the social norms vary across socio-economic groups.

We find interesting effects of the context variables, showing that social norms concerning crime not only depend on the crime itself but also on the context in which it is committed. The respondents judge a small crime committed by an underprivileged person less harshly than the same offense committed by a wealthy person. Not everyone would agree with Bishop Muskens that a poor man is allowed to steal bread, but income does play a role in people's judgment. This is true even for non-financial crimes such as speeding (see Table 3.7b, situation 4). If this is indeed the public's sentiment, then one may wonder why punishments are not income-dependent. It is not unusual to make company fines dependent on the revenue earned in a certain period, for example when breaking competition laws. Income-dependent fines for individuals are not common in The Netherlands, although they do exist in some other European countries, such as Germany and Switzerland. This study does not discuss the implications for deterrence. For example, if lower sanctions were applied to less well-off individuals, this would send a signal to other similarly placed individuals thinking about the offense. Our findings do not necessarily allow conclusions about law enforcement, despite the fact that some results (for example about repeat offenders) can be related to the law.

No doubt, one can learn much from the experiences in other countries. The current study considers only The Netherlands. Evans and Scott (1984) compared perception in two different cultures: United States and Kuwait. While violent, property, and white-collar offenses were perceived similarly, moral offenses (selling illegal drugs, prostitution, having an illegal abortion, committing perjury) were perceived very differently. A new international study involving more countries would be of great interest.

Various other extensions could also be of interest. It is likely that past victims of a (small) crime judge more harshly than subjects who have never been a victim; see the discussion on the effect of victimization on a subject's judgment in Pease (1988). Hence, including a measure of victimization may provide additional insight. In addition, a multivariate approach would identify factors driving a subject's judgment in general, hence not only in a specific situation. Finally, it would be interesting to compare the survey answers with actual behavior, for example in experiments. The fact that, for example, older people perceive small crimes as more serious than younger people, might reflect differences in interpreting the answering scales—older people might more easily call something 'severe' instead of really having a different attitude. This is an issue that has not been addressed in the survey literature on crime perception, but is prominent in subjective evaluations of aspects of well-being such as health or political efficacy (see, e.g., King et al., 2004).

3.A Vignette questions

In the vignette part of the questionnaire we consider six offenses. For each offense we study two variants. Hence we ask twelve vignette questions. [3] In all cases we randomize over men and women (and adjust the name accordingly), and over age (27, 43, or 55 years old). In the first variant, income is either low or usual (randomized) for the type of work that the vignette person does. In the second variant, income is set higher and is either usual or high (randomized) for the type of work. This is the only difference between the two variants. At the end of each question we ask whether the vignette person's behavior is absolutely not justifiable (1),..., always justifiable (10) on a scale from 1 to 10. Below we give one example for each of the six offenses, each time for the first variant (low income). Randomizations other than those mentioned above are italicized and explained.

Not having a valid (train) ticket: Jack is 27 years old and earns €2500 per month before tax, a low wage for the type of work he does. Each day he takes the train to work, a trip of about *5 minutes*. Today he is in a hurry since he does not want to arrive late at work. He jumps on the train without a valid ticket. It has *not* happened before that he knowingly did not have a valid ticket. The probability that someone will check the tickets on this route is *very small*. [There are three additional randomizations: travel time is either 5 or 20 minutes; it has not happened before or it has happened often; and probability of detection is very small or 50%.]

Breaking a coffee mug and not reporting it: Anne is 27 years old and earns €1335 per month before tax, a low wage for the type of work she does. While shopping in a department store, she accidentally drops a coffee mug, priced at €4. Anne puts the broken mug back and leaves the store without informing the owner about the accident. [No additional randomizations.]

[3] In fact, we ask fourteen questions, but two of these are not analyzed in this chapter.

Taking a bundle of printing paper: John is 27 years old and works at an office. He earns €1335 per month before tax, a low wage for the type of work he does. *John has noticed that his boss occasionally takes printing paper home for private use.* John takes a bundle of printing paper home for private use. *This is the first time that he does this.* The probability that someone will notice it is *very small*. [Three additional randomizations: ‘John has noticed that his boss occasionally takes printing paper home for private use’ or ‘John’s boss is a principled man and never takes things home from work for private use’; this is the first time or John does it often; and probability of detection is very small or 50%.]

Driving too fast on a highway: Sandra is 27 years old and earns a living by delivering packages *in her own car*. She earns €1750 per month before tax, a low wage for the type of work she does. On her way to a client she drives 170km/h on a highway where the maximum speed limit is 120km/h. It has *not happened before* that Sandra drove so fast on a highway. [Two additional randomizations: Sandra either has her own car or she works for a big courier company; and it has not happened before or it often happened before.]

Accepting a bribe: Patrick is 27 years old and works as a civil servant in a municipal department responsible for building permits. He earns €2000 per month before tax, a low wage for the type of work he does. Patrick’s boss is *known to occasionally accept gifts from building firms*. Patrick accepts a gift from someone applying for a building permit, in exchange for speeding up the procedure. This is the *first time* that Patrick does this. [Two additional randomizations: Patrick’s boss is either known to occasionally accept gifts from building firms or he is a principled man and does not accept gifts; and this is the first time or Patrick often accepts gifts.]

Reporting a lower income to the tax authorities: Linda is 27 years old and works freelance. She earns €2500 per month before tax, a low wage for the type of work she does. To the tax authorities she reports €2000 per month. This is the *first time* that Linda does this. The probability that the tax authorities check Linda's tax return is *very small*. [Two additional randomizations: This is the first time or Linda has been doing this for several years; and probability of detection is very small or 50%.]

Chapter 4

Peer Reporting and the Perception of Fairness[1]

4.1 Introduction

A young boy goes to a supermarket and sees an expensive pen which he likes a lot. He puts the pen in his pocket and walks out of the shop, but the shop assistant has seen him, grabs him, and hands him over to the police. At the police station, the boy's father is called and appears.

> *Father:* Son, why did you do this?
> *Boy:* I liked the pen so much!
> *Father:* But you know you should not steal.
> *Boy:* I liked the pen so much!
> *Father:* Why did you not tell me? I could have brought one for you from the office.

It is the father, rather than the son, who is of interest in this story. Apparently he finds taking a pen from the shop bad, but taking the same pen from his work not. Why not?

[1] This chapter is based on Douhou et al. (2011a).

Becker (1968) would explain this by saying that the expected monetary loss caused by being caught is smaller than the gain obtained by having the pen. This can be viewed as the traditional economic approach. But there are many additional or alternative views. Maybe the father's office lacks normative pressure (social norms). Normative pressure triggers guilt and shame, and this may prevent criminal activities (Weibull and Villa, 2005). A recent field study which relies on the morality of its customers is the honor-based flower picking business in the Black Forest in Germany (Schlüter and Vollan, 2011). Classical economic theory would predict that this market would break down, but it does not, even though serious money is involved. So, here is a preference for honesty in a situation where it is difficult or impossible to detect cheating. This is closely related to 'conditional cooperation': people are more likely to comply when a larger population fraction adheres to the norm (Traxler, 2010; Traxler and Winter, 2012; Weibull and Villa, 2005).

Maybe the father feels it is fair to take a pen from the office. Greenberg (1990) and Houser et al. (2011) showed that if a situation (like a pay-cut) is perceived as unfair, employees are more likely to cheat. Honesty is affected by perceptions of fairness. Or perhaps, the father works in a disorderly environment. This is the 'broken windows theory', which suggests that a disorderly environment triggers petty crime. An experiment by Keizer et al. (2008) showed that this may indeed be the case. The father may well work in a large firm. Gneezy (2005) suggested that fraudulent behavior in a large organization is considered less severe than against individuals, even if the monetary damage is similar, because the consequences of the deception are valued differently.

To take a pen from the office to give to your son is a small crime, a misdemeanor, an example of incorrect behavior. In the current chapter we study another small crime, namely to take home a bundle of printing paper from the office for private use. Employing our 2008 "Incorrect behavior in every day life" survey taken from a Dutch household panel with about 2000 respondents, two central questions drive our current study: how 'justifiable' do you (the respondent) find the behavior of someone at the office taking paper home?; and, if this person were your colleague, would you report this behavior? If so, how? If not, why not?

The answers to these questions will depend on many things. They will depend on who the person is taking printing paper home (the offender): age, gender, income, and whether the offender does this often or not. They will depend on the situation: does the offender's boss also take paper home for private use or not, is it likely that someone catches the offender or not. And they will depend on who the respondent (the reporter) him/herself is: age, gender, income, education, religion, living in town or not, his/her own history as a 'small criminal', whether the respondent has been a victim of a small or large crime, and some information on the respondent's trust and social norms. All these factors will play a role in our analysis.

In order to answer the question what determines whether the respondent would peer report or not, a major modeling issue arises, namely the fact that one of the explanatory variables (justifiability) may be endogenous, because both peer reporting and justifiability are choices of the same individuals. To solve this endogeneity issue, we propose an instrumental-variable-like approach (not exactly instrumental variables because the model is not linear). We introduce 'instruments', show that these are valid, and estimate a four-equation panel data probit model with random individual effects.

This modeling issue is one of the distinguishing features of the current chapter. In addition, unlike most of the existing literature, we combine characteristics of the reporter, the offender, and the 'small crime' with justice evaluation and information on a respondent's past victimization. Finally, our data set consists of a large representative sample of the Dutch population and is not limited to students or employees of a specific organization.

Studies in the area of peer reporting and whistleblowing have investigated, *inter alia*, factors related to the individual, the situation, the organization, social context, justice evaluation, and ethical ideology and religion. Sims and Keenan (1998) analyzed a sample of 248 full-time employees enrolled in an undergraduate or graduate business program and found that external whistleblowing was significantly related to supervisor support, informal policies, gender, and ideal values. Victor et al. (1993) used a field survey in a fast food restaurant to test the influences of social context (role responsibility and interests of group members) and justice evaluations on the respondent's inclination to report theft and the actual theft-reporting behavior. Trevino and Victor (1992) found support for a positive relation between the extent to which the offender damages the interest of group members and the inclination to peer report. King and Hermodson (2000) analyzed actual peer reporting of unethical behavior by colleagues in a sample of 197 registered nurses and found that the observer's individual characteristics, situational factors such as severity of the misdemeanor, as well as organizational issues like compliance or non-compliance with policy and procedures played a significant role. Barnett et al. (1996) analyzed peer reporting of academic cheating, focusing on the role of religion and ethical ideology, and found a positive association between peer reporting and religiosity among 267 American business students.

The structure of the remainder of this chapter is as follows. In Section 4.2 we briefly describe the survey design and the elements of the survey relevant for the current chapter. Some descriptive statistics are provided and discussed in Section 4.3. The econometric method is explained in Section 4.4. Our main equation is an equation for peer reporting, in which justifiability of the committed offense is one of the explanatory variables. To allow for confounding unobserved factors correlated with justifiability as well as peer reporting, we treat justifiability as endogenous and estimate an equation for justifiability jointly with the equation for peer reporting. Estimation results are discussed in Section 4.5. Section 4.6 concludes. Section 4.A gives details on the definitions of respondent and vignette variables used in the analysis.

4.2 Survey design

The CentERdata research institute at Tilburg University manages a panel of over two thousand 'respondents' (the CentERpanel), who participate in an online websurvey on a weekly basis, each week on a different topic. Respondents are randomly selected from a population register. If they do not have a computer with Internet access, they are provided with the necessary equipment. Detailed background information on the respondents is available from prior surveys and the response rate is generally high. Our 'small crime' survey was conducted in the Summer of 2008. A total of 1932 panel members completed the survey, amounting to a response rate of about 83%. The respondents form a representative sample of the Dutch population, aged 16 years and older.

We briefly describe the structure of the survey; a more detailed description can be found in Douhou et al. (2011b) where the same data source is used. The complete questionnaire (in Dutch) is available upon request from the authors. Our survey was divided into three blocks of questions. The first block consisted of a set of 24 small offenses, ranging from taking a ballpoint from the office for private use to accepting a bribe. The respondents were asked to rate the severity of 18 offenses and the justifiability of six other offenses.

In the second block we concentrated on six offenses: (i) not having a valid (train) ticket, (ii) breaking a coffee mug and not reporting it, (iii) taking a bundle of printing paper for private use, (iv) driving too fast on a highway, (v) accepting a bribe, and (vi) reporting a lower income than the actual income to the tax authorities. This time the offenses were described in short stories ('vignettes') concerning hypothetical persons in a hypothetical setting. Each of the six offenses was described in two vignettes with varying characteristics of the hypothetical person (the 'vignette person') committing the offense, and of the hypothetical setting. Vignettes have often been used in the social sciences. They were first introduced in economics by Van Beek et al. (1997) in the context of employer evaluations of hypothetical job applicants. An advantage of vignettes is that the characteristics (of offenses and offenders, in our case) are part of the design, making it possible to create large exogenous variation within and across respondents. Moreover, using hypothetical offenses rather than offenses actually experienced by the respondents avoids endogeneity problems (which would arise if characteristics of actually experienced offenses are correlated to unobserved respondent characteristics) as well as selection problems (possibly arising if a specific group of respondents has never experienced the type of offense). The use of vignettes makes it therefore much easier to obtain consistent and relatively efficient estimates of how justifiability and peer reporting vary with offense and offender characteristics.

A typical example (concerning offense (iii)) is:

> Anne is 27 years old and works at an office. She earns €1335 per month before tax, a low wage for the type of work she does. Anne has noticed that her boss occasionally takes printing paper home for private use. Anne takes a bundle of printing paper home for private use. This is the first time that she does this. The probability that someone will notice it is very small. Do you think Anne's behavior is never justifiable (1),..., always justifiable (10)?

In the first variant of this vignette question the vignette person (Anne) earns €1335; in the second variant €2500. Both variants were put to the respondents in the survey. Other items were randomized. In this case, the following six aspects of the vignettes were randomized:

- *Gender:* Anne or John;
- *Age:* 27, 43, or 55 years old;
- *Boss:* occasionally takes printing paper home for private use, or is a principled man and never takes things home from work for private use;
- *Frequency:* this is the first time or Anne does it often;

- *Catch:* probability of detection is very small or 50%;
- *Wage:* low or average if wage is €1335; average or high if wage is €2500.

The associated randomized binary vignette variables are presented in more detail in Section 4.A, Table A.1. Note that each respondent sees two vignettes for each crime, and that in all of these pairs the first vignette always presents a low-income person and the second vignette a high-income person. Since the order of the income levels was not randomized, there might be a 'demand effect': Respondents realize that income varies between vignettes and feel that they should react by adjusting their responses. We cannot test the existence of this effect, but speculate that the repetitive sequencing of the income levels made the low versus high income treatment variation quite salient to respondents.

In this chapter we concentrate on the above vignette question on taking a bundle of printing paper from the office, because it was the only one that was followed by a question on reporting behavior, phrased as follows:

> Suppose Anne/John is your colleague, would you report this behavior?

The respondents could then choose from the following options:

- Yes,
  - I would talk with Anne/John about it (1)
  - I would talk with my colleagues, but not with my boss (2)
  - I would immediately report this behavior to my boss (3)
  - I would report this to someone else (4);
- No,
  - because I am worried about the reaction of my colleagues (5)

  - because I am worried about my position within the company (6)
  - because I don't know to whom to report this behavior (7)
  - because this is too futile to worry about (8)
  - for some other reason (9).

The third block was designed to provide more detailed background information of the respondents. The following two questions about past victimization are particularly relevant:

- Have you been a victim of a serious crime in the past five years (i.e., burglary, holdup, violence, or something similar)?
- Have you been a victim of 'incorrect' behavior in the past five years?

If either question is answered 'yes', then a follow-up question asks to rate the severity of the most serious crime on a scale from 1: very severe to 10: not severe. We used this information to construct an index of self-reported severity of past victimization. The reason that we only ask about the past five years is to avoid a bias towards older respondents that have a higher probability of being victimized. Note that there is a subtle difference between seriousness and severity of a crime. Seriousness reflects our judgment, while severity reflects the judgment of the respondent. In the questionnaire, 'incorrect' behavior is defined as an infringement or misdemeanor which carries (almost) no punishment, but disadvantages others, such as the government, the employer, co-users of the road, or the neighbors. Since 'incorrect' behavior ranges from stealing a pen to smoking in a public place, it is highly unlikely that a respondent has never been a victim of this type of behavior. Still, only about one quarter of the respondents reported being a victim of incorrect behavior, suggesting that the answer reflects the respondent's attitude or sensitivity towards social norm violations.

Since peer reporting may be associated with trust in other people (Trevino and Victor, 1992), we used a trust index as one of our explanatory variables. Questions on trust were not included in our survey, but they were asked to the same panel of respondents in another CentERpanel survey, conducted around the same time, entitled 'Victims of (attempt to) fraud' (Oudejans and Vis, 2008). This survey was merged with our own data to obtain an index for trust. Three questions were used to construct *trust_index*:

- Would you say that most people can be trusted or that you cannot be too careful in dealing with people? Please answer on a scale from 1: you have to be careful to 11: most people can be trusted;
- Do you think that most people would try to take advantage of you if they got the chance, or would they try to be honest? Please answer on a scale from 1: most people would try to make advantage of me to 11: most people would try to be honest; and
- Would you say that most of the time people try to be helpful or that they are mostly looking out for themselves? Please answer on a scale from 1: people look mostly of themselves to 11: people try to be helpful.

4.3 Descriptive statistics

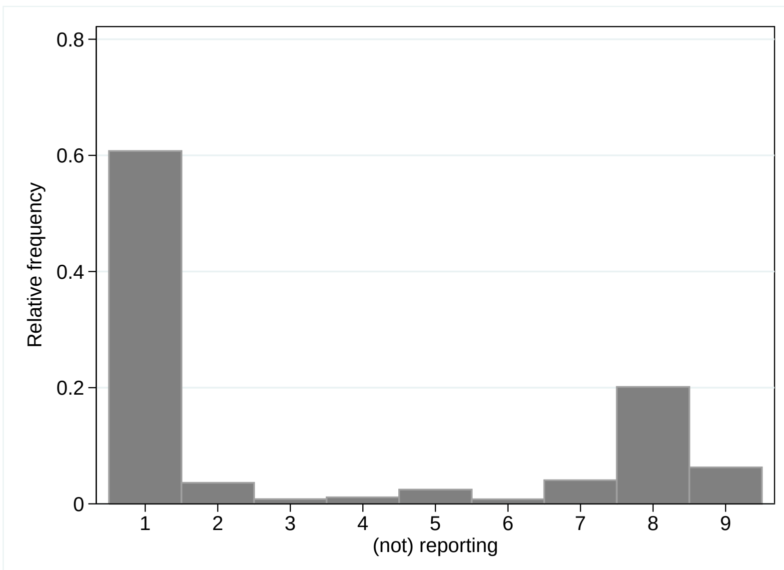
Descriptive statistics of the respondent variables used in our analysis are presented in Table 4.1. Peer reporting and justifiability are the main variables of interest (and the dependent variables in our econometric model); the other variables are used as explanatory variables for peer reporting, justifiability, or both. The corresponding variable definitions are listed in Section 4.A, Table A.2. We mentioned in Section 4.2 that the response rate is high, namely 83%. Still, the non-respondents may have an effect on the estimates due to selection bias. Upon further investigation we find that the average age of the non-respondents is 44.9 (50.68 for the respondents), *urban_middle* is 0.25 (0.20 for the respondents), and *hh_lincome* 7.79 (7.93 for the respondents). A probit regression of key respondent characteristics on the binary response variable confirms these results. Older people, in particular, are overrepresented in our sample.

Table 4.1: Descriptive statistics — respondent characteristics

Binary			*Non-binary*		
	Mean	N		Mean	N
female	0.47	1931	age	50.68	1931
edu_middle	0.39	1924	hh_lincome	7.93	1931
edu_high	0.55	1924	vict_index	1.87	1919
urban_high	0.41	1924	trust_index	21.69	1635
urban_middle	0.20	1924	social_norm	7.01	1929
religion	0.58	1932	justifiability*	3.19	3840
victim_small	0.25	1919			
victim_serious	0.12	1919			
takematerial	0.33	1919			
peer_report*	0.66	3840			

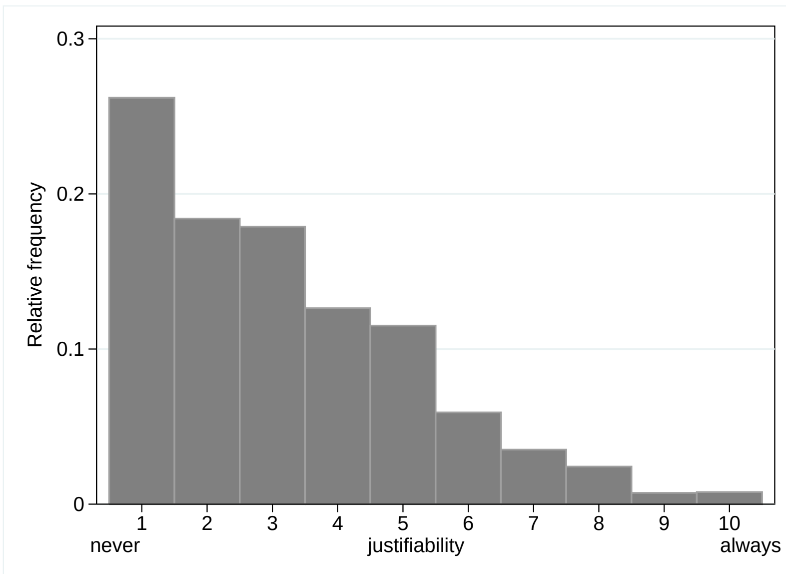
* Dependent variable

Figure 4.1: Peer reporting



Our principal dependent variable is *peer_report*. About 66% of the respondents would report a colleague if this colleague would take a bundle of printing paper from the office for private use. As explained in Section 4.2, labels 1–4 in Figure 4.1 refer to the situation where the respondent decides to report, while labels 5–9 refer to the situation where the respondent does not report. Most respondents, if they report, choose to talk to the offender (label 1). If respondents choose not to report the offense, it is usually because they find the offense too futile to worry about it (label 8).

Figure 4.2: Justifiability

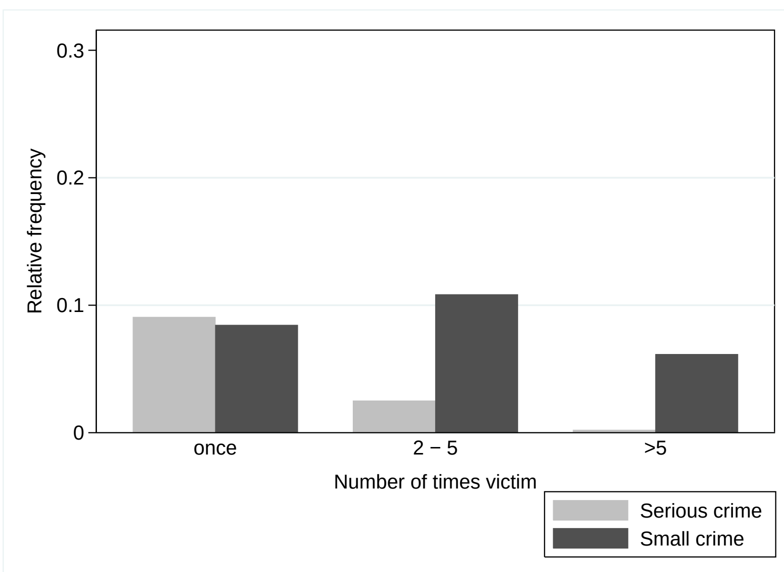


Our second variable of main interest (used both as a dependent variable and as an explanatory variable for peer reporting) is *justifiability*, and Figure 4.2 presents its empirical distribution. The mean and median are around 3. Since a low value of justifiability means that the respondent does not find the action justifiable, the figure shows that most respondents disapprove of taking a bundle of printing paper home. Some authors claim that it is the perceived severity of a small crime rather than its justifiability which should play a role in the analysis (King, 1997; King and Hermodson, 2000). The relationship between justice evaluations and the severity of a small crime was discussed by De Graaf (2010) based on interviews performed with employees of public organizations. He shows that the two concepts are closely related.

The explanatory variables include a set of basic socio-economic and demographic characteristics. The age of the respondents ranges from 15 to 93 with a mean of 51 (Table 4.1). Median household income before tax was about €2780 per month. A slight majority of the respondents is male and has at least a degree from an intermediate vocational school (*edu_high=1*). About 41% live in more urbanized areas (cities, *urban_high=1*).

The other explanatory variables are specific to the current analysis. There are three variables relating to victimization. In our sample of 1932 respondents, 488 (25%) reported that they had been victim to a 'small' crime (*victim_small*) in the past five years, and 226 (12%) that they had been victim to a 'serious' crime (*victim_serious*) during the same period. The range of 'incorrect' actions is wide, and this makes it unlikely that someone has never been a 'victim' of incorrect behavior. The fact that only one quarter of the respondents reported being a victim of incorrect behavior therefore suggests that the answer may not only reflect victimization, but also the respondent's susceptibility to harm or injustice.

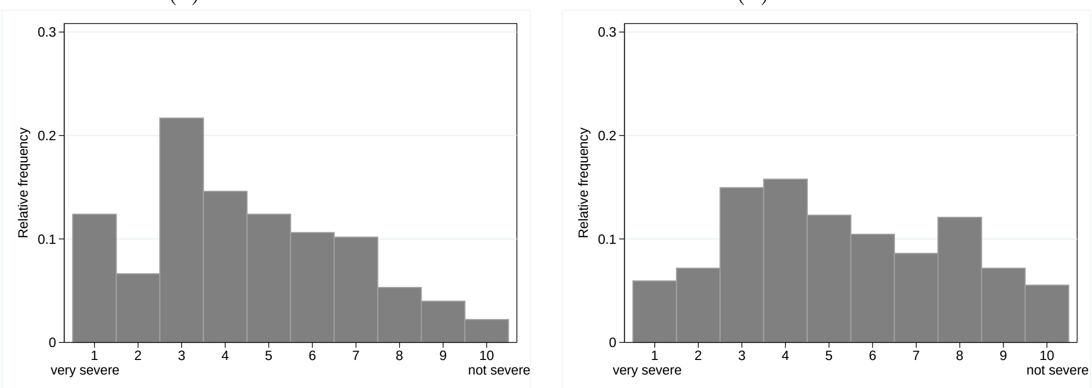
Figure 4.3: Degree of victimization



In Figure 4.3 we consider only respondents that have been a victim at least once. The figure shows that people who have been a victim of a serious crime in the past five years typically experienced a serious crime only once, while the empirical distribution of the number of small crimes is more evenly spread. If a respondent reported having been victim of a crime (small or

Figure 4.4: Severity of victimization

(a) Serious crime (b) Small crime



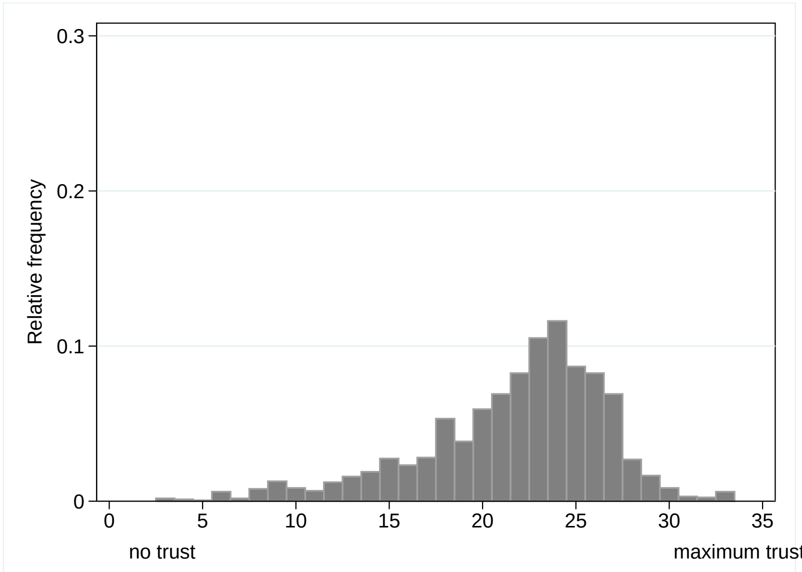
serious) in the past five years, then the perceived severity of this crime (or the worst of them, if they experienced more than one) was also asked (on a ten-point scale: 1 is very severe, 10 is not severe). Figure 4.4 shows that a few victims of a serious crime judge the crime to be very severe (1 or 2), while most respondents find the crime rather severe (mode is 3), and only a few do not find the crime severe at all. For small crimes the distribution is more even, as one would expect. The average severity of a small crime is 5.3 (median is 5), and of a serious crime 4.5 (median 4). We constructed an index for the degree of severity of victimization from these two variables (*vict_index*) ranging from 0 (not a victim of any crime) to 20 (victim of both small and serious crime and both rated as very severe).

Respondents were also asked three questions relating to their own criminal behavior. In particular, they were asked about shoplifting, taking materials from work for private use, and claiming government benefits they were not entitled to. Few respondents reported that they had committed these crimes (which may or may not be truthful), with the exception of taking work material home for private use (the variable *takematerial*). One third of the respondents admitted having done this at least once, and 26% at least twice. This variable is of interest because it relates closely to the vignette question used in our analysis, and it allows us to verify whether the respondents' own incorrect behavior in a similar situation is associated with their action in the hypothetical situation.

Ethical judgements of a situation and the reaction to it can also be influenced by religious views, social norms, and trust. The literature on moral attitudes suggests that religious people hold more traditional views on moral issues than non-religious people (Barnett et al., 1996). There is reason to believe that people with a religion may respond differently to an unethical act (in this case: taking a bundle of printing paper from the office for private use). About 58% of our respondents reported being religious (interpreted in a broad sense). Regarding social norms, we constructed a social norms index as the average of the responses on severity (on a scale from 1: not severe to 10: very severe) of a list of 18 offenses that differ in the level of damage caused; see Table 2 in Douhou et al. (2011b) for the 18 questions and the mean answer to each of them. The overall mean (and the mean of our index) is 7.01. A low value of the index means that the respondent considers small crimes as less severe, indicating a lower value placed on social norms.

Finally, a variable measuring how much trust the respondent has in other people can be important for one's actions and beliefs in general (Deutsch, 1958), and for peer reporting in particular (Trevino and Victor, 1992). The variable *trust_index* is constructed as the sum of three variables, formulated at the end of Section 4.2, that measure several aspects of a person's trust, each on a scale from 1 to 11 (a higher value means more trust), so that the trust index ranges from 3: very low trust to 33: maximum trust level. Since these questions come from a different CentERpanel survey, they were asked in a different week, and therefore they were not answered by all respondents

Figure 4.5: Trust



who answered our peer reporting and justifiability questions. This explains why for this variable we have fewer observations.[2] Figure 4.5 with a mode of 24 and a mean of 21.7 shows that respondents on the whole seem to have trust in others.

4.4 Models

Each respondent i answers questions on two vignettes describing taking home a bundle of printing paper from work for private purposes. In the first variant ($t = 1$) the offender's income is €1335; in the second variant ($t = 2$) it is €2500. In addition, several other aspects of the vignettes differ in a randomized way, as described in Section 4.2. Our main dependent variable is peer reporting (*peer_report*, y_{it}), and this is a binary variable: respondents choose to report ($y_{it} = 1$) or not to report ($y_{it} = 0$) the offense for each of the two vignettes. Observations on different respondents i are all assumed to be

[2]Respondents who answered the trust questions but did not participate in our small crime survey are not included.

independent of each other, but it is very likely that there is a positive correlation between the two answers of the same respondent ($t = 1$ and $t = 2$), and we shall take this correlation explicitly into account.

For this purpose, we use the following bivariate probit model (which is similar to a panel data probit model with random individual effects, where $t = 1$ and $t = 2$ are the (two) time periods):

$$\begin{aligned} &y_{it}^* = \beta_0 + x_{it}^{'}\beta + \delta z_{it} + \epsilon_{it} \qquad (i = 1, \ldots, N; \;\; t = 1, 2); \\ &y_{it} = 1 \text{ if } y_{it}^* > 0, \quad y_{it} = 0 \quad \text{if} \quad y_{it}^* \leq 0. \end{aligned} \tag{4.1}$$

In our specification there are 21 regressors in the model: the constant term, 19 regressors $\{x_{it}\}$ (vignette characteristics and respondent characteristics and attitudes), and the justifiability assessment z_{it}, which plays a special role (see below). Regarding the unobserved error terms ϵ_{it} we assume that

$$\epsilon_i = \begin{pmatrix} \epsilon_{i1} \\ \epsilon_{i2} \end{pmatrix} \sim_{iid} \mathrm{N}_2(0, \Sigma), \qquad \Sigma = \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix},$$

and also that ϵ_i is independent of x_{it}. The specification implies that $\mathrm{var}(\epsilon_{i1}) = \mathrm{var}(\epsilon_{i2})$; the fact that both are equal to one is a harmless normalization. The parameter ρ_1 is expected to be positive since ϵ_{i1} and ϵ_{i2} contain a common individual-specific component (a random individual effect in panel data modeling terminology).

In our first model, given in Equation (4.1), we assume that justifiability z_{it} is exogenous. This exogeneity assumption may, however, be criticized, since both justifiability and peer reporting are choices of the same individuals, and it seems plausible that there are unobserved confounding factors — unobserved variables that have an influence on both justifiability and peer reporting. This leads to a correlation between z_{it} and ϵ_{it}, making justifiability potentially endogenous. In a linear model it would be natural to use an instrumental variables approach to deal with the endogeneity problem. Our approach is similar in terms of identifying assumptions, but because of the nonlinear nature of the model, we do not use instrumental variable estimation as such. Instead, we add equations for assessed justifiability of the two vignette offenses and estimate these equations jointly with the equations for peer reporting (using maximum likelihood). By allowing for arbitrary correlations between the error terms of the peer reporting and the justifiability equations, we allow z_{it} to be endogenous in the equation for y_{it}.

To identify the model (other than through functional form assumptions), we have to exclude at least one variable from the equation for y_{it} that appears in the equation for z_{it}. For this purpose, we include three vignette variables (a vector w_{it}, our 'instruments') in the justifiability equation that are not included in Equation (4.1): two dummies describing the relative wage of the vignette person (*vign_wage_low* and *vign_wage_high*) and the probability of getting caught given in the vignette (*vign_catch*). These instruments indeed contribute to explaining justifiability of the offense described in the vignette (see Section 4.5), giving them enough power to serve as instruments. The key identifying assumption that makes these three variables suitable instruments is that they do not to have a direct effect on peer reporting (keeping justifiability constant). This seems a plausible assumption. There is no apparent reason why there should be such a direct effect. Note that these variables are part of the randomized design (they are vignette characteristics and not respondent characteristics), so that they are by construction independent of the unobserved confounding factors leading to correlation between z_{it} and ϵ_{it}. This also applies to the other vignette variables, but these might have a direct effect on peer reporting. For example, behavior of the supervisor (*vign_boss*) may matter since a respondent may decide not to peer report if the behavior of the supervisor indicates that the incorrect behavior is apparently common in the organization, even though justifiability does not change. For the three variables in w_{it} no such argument applies.

The equation for justifiability is specified as follows:

$$\begin{aligned} z_{it}^{*} &= x_{it}^{'}\alpha + w_{it}^{'}\gamma + \zeta_{it} && (i = 1, \ldots, N;\;\; t = 1, 2), \\ z_{it} &= j \quad \text{if } \lambda_{j-1,t} < z_{it}^{*} \leq \lambda_{j,t} && (j = 1, \ldots, 10;\;\; t = 1, 2), \end{aligned} \tag{4.2}$$

where

$$\zeta_i = \begin{pmatrix} \zeta_{i1} \\ \zeta_{i2} \end{pmatrix} \sim_{iid} \mathrm{N}_2(0, \Omega), \qquad \Omega = \begin{pmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{pmatrix},$$

and ζ_i is assumed to be independent of (x_{it}, w_{it}). Again, there is no loss of generality in normalizing the Ω matrix. Like ρ_1, we expect ρ_2 to be positive, because of an individual-specific component in both justifiability assessments. We allow ζ_i to be correlated with ϵ_i. More precisely, we assume that the vector $(\epsilon_{i1}, \epsilon_{i2}, \zeta_{i1}, \zeta_{i2})'$ is multivariate normal with variances normalized to one and with unrestricted correlation coefficients $\rho_{st} = \mathrm{corr}(\epsilon_{is}, \zeta_{it})$. Since unobserved respondent characteristics that are associated with a stronger tendency of peer reporting are likely to be also associated with harsher assessments of the vignette offenses, that is, to lower scores on the justifiability scale (which runs from never justifiable to always justifiable), we expect the four ρ_{st} correlations all to be negative.

The six correlations ρ_1, ρ_2, and ρ_{st} $(s, t = 1, 2)$ are auxiliary model parameters to be estimated, as well as the thresholds $\lambda_{j,t}$ $(j = 1, \ldots, 9;\ t = 1, 2)$. We set $\lambda_{0,t} = -\infty$ and $\lambda_{10,t} = \infty$. By means of normalization, there is no constant term in (4.2). The four equations (4.1) and (4.2) are estimated jointly by maximum likelihood using Roodman's (2009) conditional mixed process (CMP) routine.

4.5 Results

We present the estimation results in Tables 4.2 (for the equation with justifiability as the dependent variable) and 4.3 (for the equation in which peer reporting is the dependent variable). In the second and third columns of Table 4.3, labeled 'exogeneity', we assume that justifiability is exogenous and explain peer reporting from the bivariate probit model (4.1) with exogenous z_{it}. In the fourth and fifth columns, labeled 'endogeneity', we allow justifiability to be endogenous and present the estimates of the peer reporting equation in the complete model given by (4.1) and (4.2). Table 4.2 reports the estimates of the justifiability equation in this complete model. Table 4.4

presents the estimated correlation structure of the error terms in the complete model.

The number of observations is always 1615, which is lower than the number of respondents to our survey because we also used data from another survey (see Sections 4.2 and 4.3), and not all respondents of our small crime survey participated in this other survey.

From the three tables, we can draw three broad conclusions. First, most of the exogenous variables have both a direct and an indirect (via justifiability) effect on peer reporting. Second, the correlations between the error terms of (4.1) and (4.2) in Table 4.4 are negative and significant, confirming our hypothesis that justifiability should be treated as an endogenous variable. Third, in spite of this finding, the differences between the estimates of the peer reporting equation allowing and not allowing for endogeneity of justifiability are generally rather small. We also note that ρ_1 and ρ_2 are close to one and that $\rho_{st} \approx -0.2$ in all four cases, irrespective of whether $s = t$ or not (Table 4.4). This suggests that the individual effects play a much larger role than the vignette-specific idiosyncratic error terms.

4.5.1 Justifiability

Although our main interest is in the peer reporting estimates (the second column in Table 4.3), let us briefly consider Table 4.2, which reports the estimates when justifiability is the dependent variable.

The behavior of the boss is important: if the offender's boss behaves incorrectly according to the vignette, then the offense is considered more justified. First-time offenders are evaluated less harshly. When the probability of getting caught is higher, the incorrect behavior is considered less justified. If the offending employee in the vignette receives a relatively low wage for the work he or she does, the offense is considered more justifiable than if the employee receives a usual or high wage (keeping other variables constant, including the absolute wage level). Both of these variables (two of the three variables used as instruments in the peer reporting equation, see Section 4.4)

Table 4.2: Regression results — justifiability

vign_female	0.014	(0.024)
vign_43y	0.046	(0.029)
vign_55y	0.045	(0.029)
vign_boss	−0.253***	(0.024)
vign_freq	−0.188***	(0.024)
vign_catch	−0.064***	(0.024)
vign_wage_low	0.073**	(0.034)
vign_wage_high	−0.022	(0.034)
female	0.032	(0.052)
age	−0.001	(0.002)
hh_lincome	0.002	(0.019)
edu_middle	−0.036	(0.107)
edu_high	−0.116	(0.105)
urban_high	0.028	(0.057)
urban_middle	−0.040	(0.068)
religion	−0.001	(0.051)
vict_index	−0.007	(0.016)
trust_index	−0.015***	(0.005)
social_norm	−0.487***	(0.022)
victim_small	−0.121	(0.101)
victim_serious	0.020	(0.101)
takematerial	0.280***	(0.058)

*** = $\{p < 0.01\}$; ** = $\{0.01 \leq p < 0.05\}$; * = $\{0.05 \leq p < 0.10\}$ Standard errors in parentheses. Dependent variable is *justifiability*.

are significant and the three instruments are also jointly significant, confirming that our instruments have sufficient predictive power (conditional on the exogenous variables x_{it}) for the justifiability variable that is instrumented.

Neither having been a victim of a serious or a small crime, nor the victimization index are significant, so that victimization has no apparent influence on the justifiability assessments (keeping other variables constant). As expected, own involvement in employee theft (*takematerial*) is associated with judging the hypothetical offender more lightly. A lower score on the social norm index implies that a respondent considers small crimes as relatively less severe. Respondents with higher trust in others (a higher score on the variable *trust_index*) also tend to assess the offenses in the vignettes significantly more harshly.

4.5.2 Peer reporting

In discussing the estimates of the peer reporting equation in Table 4.3, we distinguish between three types of explanatory variables, following the analysis of Mesmer-Magnus and Viswesvaran (2005) in the context of whistleblowing: characteristics of the offense, context of the offense, and characteristics of the reporter. Before we discuss these types, one by one, we comment briefly on the validity of our 'instruments'.

Table 4.3: Regression results — peer reporting

	Exogeneity		Endogeneity	
vign_female	−0.008	(0.028)	−0.008	(0.029)
vign_43y	0.002	(0.033)	−0.002	(0.034)
vign_55y	0.027	(0.033)	0.024	(0.033)
vign_boss	0.010	(0.028)	0.029	(0.030)
vign_freq	0.098***	(0.027)	0.116***	(0.029)
female	−0.160***	(0.051)	−0.180***	(0.068)
age	0.003	(0.002)	0.003	(0.002)
hh_lincome	0.000	(0.021)	0.001	(0.024)
edu_middle	0.156	(0.100)	0.108	(0.138)
edu_high	0.229**	(0.098)	0.219	(0.136)
urban_high	0.008	(0.055)	0.025	(0.075)
urban_middle	−0.105	(0.066)	−0.118	(0.088)
religion	0.023	(0.051)	0.028	(0.067)
vict_index	−0.022	(0.016)	−0.031	(0.021)
trust_index	0.013**	(0.005)	0.015**	(0.007)
social_norm	0.043*	(0.023)	0.092**	(0.036)
victim_small	0.337***	(0.101)	0.403***	(0.138)
victim_serious	0.226**	(0.103)	0.283**	(0.137)
takematerial	−0.116*	(0.063)	−0.160**	(0.076)
justifiability	−0.207***	(0.014)	−0.161***	(0.032)

*** = $\{p < 0.01\}$; ** = $\{0.01 \leq p < 0.05\}$; * = $\{0.05 \leq p < 0.10\}$ Standard errors in parentheses. Dependent variable is *peer_report*.

Table 4.4: Regression results — correlations

	ρ_1	ρ_2	ρ_{11}	ρ_{12}	ρ_{21}	ρ_{22}
Exogeneity	0.97					
Endogeneity	0.97	0.81	−0.15	−0.23	−0.16	−0.22

Dependent variable is *peer_report*.

Characteristics of the offense

There is only one variable in this group, namely justifiability. We know from Figure 4.2 that most respondents disapprove of taking a bundle of printing paper home. Justifiability has a significant negative effect on reporting: respondents who disapprove more are more likely to report (keeping other variables constant). This is not as trivial a result as it may appear, because it shows that the potential respondent's moral judgement is much involved in the decision on whether or not to report. In our case, most respondents find the 'crime' of taking a bundle of printing paper home too futile (see Section 4.3), and would therefore not report it. Including justice evaluation as a possible explanation for peer reporting was considered by Victor et al. (1993), who distinguished between different forms of justice evaluations (distributive, procedural, and retributive justice) and concluded that justice evaluations matter for peer reporting. This is in line with our findings.

The magnitude of the estimated coefficient (−0.161) implies that, for a benchmark respondent with average peer reporting probability, an increase of 1 in the justifiability score leads to a reduction of 0.054 in the probability of peer reporting, keeping x_{it} constant. Since the sample standard deviation of the justifiability scores is 2.05, a one standard deviation increase would lead to a fall in the probability of peer reporting of about 11 percentage points. The effect is therefore not only statistically but also economically significant. According to the estimates in the second column of Table 4.3, the effect of justifiability would be even larger if we assume peer reporting to be exogenous.

Context of the offense

The context is captured by five vignette characteristics, relating peer reporting to the hypothetical situation (for example, behavior of the boss) and to the hypothetical offender (for example, age and gender). Interestingly, we find no evidence that peer reporting is influenced by the age of the offender, nor by the fact whether the offender is a man or a woman. The behavior of the boss does not matter, *ceteris paribus*. The only thing which does matter is whether the offender has engaged in this type of incorrect behavior before or not (*vign_freq*).

Characteristics of the reporter

While we find no evidence that peer reporting is influenced by the age or gender of the offender, the gender of the potential reporter does matter: Men are significantly more likely to report than women (keeping other characteristics constant, including justifiability and personal traits like trust and social norms). This corresponds with other findings (Near and Miceli, 1985; Sims and Keenan, 1998), although the reason for the different reporting behavior of men and women is not clear. We find no significant effect for age. The literature is also ambiguous in this respect (Jones and Kavanagh, 1996; Mesmer-Magnus and Viswesvaran, 2005; Sims and Keenan, 1998). Neither do we find a significant effect of income. If we assume that justifiability is exogenous then we find that higher-educated respondents are more likely to report than respondents with less education (column 2 of Table 4.3), but if we assume endogeneity then this effect is no longer significant. The literature on the effect of education is mixed. Mesmer-Magnus and Viswesvaran (2005) cite studies that find an education effect, but Sims and Keenan (1998) find no significant effect. Whether the respondent lives in a city or in the country does not matter either. We find no evidence that religious people are more likely to report than non-religious people, possibly because religion has an indirect effect on reporting, through ethical ideology (Barnett et al., 1996).

Trust is significantly associated with peer reporting: More trust in others significantly increases the likelihood of peer reporting, probably because a violation of trust affects trusting people more than it affects suspicious people. Important is also *social_norm*, which measures the perceived severity of a wide range of situations of incorrect behavior. We find, as expected, that someone who judges incorrect behavior mildly (low value of *social_norm*) is significantly less likely to report such behavior, keeping justifiability and other variables constant. The size of the parameter estimate implies, for example, that a one standard deviation decrease in *social_norm* reduces the probability of peer reporting by about 6 percentage points for an average respondent. The effect of social norms is much stronger in the model allowing for endogeneity than in the model assuming that justifiability is exogenous. While the existing literature emphasizes the importance of social context (Victor et al., 1993), we are not aware of other studies on peer reporting that incorporate social norms.

New in the literature on peer reporting is also to consider past victimization of the potential reporter. We include a victimization index (*vict_index*) that measures the perceived severity of the different types of crime a respondent has possibly been a victim of, and we also include the fact whether a respondent has been a victim of a small or a serious crime or not. We find that victims of serious crimes and victims of small crimes are more likely to report. The marginal effect of having been a victim of a small crime (an increase of about 13 percentage points in the probability of reporting, for the average respondent) seems to be larger than the effect of *victim_serious* (an increase of about 9 percentage points). Regarding the impact on one's behavior regarding a small crime, this implies that victimization of a small crime has a larger impact than victimization of a serious crime.

Finally, we included a variable *takematerial* which measures whether the respondent him/herself has taken material from work for private use at home. This allows us to see whether a person's own past behavior in a similar situation is of influence on the reporting decision. Note that *takematerial* is negative and significant, which means that respondents that have been in a similar situation as the offender in the vignette are less likely to report.

4.6 Concluding remarks

In this chapter we have considered one 'small crime', namely taking printing paper home from work for private use, and asked whether or not a colleague would report this crime. Peer reporting is viewed as a behavioral response to the perception of fairness (i.e., justifiability) regarding employee theft, because it may be considered an additional task for the employee to help the management or to do justice (see Victor et al., 1993). We learn about the perception of fairness from the vignette question, where the CentERpanel respondents were asked to rate the justifiability on a 10-point scale. We find that situational characteristics, such as the behavior of the offender's boss and the probability of getting caught, influence fairness perception. This perception is also influenced by characteristics of the respondent him/herself, such as the level of trust in others and whether or not the respondent committed employee theft him/herself. Fairness perception and peer reporting are not influenced by age, income or education, but they are influenced by gender: women are less likely to report than men.

The most important aspect triggering peer reporting is the internal attitude towards incorrect behavior. Other important aspects are fairness perception, trust in others, and the potential reporter's own behavior in a comparable situation of employee theft.New in the literature of peer reporting is that we look at the reporter's past victimization. We consider victimization of incorrect behavior in general, and also victimization of a serious crime. We find that the first type of victimization is mainly an attitude variable towards misdemeanors in daily life. The range of misdemeanors a person could possibly have been a victim of in the past five years is so wide that it would seem impossible to find a person that never encountered such a situation. However, only one quarter of the respondents reported being a victim of incorrect behavior, from which we conclude that this group contains people with a greater awareness or sensitivity to social norms. We also find evidence that serious crime victimization changes a person's willingness to report although this effect is smaller than the effect of small crime victimization.

We also looked at reasons for people not to report a misdemeanor. The most important reason for respondents not to report is that the misdemeanor is not important enough to worry about. The loss to a company as a result of stealing a bundle of printing paper is considered to be very small. This is a well-known result: in general, people consider theft from a victim with larger assets (in this case a company) easier to excuse (Greenberg and Scott, 1996).

We mention four possible extensions. First, one could consider group dynamics such as group norms and role responsibility. Such aspects have been found to have an important impact on peer reporting (Victor et al., 1993), but they are difficult to implement in the context of vignette questions, because the description of the hypothetical situation would become too long and too complex. Second, one could look at more serious types of employee theft (in terms of monetary losses to the employer), and ask whether peer reporting happens more often in large than in small organizations or vice versa. Third, it may be the case that organizations with an established ethics program have lower employee theft than organizations without such a program (Greenberg, 2002). Possibly, an ethics program stimulates awareness to social norms in a company and creates a more open environment for allowing employees to report. Fourth, while taking printing paper home for private use would generally be considered as a very minor crime, two-thirds of respondents would report it on average. Our current questionnaire does not enable us to answer the question how this behavior changes with the severity of offenses, since we observe peer reporting behavior only for one situation. Still, this question is of interest and it would also help in differentiating with justifiability.

4.A Variables with explanation

Table A.1: Binary vignette variables with explanation

vign_female	1 if vignette person (vp) is a woman
vign_27y	1 if vp is 27 years old
vign_43y	1 if vp is 43 years old
vign_55y	1 if vp is 55 years old
vign_boss	1 if the boss of the vp behaves correctly
vign_freq	1 if small crime has been committed more often before
vign_catch	1 if the probability of getting caught is 50% (0 if very small)
vign_wage	1 if vp has a high wage
vign_wage_low	1 if vp receives low wage for type of work, given vign_wage = 0
vign_wage_high	1 if vp receives high wage for type of work, given vign_wage = 1

Table A.2: Respondent variables with explanation

Non-binary variables	
age	age of respondent (in years)
hh_lincome	log of gross monthly household income
vict_index	severity of crime respondent has been victim of (0 if no victim)
trust_index	degree of trust in other people (0 if no trust)
social_norm	average of answers to short questions on severity of 18 small crimes on a scale from 1: not severe to 10: very severe
justifiability	1 = crime is never justifiable, 10 = — always justifiable
Binary variables	
female	1 if respondent is a woman
edu_middle	1 if respondent's highest education is secondary school
edu_high	1 if — at least vocational school
urban_high	1 if respondent lives in an urbanized area
urban_middle	1 if — in an area with intermediate urban character
religion	1 if respondent has a religion
victim_small	1 if respondent was victim of incorrect behavior
victim_serious	1 if — of a serious crime
takematerial	1 if respondent took material from the workplace
peer_report	1 if respondent would peer report

Chapter 5

Explaining Subjective Well-Being: The Role of Victimization, Trust, Health, and Social Norms[1]

5.1 Introduction

No scholar would disagree with the statement that crime is costly. *How* costly crime is has not led to an unambiguous answer, as different methodologies and definitions of crime have led to different results. Scholars have relied on three types of methodologies to estimate the cost of crime: (i) revealed preference methods (mainly using the impact of crime on housing prices; see, e.g., Gibbons, 2004) (ii) stated preference methods (leading to 'willingness-to-pay' estimates for avoiding crime; see, e.g., Dolan et al., 2005), and (iii) subjective well-being surveys (see, e.g., Di Tella and MacCulloch, 2008). The costs of crime can be classified as either direct, as a result of law enforcement and deterrence, or indirect, by means of, for example, lower housing prices

[1] This chapter is based on joint work with Arthur van Soest.

or costs of medical care to fearful non-victims. Dolan and Moore (2007), for example, distinguish between tangible and intangible victim costs in this respect, while Cornaglia and Leigh (2011) call this economic and social costs of crime, which essentially means the same. Research on the costs of crime is important as it provides insight in where losses from crime are highest and, therefore, helps to analyze policy measures to reduce the economic and social burden that crime puts on society.

This chapter uses the third method to analyze the importance of crime. Our goal is to rethink and estimate the relation between crime measures and well-being (or, happiness), thereby also considering other variables that affect happiness. We will use a cross-sectional survey data to analyze the association between crime and subjective well-being. Since there is no single measure of crime that captures all concepts related to a criminal action, we will look at different types of measures of crime. We will use data on personal victimization where we distinguish different types of crime, but we will also consider the effects of the frequency of crimes and the fear for crimes in the region.

Victimization is of a complex nature as it influences well-being in many ways: in terms of physical and mental health, but also economically and through the individuals' perception of their surroundings. It is a misconception that this only holds for victims: non-victims suffer from fear of crime in their neighborhood and as a result display lower mental health (see, e.g., Cornaglia and Leigh, 2011) and take precautionary measures against victimization. We find that victimization is not only related to the usual variables that capture personal victimization and fear in the area of residence but that it is also associated with health and social capital. The cross-sectional nature of our data and limited information available for our survey respondents do not allow us to determine whether such associations are causal. This makes our analysis less ambitious than, for example, Cornaglia and Leigh (2011) who use panel data to identify the causal effect of crime on mental health.

A second aim of this chapter is to look at the well-known victimization–

fear paradox: a general finding in crime surveys is the large gap between fear of crime and actual victimization. We find that indeed women and elderly are the least victimized, and estimate how men and women and younger and older respondents differ in terms of the association of victimization with fear in their area of residence on their well-being. The results show that the relations are different for the subgroups under investigation.

The data we use in this chapter come from several sources. We matched survey responses from a survey on incorrect behavior (see Douhou et al., 2011b) with other surveys that have been set out in the same pool of respondents (the CentERpanel) in the same year (2008). Furthermore, we matched these data to administrative data on victimization and fear of crime figures in their region of residence. Other than existing studies, we use broad measures of personal victimization and distinguish between two crime types: serious (assault, robbery, etc.) and small (breaking a mug, littering, etc.). In addition, we also consider the roles of health, trust, and social norms in driving well-being and investigate whether controlling for these factors changes the relation between victimization and subjective well-being.

We try to contribute to three strands of the existing literature. First, studies that look at the association between crime and individuals' subjective well-being (see, e.g., Di Tella and MacCulloch, 2008; Michalos and Zumbo, 2000; Møller, 2005). Second, the literature on the effect of social capital, which is assumed to be a combination of trust, social norms, and associational activity, on well-being (see, e.g., Bjørnskov, 2003; Helliwell, 2003). The analysis of the relation between (self-reported) health and personal victimization is the third literature stream: some references are Koss et al. (1991, 1990) and Britt (2001). Section 5.2 briefly discusses the main mechanisms that lead to an association between crime, well-being, social norms, trust, and health. In Section 5.3 we provide more details of the literature we try to connect our research to. Section 5.4 describes the data and provides some summary statistics. The empirical methodology and the empirical results are presented in Section 5.5. Section 5.6 concludes.

5.2 Crime, fear of crime, trust, health, social norms, and well-being

It is impossible to capture the consequences of crime to which a person is exposed to in one measure. We use personal victimization experience and victimization and fear of crime rates in the respondent's area of residence as direct measures of crime. The latter two variables are also relevant to non-victims as frequent crimes in the neighborhood may lead to a drop in subjective well-being. In addition to creating feelings of fear and anxiety, frequent crime may make people feel less free in their daily routine and may make them take precautionary measures to deter future victimization. On the other hand, it might also be the case that people avoid living in certain areas because they are concerned about crime, leading to a sorting effect of an individual's attitude towards crimes on the crime rate in the area. Denkers and Winkel (1998) find that people with lower well-being are more likely to be victimized and people with lower happiness are living in areas with higher crime. In this chapter we will distinguish between two types of crime a respondent can be a victim of: (i) serious crime (e.g., assault and robbery) and (ii) incorrect behavior (or, small crime) (e.g., damaging a car and fare dodging). We expect the association between well-being and serious crime victimization to be stronger than the association between well-being and small crimes. The latter are more widespread and we expect their effect on well-being to be more of a transitory kind.

Happiness is about how we think and feel about our lives and is therefore related to perception of safety and security, norms and values, and (self-reported) health. We will consider indexes measuring these concepts and their association with subjective well-being, controlling for individual characteristics such as age and income. Someone who has been the victim of a crime may experience lower mental health, and perhaps also lower physical health. Moreover, a person's perception of life may change – changing trust in others or the person's social norms.

5.3 Background

Long before economists started to get interested in 'happiness', researchers in the field of psychology were already working on this topic; see, for example, the review articles of Diener et al. (1999) and Frey and Stutzer (2002). The paradox that is revealed in Easterlin (1974) regarding the relation between income and happiness triggered the interest of economists, starting with Inglehart (1996) and Blanchflower and Oswald (2004). Economists have not only looked at the link between happiness and income, but also at the relation between happiness and, to name a few, unemployment (Clark and Oswald, 1994; Winkelmann and Winkelmann, 1998), macroeconomic volatility (Di Tella et al., 2003), airport noise (Van Praag and Baarsma, 2005), social capital (Bjørnskov, 2003, 2006; Helliwell, 2003, 2006), and inequality (Alesina et al., 2004). Other studies look at alternative measures of well-being such as indicators of mental health problems. A recent example is Cornaglia and Leigh (2011) who look at the crime-mental health interaction. In this chapter we will focus on happiness as the measure of well-being and we will focus on several determinants of happiness: crime victimization, trust, (physical) health, and social norms.

The analysis of the link between well-being and crime also has its roots in psychology and sociology. The focus has mainly been on the psychological effects of having been a victim of a crime on well-being, e.g. through anxiety and fear; see the studies cited in Powdthavee (2005) and Di Tella et al. (2008). Some studies have also analyzed the effects of crime on (subjectively measured physical) health (see Britt, 2001; Koss et al., 1991, 1990). Their main conclusion is that the expected negative association between victimization and health exists: people have significantly worse health after they have been the victim of a crime and more severe crimes are associated with health problems. Other studies have focussed on crime victimization and well-being: Michalos and Zumbo (2000) look at the relation between quality of life and crime-related issues such as fear and actual cases of victimization, neighborhood safety, and beliefs about increases in local crime. They find that victims and non-victims differ in their satisfaction with life but not in a convincing way. With regards to neighborhood satisfaction, the reported difference between victims and non-victims is much higher. Furthermore, these authors find that crime-related issues account for only 7% of the variation in satisfaction with life while explaining 38% of the variation in neighborhood satisfaction. Møller (2005) conducts a similar study using South African data and finds that actual victimization is not as good a predictor of well-being as fear of victimization or personal safety.

Powdthavee (2005) analyzed South African survey data among heads of households regarding the perceived quality of life of the household as a whole. The author relates subjective well-being to information on victimization in the past 12 months of one of the household members in a multiple regression analysis, controlling also for socio-economic characteristics of the household head. Victimized households report significantly lower well-being and if a household lives in a region with a high crime rate this also appears to have a negative effect on well-being. A similar study by Kingdon and Knight (2003), also using South African data, confirmed that household victimization has a significant and negative effect on well-being. In a similar vein, Davies and Hinks (2010) use Malawian survey data and include victimization of the household head and the regional crime rate, but also whether the respondent feels unsafe. As expected, feelings of insecurity and victimization (personal and regional) have a detrimental impact on happiness.

Denkers and Winkel (1998) focus on the influence of victimization on well-being and fear using a sample from the Dutch population.[2] They found no difference between the well-being of victims of violent crime and property crime, but a significant difference between victims and non-victims. Moreover, they found that victims of a crime already appear to be more fearful before they become the victim of a crime and their fear does not seem to change after the crime.

[2] This survey was carried out in the Telepanel, a predecessor of the CentERpanel which was used to collect our data; see section 5.4.

Di Tella and MacCulloch (2008) use happiness responses from a random sample of Europeans (Euro-Barometer Survey Series) and Americans (General Social Survey) for the period 1975-1997. They include aggregate measures as they want to investigate the effect of macroeconomic indicators such as income, unemployment, inflation, and the (violent) crime rate on happiness. The effect of the crime rate is negative in the combined European and American sample but not significant in the regression that includes only European respondents. Cohen (2008, p. 3) notes that due to the nature of this crime rate and since no other crime-related variables are included, this result does not necessarily prove that violent crime has a negative impact; it rather suggests that 'crime and social disarray in general' have a negative impact on well-being. Alesina et al. (2004) include the crime rate as a control variable (since it is correlated with their main variable of interest: inequality) and find no significant effect on happiness.

The paper by Di Tella et al. (2008) has a more specific focus on crime and well-being, investigating correlations between crime-related variables and well-being and different measures of positive and negative emotions (e.g., anger, worry, smiling) for a sample from the Gallup World Poll in 2006 and 2007 covering a large number of countries. Results (excluding Latin American countries) show that victimization is negatively related to well-being.

A study by Cohen (2008) combines previous research by looking at the regional crime rate, perceived neighborhood safety, and personal victimization in the U.S. over the period 1993-2004 (using the General Social Survey). The author concludes that crime rate and neighborhood safety have little impact on well-being. Victimization is only negative and significant for the specific case of victims of burglary, while the more general measure, victim of a violent crime, is not significant. Taking all these studies in consideration we can conclude that the relation between victimization and well-being is not straightforward. The literature agrees that the effect of victimization should be negative but the relation is not always significant and crime-related measures in general are not the most important contributors to explaining the variation in happiness.

Putnam (1993, p. 167) provides an appealing and intuitive definition of social capital: 'features of social organization, such as trust, norms, and networks, that can improve the efficiency of society by facilitating coordinated actions'. Social capital is hypothesized to improve life satisfaction as it makes life easier to have more trust in others and more social interaction. A study by Stutzer and Lalive (2004), on the other hand, showed that social norms are not always a blessing: social work norms put pressure on the unemployed, reducing their life satisfaction. Bjørnskov (2003) looked at cross-country differences in social capital and their association with happiness. He used a social norm index that captures the three elements of social capital and found that it is positively related to happiness. To identify micro and macro measures that influence well-being at the individual and the national level, Helliwell (2003, 2006) included three separate measures of social capital and found that all three are significant and have a positive influence on well-being. Bjørnskov (2006) found, however, that only trust contributes significantly to subjective well-being: adding the other social capital indexes did not lead to significant improvement compared to a model including trust as the only social capital measure.

The association between health and subjective well-being is not obvious since health consists of different dimensions. Dolan et al. (2008) argues that physical health and well-being are positively associated and the causality is most likely to be from health to well-being. As health is considered to be one of the domains of well-being, many studies include a (self-assessed) health measure in happiness regressions, for example, Ravallion and Lokshin (2001) and Cohen (2008). Both find a positive relation between health and subjective well-being. Clark and Oswald (1994) use *mental* well-being as a measure of happiness in relation to unemployment.

The studies discussed above typically use cross-section data and analyze the association between well-being and crime victimization and other variables, without considering potential causality or endogeneity issues. More recent work by Dustmann and Fasani (2011) and Cornaglia and Leigh (2011) is more ambitious and tries to isolate causal from non-causal effects. These studies use mental well-being/health instead of happiness but now in relation to crime measures. They argue that damage of crime can also be inflicted by non-victims, which may add significantly to the costs of crime. Cornaglia and Leigh (2011, p. 20) acknowledge endogeneity of the crime variable(s) —coined as a sorting problem— as people 'with mental distress symptoms are at the same time more likely to react more strongly to crime, or live in areas with higher crime rates'. They account for this by estimating panel data models with fixed effects. They find that sorting is indeed a problem but nevertheless the impact of (area) crime on mental well-being remains significantly negative when sorting is taken into account. Our data do not allow us to use this identification strategy so that we cannot account for potential endogeneity of crime or other variables in our happiness regressions.

5.4 Data and descriptives

5.4.1 Data design

Our data set is based upon several surveys conducted in the Netherlands in June/July 2008 through CentERpanel (CP). CP consists of about 2000 households —representative of the Dutch population— aged 16 years and older, that are repeatedly invited to participate in web-based surveys.[3] The main source of information is a survey entitled "Incorrect Behavior in Everyday Life". See Douhou et al. (2011a) and Douhou et al. (2011b) for a detailed description of the complete survey. In this chapter, one of our main interests is personal victimization experiences of our respondents, which are asked as follows:

- Have you been a victim of a serious crime in the past five years (i.e., burglary, holdup, violence, or something similar)?

[3]Households that have no access to the Internet are provided the necessary means to participate in surveys.

- Have you been a victim of 'incorrect' behavior in the past five years?

If either question is answered 'yes', then a follow-up question asks to rate the severity of the most serious crime on a scale from 1: very severe to 10: not severe. We use this information to construct four dummy variables that distinguish crime types (serious and small) and severity of the crime (severe if the score is 4 or lower and not severe if it is 5 or higher). The reason that we only ask about the past five years is to avoid a bias towards older respondents that have a higher probability of having been a victim in the past.

Most respondents in our small crime survey also participated in several other surveys in the same year. We exploit this to get more detailed background information. Questions on social trust and perceived norms of reciprocity, which we use to construct a trust index, are taken from the CP survey "Victims of (attempt to) Fraud" (Oudejans and Vis, 2008). These questions were phrased as follows:

- *trust1:* Would you say that most people can be trusted or that you cannot be too careful in dealing with people? Please answer on a scale from 1: you have to be careful to 11: most people can be trusted;

- *trust2:* Do you think that most people would try to take advantage of you if they got the chance, or would they try to be honest? Please answer on a scale from 1: most people would try to make advantage of me to 11: most people would try to be honest; and

- *trust3:* Would you say that most of the time people try to be helpful or that they are mostly looking out for themselves? Please answer on a scale from 1: people look mostly of themselves to 11: people try to be helpful.

Health is one of the domains of (satisfaction with) life and is frequently included as a control variable in happiness regressions. From the DNB Household Survey (DHS), an annual survey also administered to respondents in the CP, we use a question on self-assessed health: 'What is the general status of your health?' Our health index simply codes the five answers from 1: poor to 5: excellent, so that higher values indicate better self-assessed health. This survey is conducted between February and September 2008, with most questionnaires completed in April 2008.

The impact of crime-related issues consists not only of actual victimization but also of neighborhood problems, fear of victimization, etcetera (see Michalos and Zumbo, 2000). Since we do not have this information at the individual level, we use data on feelings of fear and the rate of victimization at a regional level. The aggregation is at the level of police regions; the Netherlands is divided into 25 police regions. A police region usually consists of one big city with its surrounding areas.[4] These data come from "Veiligheidsmonitor Rijk" 2008 (VMR), obtained from Statistics Netherlands and conducted mid-2008.

The measure of perceived well-being comes from a CP survey conducted in November/December 2008 entitled "World Perceptions, Technology, and Environment" and is based upon the question: 'Generally speaking, would you say that you are ... 1: very unhappy ... 10: very happy? The respondents were shown a table with a ten point scale but only the extreme values 1 and 10 are provided with verbal labels.[5] All survey data have been collected in the same year. Since all surveys except the World Perceptions Survey are conducted within a period of just a few weeks, we assume that these time differences will not influence our conclusions: it seems highly unlikely that in the few weeks in between these surveys important shocks have taken place that may have affected response behavior. The World Perceptions survey was administered near the end of 2008. This time difference has the advantage that the concern that potential feedback mechanisms from subjective well-being to some of the explanatory variables would be mitigated. It does

[4] These regions are based on population density and crime rate; that is, higher crime rate and/or higher population density lead to a geographically smaller police region. Unfortunately, figures at a more detailed regional level were not available.

[5] The respondents did not have the possibility to answer 'don't know' or 'no answer'.

not, however, take away the concern that common unobserved factors drive well-being as well as, for example, victimization, so that endogeneity is still a potential problem (cf. Section 5.3).

5.4.2 Descriptive statistics

Descriptive statistics are shown in Table 5.1. This tells us that the majority of the respondents are male, the majority finished at least a vocational education, have a partner, and do paid work. Not all respondents from the small crime survey participated in the other CP surveys, resulting in missing values for several measures gathered from other surveys, as can be seen in the third column of the table. The means in the table are very similar to those for the subsample without any missing values, suggesting that non-participation in one of the surveys does not lead to selection problems.

Figure 5.1 shows the empirical distribution of our subjective well-being or 'happiness' variable (Sumner, 1996). The average score is 7.51 (Table 5.1), which says that respondents are fairly happy on average. About 3.3% of the respondents report a happiness level of 3 or lower while the majority of the respondents are at the higher end of the scale: 58.9% reports a happiness level of 7 or higher.[6]

Figure 5.2 shows that people who have been a victim of a serious crime in the past five years typically experienced a serious crime only once, while the empirical distribution of the number of small crimes is more evenly spread. This means that multiple victimization is a more common phenomenon for small crimes than for serious crimes, as expected. The number of unique victims (whether of a serious or of a small crime) is 618 and there are 96 respondents who report that they have been a victim of both a serious crime and a small crime in the past five years. If a respondent reported having been a victim of a crime (small or serious) in the past five years, then the perceived severity of this crime (or the worst of them in the case of multiple

[6]The empirical results presented below did not change when combining the lowest categories.

Table 5.1: Descriptive statistics

	mean	std	*N*
Non-binary			
age	51.33	15.84	1735
health	3.85	0.72	1441
hh_lincome	7.90	1.41	1735
size_hh	2.60	1.27	1735
social_norm	7.02	1.33	1734
trust1	7.34	2.06	1536
trust2	7.41	1.83	1520
trust3	6.94	1.93	1529
trust_index	21.69	5.04	1516
fear_rate	0.20	0.04	1736
vict_rate	0.25	0.04	1734
well-being	7.51	1.35	1736
Binary			
female	0.47		1735
edu_middle	0.39		1730
edu_high	0.54		1730
occup_pension	0.24		1735
occup_indep	0.04		1735
occup_nowork	0.24		1735
partner	0.78		1736
urban_high	0.41		1729
urban_middle	0.20		1729
victsmall_sev	0.11		1725
victsmall_notsev	0.15		1725
victserious_sev	0.06		1725
victserious_notsev	0.06		1725

victimization) was asked using a ten-point scale (1: very severe, 10: not severe). The distribution of the reported answers is presented in Figure 5.3.

It shows that some victims of a serious crime judge the crime to be very severe (1 or 2), while most respondents find the crime rather severe (the modal answer is 3), and only a few do not find the crime severe at all. For small crimes the distribution is more even, as one would expect. The average severity of a small crime is 5.3 (the median is 5), compared to 4.5 (median 4) for a serious crime. For our empirical analysis, we constructed four binary

Figure 5.1: Subjective well-being

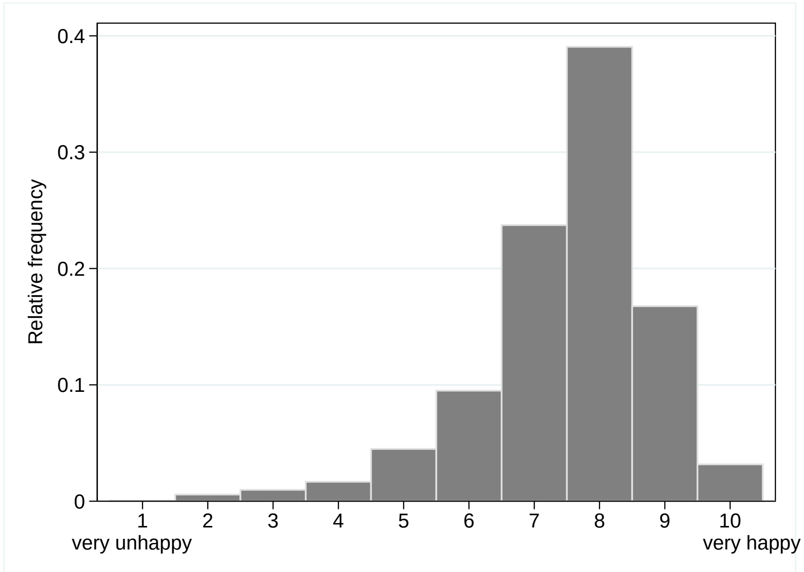


Figure 5.2: Relative frequency of victimization (victims only)

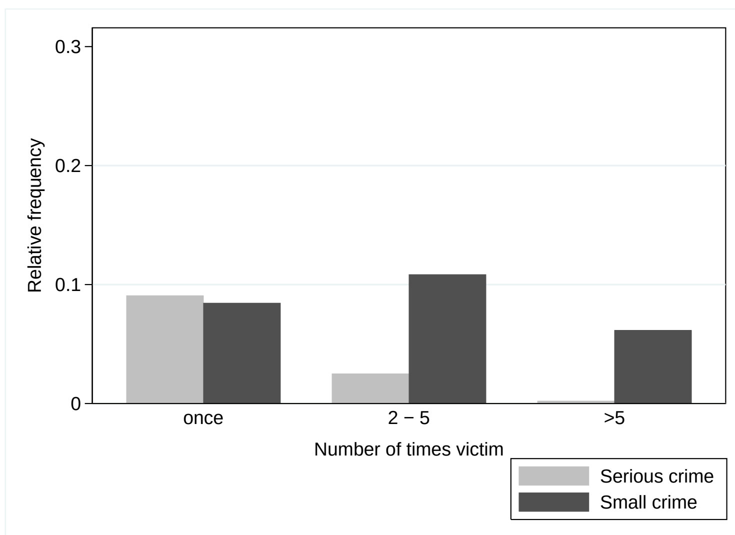
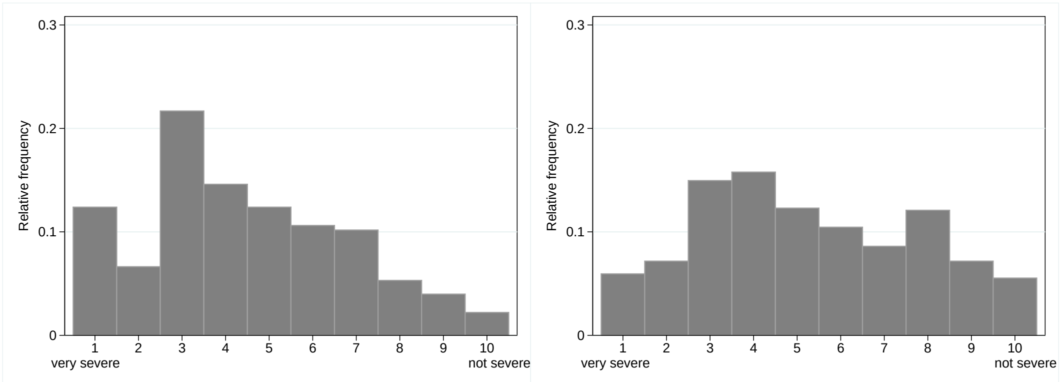


Figure 5.3: Severity of victimization

(a) Serious crime (b) Small crime



variables to distinguish between types of crime and severity of a crime type. A (serious or small) crime is considered severe if the perceived severity (of the worst serious or small crime in the past five years) is rated 4 or lower and not severe if the severity is rated 5 or higher. The reference category are respondents who were not a victim of any serious or small crime. For example, the variable *victsmall_sev* is 1 if a victim of a small crime gives the crime a rating of 4 or lower and 0 otherwise.

In order to provide more insight into the raw data, we present the number of victimized respondents for different groups in Table 5.2. Men and younger people (aged below 55) are more likely to be a crime victim. This is a common result in the empirical literature on crime victimization: the most fearful groups of society (women and elderly) are the least victimized. When we only look at victimization of a small crime the difference between men and women is very small. Furthermore, the elderly are much less likely to report that they have been the victim of a small crime.

In Table 5.3 we present mean scores for well-being of victims and non-victims by gender. Consistent with other studies we find that non-victims report a higher subjective well-being than victims. In addition, well-being for victims of a serious crime is lower than for victims of small crime. The

Table 5.2: Victims of crime by gender and age*

	female	male	<55 years	55+ years
victim_small	24.9	25.6	27.8	21.9
victim_serious	10.9	12.4	12.3	11.0
victim	31.0	32.9	34.5	28.7

* Values are percentage of victims within a subgroup.

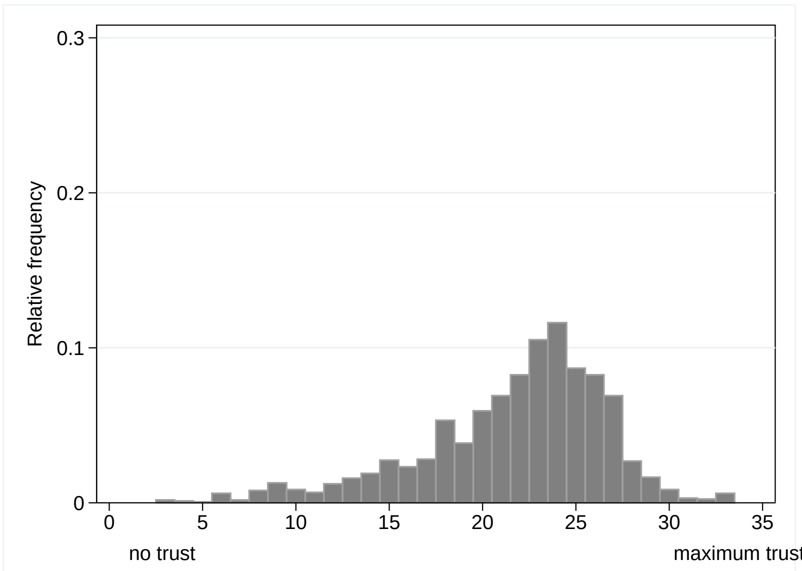
difference in mean scores for women is less obvious: it looks like it does not matter much whether women have been victimized or not. The male group is not the same in this respect: male victims report clearly a lower well-being than male non-victims. Despite the absolute differences in well-being we find that none of the mean differences are significant. This shows us that the victimization–subjective well-being relation is not expected to be strong. This does not discard our main interest as the focal point of our research is not on victimization.

Table 5.3: Victim and subjective well-being mean scores

	mean	std	*N*
victim	7.46	1.34	556
non-victim	7.53	1.36	1169
victim_serious	7.41	1.47	207
non-victim_serious	7.52	1.34	1518
victim_small	7.45	1.29	441
non-victim_small	7.53	1.38	1284
female victim	7.52	1.34	250
female non-victim	7.53	1.30	556
male victim	7.42	1.34	306
male non-victim	7.55	1.41	612

How much trust the respondent has in other people can be important for actions and beliefs in general (Deutsch, 1958) as well as for subjective well-being, since more intense social linkages are expected to make people happier (see Bjørnskov, 2003, 2006; Helliwell, 2003, 2006). The variable *trust_index* is constructed as the sum of three variables that measure several aspects of a person's trust, each on a scale from 1 to 11 (a higher value means more trust in others), so that *trust_index* ranges from 3: very low trust to 33: maximum trust level. Figure 5.4, with a mode of 24 and a mean of 21.7, shows that respondents in general tend to have trust in others.

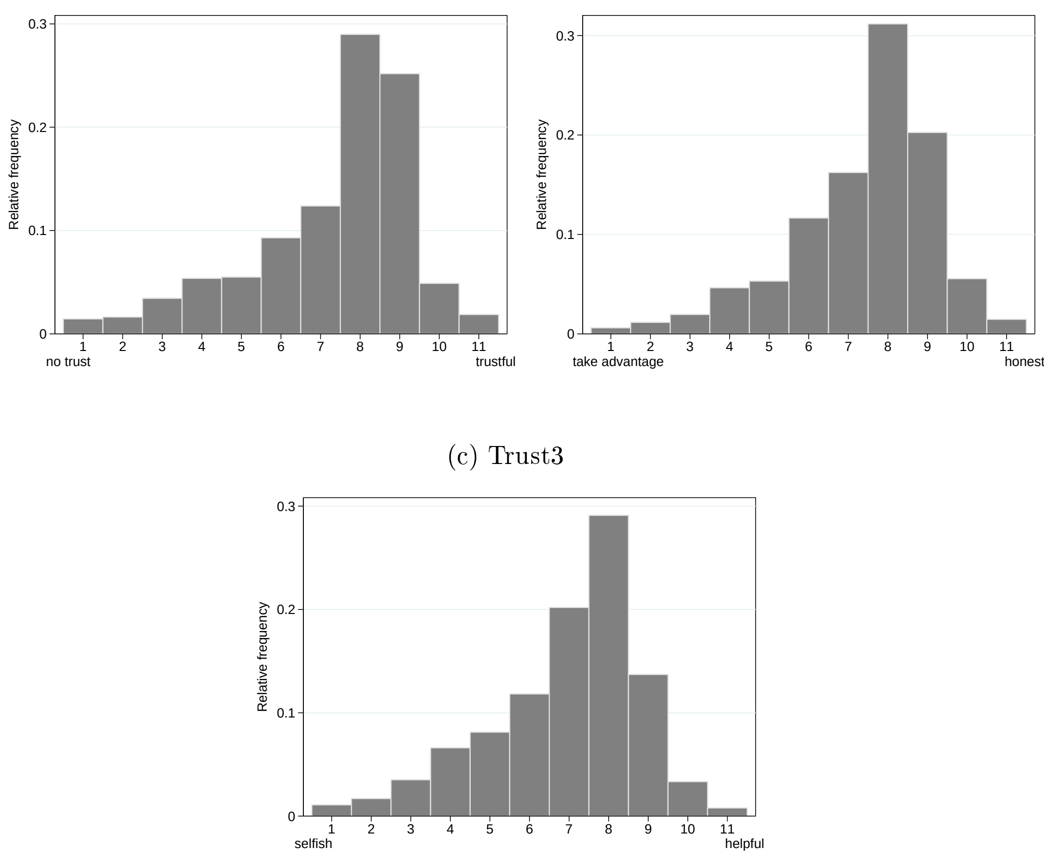
Figure 5.4: Trust_index



See Table A.1b in Section 5.A for details on the three questions. We present the distribution of these three separate trust measures in Figure 5.5, which shows that there are no large differences between the three distributions (see also Table 5.1).

We constructed a social norm index as the average of the responses on severity (on a scale from 1: very severe to 10: not severe) of a list of 18 offenses that differ in the level of damage caused (from stealing a pen to damaging a car and not informing the owner); see Table 2 in Douhou et al. (2011b) for the 18 questions and the mean answers to all of them. To simplify interpretation, our index is defined as 11 minus the mean of the 18 answers, so that a higher value reflects a higher social norm; in the sense of finding crimes less justifiable or more severe. The overall mean of our index is 7.02.

Figure 5.5: Histograms of trust variables



(a) Trust1

(b) Trust2

(c) Trust3

5.5 Regression results

5.5.1 Model

Standard economic theory assumes that individual preferences can be described with a utility function. Following Powdthavee (2005) we assume there exists a utility function for each respondent that describes subjective well-being and has as inputs socio-economic and demographic characteristics, including age, gender, household size, marital status, trust, and past victimization. We will also interact some characteristics to study the effect of victimization for socio-demographic subgroups (defined by age and gender). We obviously cannot observe true well-being, only reported well-being. The literature on psychology shows convincing evidence that reported well-being is correlated with physical reactions that are in turn related to true well-being (see Di Tella and MacCulloch, 2008). According to Frey and Stutzer (2002, p. 405) 'it is a sensible tradition in economics to rely on the judgement of the persons involved'. Hence, we assume that respondents can communicate a level of well-being that is close to their true well-being.

Since the response scale of subjective well-being is discrete and ordered (ranging from 1: very unhappy to 10: very happy), we use an ordered probit model.[7] This model describes the reported evaluation as the category containing the value of an unobserved (latent) continuous variable y_i^*, which is driven by a vector of explanatory variables x_i and an error term ϵ_i:

$$\begin{aligned} y_i^* &= x_i'\beta + \epsilon_i \\ \epsilon_i &\sim N(0,1), \text{ independent of } x_i \\ y_i &= j \quad \text{if} \quad \alpha_{j-1} < y_i^* \leq \alpha_j \end{aligned} \tag{5.1}$$

where $i = 1, \ldots, N$ denotes the respondent, and $j = 1, \ldots, 10$ are the possible values that y_i can have. In the next subsections we will discuss, in turn, the main variables we have in mind for the mechanism discussed in Section 5.2.

5.5.2 Victimization

To show how victimization varies with individual characteristics, Table 5.4 presents regression results with the four personal victimization dummies as dependent variables and some basic respondent and area characteristics as regressors.[8] Not many variables are statistically significant. Living in a

[7] An ordered logit model leads to very similar results.

[8] We also ran a multivariate probit regression and found hardly any differences with the results in Table 5.4.

highly urbanized area significantly increases the probability of being victimized compared to living in a non-urbanized area in three out of four cases. People living in an area with an intermediate urbanization are more likely to be the victim of a severe serious crime but less likely to be the victim of another type of crime than those living in big cities. Respondents with their own (small) business ($occup\_indep = 1$) are significantly more likely to be the victim of a severe (small or serious) crime than employees. This may be because small businesses are vulnerable to burglaries and incorrect behavior by customers. Non-workers less often than employees report to be the victim of a non-severe small crime. We find no relation between living in an area that has a high rate of victimization and/or fear of crime and actual victimization at the individual level. This is not so surprising considering that the local crime-related measures are defined for a relatively broad region, which makes it difficult to find a direct link with personal victimization.

5.5.3 Trust, health, and social norms

Socio-economic variables like gender and income are widely considered as control variables in the well-being literature. We introduce trust, health, and social norms as additional controls, but first analyze whether they are related to crime-related measures. This is important since if they are, crime may affect well-being through these measures or directly. Personal victimization might have an effect on a person's trust in others and their judgement of other crimes. Existing studies show, in addition, that victimization has a negative influence on one's perceived physical health (see Britt, 2001; Koss et al., 1991, 1990). Research on the relation between mental health and victimization (see Cornaglia and Leigh, 2011; Dustmann and Fasani, 2011) comes to the same conclusion: victimization is detrimental to one's (mental) health.[9] This suggests that personal victimization can have an indirect rela-

[9] Admittedly, the meaning of mental health is ambiguous as it can be related to physical health (people that are physically ill are more likely to be depressed and vice versa) and subjective well-being (feeling bad is expected to make less happy and vice versa).

Table 5.4: Probit regression of personal victimization

	victsmall				*victserious*			
	severe		not severe		severe		not severe	
hh_linc	0.136	(0.091)	0.066	(0.082)	−0.145	(0.097)	0.087	(0.112)
age	0.061***	(0.019)	−0.004	(0.016)	0.021	(0.020)	0.006	(0.021)
age^2	−0.001***	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
female	−0.063	(0.085)	0.045	(0.078)	−0.085	(0.100)	−0.141	(0.107)
edu_middle	−0.067	(0.181)	−0.230	(0.162)	−0.037	(0.204)	−0.435**	(0.196)
edu_high	0.032	(0.179)	−0.072	(0.159)	−0.026	(0.204)	−0.254	(0.192)
urban_high	0.194**	(0.098)	0.202**	(0.091)	0.251**	(0.119)	0.023	(0.125)
urban_middle	0.126	(0.110)	0.075	(0.104)	0.301**	(0.130)	0.184	(0.135)
occup_pension	−0.023	(0.166)	−0.195	(0.159)	−0.275	(0.201)	−0.264	(0.216)
occup_indep	0.283*	(0.169)	−0.027	(0.173)	0.503***	(0.183)	−0.018	(0.231)
occup_nowork	0.040	(0.115)	−0.249**	(0.110)	0.057	(0.133)	0.083	(0.143)
size_hh	−0.008	(0.040)	−0.046	(0.037)	0.024	(0.048)	−0.050	(0.050)
partner	−0.136	(0.114)	0.088	(0.107)	−0.224*	(0.130)	0.000	(0.140)
vict_rate	0.165	(1.798)	1.574	(1.695)	0.677	(2.108)	3.439	(2.144)
fear_rate	−0.536	(1.802)	−2.110	(1.681)	0.029	(2.136)	−0.721	(2.180)
constant	−3.619***	(0.881)	−1.203	(0.787)	−1.115	(0.960)	−2.665**	(1.074)
N	1820		1820		1820		1820	
pseudo R^2	0.025		0.026		0.038		0.032	

*** = $\{p < 0.01\}$; ** = $\{0.01 \leq p < 0.05\}$; * = $\{0.05 \leq p < 0.10\}$ Standard errors in parentheses. We included a dummy for zero income.

tion with happiness via self-assessed health, trust, and social norms.

In Table 5.5 we present regression results with trust, health, and social norms as dependent variables. We find that females, older persons, higher educated respondents, and people with a high income have more trust in others compared to their counterparts (males, younger persons, etcetera). Females and older age groups are also found to have higher social norms, in the sense that they find small crimes less justifiable than males and younger age groups. The gender difference is consistent with a fair part of the literature on ethical decision-making, but the results for age in the existing literature are ambiguous (O'Fallon and Butterfield, 2005). As expected, richer, higher educated, and younger people give themselves a better health rating, while people without full-time work ($occup\_nowork = 1$) have lower self-assessed health.

The last column of Table 5.5 shows that social trust and health are positively associated at the individual level, controlling for socio-economic variables. This is in line with the existing literature (Barefoot et al., 1998; Poortinga, 2006; Rose, 2000). As emphasized before, we cannot claim that this reflects a causal effect in a given direction: Poortinga (2006, p. 301) notes that poor health may lead to social exclusion and lower trust, but Rose (2000) finds an effect of social trust on health.

Being the victim of a severe small or serious crime is negatively related to trust in others and health, while a positive association with social norms, i.e., victimization seems to make the respondent *more* dismissive of crimes. On the other hand, being the victim of a not so severe small crime makes one's judgement of small crimes milder or, in other words, it lowers social norms. Being a victim of a serious crime has no significant effect on health, while a not severe crime victimization is negatively related to social norms.

Table 5.5: Regressions with trust, social norms, and health as dependent variables

	trust_index		*social_norm*		*health*		*health*	
hh_linc	1.319***	(0.286)	0.015	(0.063)	0.278***	(0.068)	0.250***	(0.068)
age	0.127**	(0.051)	0.034***	(0.012)	−0.023*	(0.012)	−0.030**	(0.012)
age^2	−0.001*	(0.001)	0.000	(0.000)	−0.000	(0.000)	0.000	(0.000)
female	1.249***	(0.264)	0.324***	(0.061)	−0.076	(0.064)	−0.132**	(0.065)
edu_middle	0.722	(0.543)	0.111	(0.124)	0.324***	(0.123)	0.291**	(0.123)
edu_high	1.677***	(0.541)	0.141	(0.124)	0.356***	(0.123)	0.291**	(0.124)
urban_high	0.211	(0.304)	−0.063	(0.071)	−0.057	(0.073)	−0.050	(0.074)
urban_middle	−0.664**	(0.335)	−0.071	(0.079)	−0.041	(0.081)	−0.013	(0.081)
occup_pension	0.912*	(0.500)	0.032	(0.116)	−0.076	(0.118)	−0.108	(0.119)
occup_indep	0.669	(0.631)	−0.428***	(0.139)	−0.101	(0.149)	−0.077	(0.151)
occup_nowork	−0.387	(0.352)	0.028	(0.083)	−0.368***	(0.085)	−0.362***	(0.086)
size_hh	0.168	(0.123)	0.075**	(0.029)	0.060**	(0.030)	0.052*	(0.030)
victsmall_sev	−0.879**	(0.411)	0.202**	(0.095)	−0.223**	(0.099)	−0.231**	(0.100)
victsmall_notsev	0.238	(0.356)	−0.251***	(0.083)	−0.081	(0.086)	−0.077	(0.087)
victserious_sev	−0.963*	(0.529)	0.047	(0.122)	−0.156	(0.125)	−0.133	(0.126)
victserious_notsev	0.195	(0.539)	−0.216*	(0.127)	0.136	(0.132)	0.163	(0.133)
partner	−0.709**	(0.351)	−0.196**	(0.082)	0.025	(0.084)	0.057	(0.084)
vict_rate	0.070	(5.651)	0.564	(1.315)	−0.038	(1.374)	0.021	(1.380)
fear_rate	−1.957	(5.540)	−1.369	(1.302)	−0.194	(1.364)	−0.219	(1.369)
trust_index							0.032***	(0.006)
social_norm							0.067***	(0.024)
constant	6.178**	(2.702)	5.041***	(0.603)				
N	1576		1820		1510		1510	
(pseudo) R^2	0.074		0.176		0.051		0.062	

*** = $\{p < 0.01\}$; ** = $\{0.01 \leq p < 0.05\}$; * = $\{0.05 \leq p < 0.10\}$ Standard errors in parentheses. We included a dummy for zero income. Ordered probit is used for health and OLS for trust and social norms.

5.5.4 Happiness

We distinguish two models for subjective well-being: a baseline model and an extended specification.[10] The results are presented in Table 5.6. The baseline model shows a marginally significant negative association between subjective well-being and being the victim of a not severe small crime. The other three victimization dummies are insignificant. The control variables that have a significant relationship with subjective well-being are similar to what is found in the happiness literature: women are in general happier than men with the same socio-economic characteristics and having a partner increases one's happiness. Additional household members are also significantly associated with more happiness but this effect is much smaller than that of having a partner. Higher household income is also associated with more happiness.[11] Retired people have more time for leisure which can explain why they are happier: the effect of *occup_pension* is positive and significant (the reference group consists of people on a payroll).

The second specification extends the basic model with indexes for trust, social norms, self-assessed health, and regional crime-related measures. In the extended specification the explained variance (pseudo R^2) increases from roughly 0.03 to 0.07, which is close to the results found in related studies. The results for the socio-demographic characteristics in the extended model are generally comparable to those in the basic model, though gender and

[10]Dolan et al. (2008) criticize studies on subjective well-being for including a single specification only and not showing what the impact is when other or more controls are added. With this set-up we try to meet this criticism.

[11]Easterlin (1974) showed that happiness and income are positively correlated but that over time, as average income levels increased, happiness did not increase accordingly. This result, referred to as the Easterlin paradox, stirred a lot of research on how to measure income to capture an income effect in a well-being regression. Since our data are of cross sectional nature we will keep matters simple and include the (log of) absolute income level to account for the fact that people with higher income have more means to satisfy their needs and are therefore expected to be happier. In addition, gross monthly income is censored at 10,000 euros to account for outliers; since zero incomes may be misreported (and thus reflect missing values) we also include a dummy variable for zero reported income (not reported in the table).

urbanization that were significant in the basic model are no longer significant. Health and happiness are found to be strongly positively related, which is in line with expectations and the existing literature.

Table 5.6: Ordered probit regression: basic and extended model for total sample

	Basic		Extended	
hh_linc	0.249***	(0.055)	0.153***	(0.056)
age	−0.015	(0.011)	−0.013	(0.011)
age^2	0.000	(0.000)	0.000	(0.000)
female	0.108**	(0.054)	0.057	(0.055)
edu_middle	0.071	(0.109)	−0.059	(0.110)
edu_high	0.076	(0.109)	−0.099	(0.111)
urban_high	−0.006	(0.058)	0.018	(0.063)
urban_middle	−0.118*	(0.069)	−0.091	(0.070)
occup_pension	0.212**	(0.102)	0.231**	(0.103)
occup_indep	0.101	(0.125)	0.099	(0.126)
occup_nowork	−0.109	(0.072)	0.003	(0.073)
size_hh	0.061**	(0.026)	0.046*	(0.026)
victsmall_sev	−0.065	(0.084)	0.025	(0.085)
victsmall_notsev	−0.136*	(0.073)	−0.127*	(0.074)
victserious_sev	0.060	(0.107)	0.130	(0.108)
victserious_notsev	−0.018	(0.110)	−0.046	(0.111)
partner	0.439***	(0.073)	0.491***	(0.074)
vict_rate			−3.276***	(1.170)
fear_rate			3.145***	(1.164)
health			0.483***	(0.041)
trust_index			0.046***	(0.006)
social_norm			0.044**	(0.021)
N	1713		1711	
pseudo R^2	0.026		0.068	

$^{***} = \{p < 0.01\}$; $^{**} = \{0.01 \leq p < 0.05\}$; $^{*} = \{0.05 \leq p < 0.10\}$
Standard errors in parentheses. We included dummies for missing observations for *health* and *trust_index* and a dummy for zero income.

We also find a strong and significant positive link between our broad index of trust and subjective well-being.[12] Another component of social cap-

[12] We also looked at a specification where we included the three trust-type of variables

ital, the social norms index, also positively and significantly contributes to well-being. This is consistent with the empirical literature on social capital.

We use the regional rate of victimization[13] and the rate of fear of crime in the respondent's region of residence to capture area-specific relations between crime and well-being. The rate of victimization is significant and has the expected sign: respondents living in an area with a high victimization rate are less happy than others, *ceteris paribus*. On the other hand, we find a strong and *positive* relation between $fear\_rate$ and well-being. This result seems counterintuitive. Cohen (2008) offers an explanation for this result: people who live in unsafe areas are compensated for the higher risk of victimization via lower costs of living or adapt their behavior, which might result in a higher well-being compared to people who live in areas considered safer.

Regarding actual victimization we find similar results as for the baseline model: victims of a not so severe small crime have a lower well-being, although this association is only significant at a 10% level. The other victimization dummies remain insignificant. The coefficients of victimization in the extended model reflect the direct relation between personal victimization and well-being only (keeping trust, social norms, and health constant), while victimization in the basic model measures the sum of the direct and indirect relation between victimization and well-being. As trust, health, and social norms are associated with victimization (see Table 5.5) and well-being (see Table 5.6), we expected an indirect relation to exist. Apparently, this is not strong enough to lead to a substantial difference between the coefficients on victimization in the two models.[14]

---

separately (instead of combining them into one index) and found that the effect of *trust2* (honesty by others) is slightly larger than that of *trust1* and *trust3*. However, a likelihood ratio test did not reject the assumption that the three trust variables have the same coefficient, which is what we assumed in the model presented here.

[13] This measure includes victimization from violent and property crimes and from vandalism; we did a similar analysis including separate victimization rates for each crime type and find no significant results.

[14] Since victimization is not correlated with $fear\_rate$ and $vict\_rate$ (see Table 5.4) we can safely say that the changes to the coefficients of victimization in Table 5.6 when moving from the basic to the extended model reflect possible indirect relations.

The relations we find between personal victimization and well-being are not as strong as we expected. This is in line with Hanson et al. (2010, p. 193) who conclude in a literature review on the (functional) impact of victimization on subjective well-being that the findings are 'not robust'. There can be several explanations for this. First, endogeneity as a result of unobserved individual characteristics influence the results. Second, the way personal victimization is measured: the victimization window in the survey is five years, which may be considered too long to capture a (robust) association with subjective well-being. Moreover, other measurement errors, such as the definition of the crime types and telescoping, may be at work here. Despite this we find very convincing results for the association of happiness with health, trust, and social norms.

5.5.5 Results by age and gender

The majority of the victims in our sample are males younger than 55. Female and elderly groups are known for displaying the highest fear of victimization although crime statistics show that they have the lowest probability of being victimized. This suggests that the role of victimization may differ for men and women and for younger and older respondents. We therefore also estimated the models separately by gender and age group (younger than 55 versus 55 years or older). The results are presented in Tables 5.7 and 5.8. Again we distinguish a baseline and an extended specification. The results by age group in Table 5.7 show that victimization is negatively related to subjective well-being for older respondent in both specifications. In addition, retired people are more satisfied than people (of the same age) who did not retire yet. Looking at Table 5.8 it is interesting to see that the positive association between income and the happiness only applies to men. The effect of severe small crimes is positive and marginally significant for women in the extended model while it is negative for men. Could this indicate adaptive behavior after a negative experience by women and not by men? The association between not working (which includes students, unemployed, incapacitated

Table 5.7: Ordered probit regression: basic and extended model by *age*

	Basic				Extended			
	<55 years		55^+ years		<55 years		55^+ years	
hh_linc	0.270***	(0.082)	0.237***	(0.077)	0.163*	(0.084)	0.147*	(0.079)
age	−0.031	(0.029)	−0.005	(0.090)	−0.020	(0.029)	−0.020	(0.091)
age^2	0.000	(0.000)	0.000	(0.001)	0.000	(0.000)	0.000	(0.001)
female	0.197***	(0.072)	−0.026	(0.086)	0.178**	(0.074)	−0.155*	(0.089)
edu_middle	0.198	(0.163)	−0.068	(0.152)	0.049	(0.165)	−0.219	(0.154)
edu_high	0.183	(0.167)	−0.022	(0.154)	0.015	(0.169)	−0.281*	(0.157)
urban_high	0.017	(0.082)	0.007	(0.085)	−0.006	(0.087)	0.064	(0.093)
urban_middle	−0.177*	(0.095)	−0.030	(0.104)	−0.182*	(0.095)	0.026	(0.105)
occup_pension			0.166	(0.136)			0.228*	(0.138)
occup_indep	0.084	(0.154)	0.088	(0.220)	0.046	(0.157)	0.200	(0.222)
occup_nowork	−0.232**	(0.098)	−0.052	(0.126)	−0.076	(0.101)	0.140	(0.128)
size_hh	0.113 ***	(0.030)	−0.078	(0.070)	0.076**	(0.031)	0.011	(0.071)
victsmall_sev	−0.132	(0.115)	−0.024	(0.126)	−0.060	(0.116)	0.097	(0.127)
victsmall_notsev	−0.086	(0.092)	−0.206*	(0.122)	−0.090	(0.093)	−0.217*	(0.124)
victserious_sev	0.054	(0.141)	0.106	(0.167)	0.102	(0.142)	0.184	(0.168)
victserious_notsev	0.073	(0.145)	−0.122	(0.170)	0.048	(0.146)	−0.155	(0.174)
partner	0.455***	(0.103)	0.505***	(0.120)	0.575***	(0.106)	0.389***	(0.122)
vict_rate					−3.099*	(1.630)	−3.369*	(1.717)
fear_rate					4.428***	(1.576)	1.820	(1.765)
health					0.507***	(0.058)	0.471***	(0.061)
trust_index					0.039***	(0.008)	0.056***	(0.008)
social_norm					0.038	(0.028)	0.043	(0.032)
N	936		777		936		775	
pseudo R^2	0.038		0.022		0.076		0.072	

*** = $\{p < 0.01\}$; ** = $\{0.01 \le p < 0.05\}$; * = $\{0.05 \le p < 0.10\}$ Standard errors in parentheses. We included dummies for missing observations for *health* and *trust_index* and a dummy for zero income.

Table 5.8: Ordered probit regression: basic and extended model by *gender*

	Basic				Extended			
	women		men		women		men	
hh_linc	0.109	(0.078)	0.370***	(0.079)	0.029	(0.079)	0.262***	(0.080)
age	−0.025	(0.016)	−0.006	(0.015)	−0.021	(0.016)	−0.003	(0.015)
age^2	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
female								
edu_middle	0.013	(0.157)	0.108	(0.155)	−0.255	(0.160)	0.098	(0.156)
edu_high	−0.018	(0.161)	0.126	(0.153)	−0.279*	(0.164)	0.012	(0.155)
urban_high	−0.018	(0.085)	0.003	(0.081)	0.053	(0.091)	−0.018	(0.087)
urban_middle	−0.096	(0.101)	−0.147	(0.097)	−0.090	(0.101)	−0.099	(0.098)
occup_pension	−0.017	(0.158)	0.322**	(0.139)	0.082	(0.160)	0.274*	(0.141)
occup_indep	0.149	(0.197)	0.065	(0.162)	0.130	(0.200)	0.056	(0.164)
occup_nowork	−0.193**	(0.095)	−0.096	(0.129)	−0.070	(0.097)	0.101	(0.132)
size_hh	0.097***	(0.037)	0.037	(0.037)	0.076**	(0.038)	0.025	(0.037)
victsmall_sev	0.123	(0.128)	−0.209*	(0.113)	0.246*	(0.129)	−0.130	(0.115)
victsmall_notsev	−0.142	(0.107)	−0.111	(0.101)	−0.171	(0.108)	−0.071	(0.102)
victserious_sev	0.135	(0.159)	0.038	(0.146)	0.303*	(0.160)	0.020	(0.148)
victserious_notsev	−0.041	(0.167)	−0.024	(0.147)	−0.069	(0.171)	−0.036	(0.148)
partner	0.507***	(0.102)	0.368***	(0.109)	0.553***	(0.103)	0.434***	(0.110)
vict_rate					−4.295**	(1.784)	−2.590*	(1.566)
fear_rate					3.626**	(1.780)	2.974*	(1.556)
health					0.507***	(0.063)	0.490***	(0.057)
trust_index					0.045***	(0.009)	0.049***	(0.008)
social_norm					0.020	(0.030)	0.070**	(0.029)
N	801		912		800		911	
pseudo R^2	0.030		0.028		0.072		0.073	

$^{***} = \{p < 0.01\}$; $^{**} = \{0.01 \leq p < 0.05\}$; $^{*} = \{0.05 \leq p < 0.10\}$ Standard errors in parentheses. We included dummies for missing observations for *health* and *trust_index* and a dummy for zero income.

for work, or otherwise) and happiness is negative for women in the basic specification. This effect is comparable to the well-known negative effect of unemployment (see Clark and Oswald, 1994; Winkelmann and Winkelmann, 1998), although our measure of non-employment is broader than (involuntary) unemployment.[15] Higher social norms are significantly positively related to the well-being of men but insignificant for women.

It can be concluded that the relation between personal victimization and well-being is weak for all groups. In the extended specification, we find some negative effects that are marginally significant, but almost as many marginally significant counterintuitive positive effects. Living in a region with a high rate of victimization is significantly negatively associated with subjective well-being for all subgroups, and the association is particularly strong for women. On the other hand, we find a positive effect of the regional fear of crime rate which is particularly strong for younger individuals. The effects of trust and health are significantly positive for all groups.

5.5.6 Some sensitivity checks

Up to now we modeled our respondents as independent from each other while they are actually part of a household where interdependencies regarding well-being may exist (Winkelmann, 2005). A first attempt to correct for this is presented in Table 5.9 by means of clustered standard errors within a household. We see that this slightly elevates standard errors but no real differences appear when we compare the results with Table 5.6.

Another way to correct for household interdependencies is to explicitly model it by using an ordered probit model with household specific random effects. The results in Table 5.10 show that the personal victimization variables are not significant anymore while all other results are similar to what we have found before.

[15] Occupational status and students or others in the 'non-employment' group can have a small job. Still, Bardasi and Francesconi (2004) show that seasonal or casual work has a negative effect on well-being.

Table 5.9: Ordered probit regression: basic and extended model using clustered errors

	Basic		Extended	
hh_linc	0.249***	(0.055)	0.153***	(0.056)
age	−0.015	(0.011)	−0.013	(0.011)
age^2	0.000	(0.000)	0.000	(0.000)
female	0.108**	(0.049)	0.057	(0.051)
edu_middle	0.071	(0.123)	−0.059	(0.122)
edu_high	0.076	(0.122)	−0.099	(0.120)
urban_high	−0.006	(0.066)	0.018	(0.070)
urban_middle	−0.118	(0.080)	−0.091	(0.078)
occup_pension	0.212*	(0.109)	0.231**	(0.109)
occup_indep	0.101	(0.113)	0.099	(0.117)
occup_nowork	−0.109	(0.074)	0.003	(0.073)
size_hh	0.061**	(0.029)	0.046	(0.029)
victsmall_sev	−0.065	(0.087)	0.025	(0.091)
victsmall_notsev	−0.136*	(0.071)	−0.127*	(0.070)
victserious_sev	0.060	(0.121)	0.130	(0.121)
victserious_notsev	−0.018	(0.126)	−0.046	(0.123)
partner	0.439***	(0.074)	0.491***	(0.077)
vict_rate			−3.276**	(1.323)
fear_rate			3.145**	(1.337)
health			0.483***	(0.049)
trust_index			0.046***	(0.007)
social_norm			0.044*	(0.024)
N	1713		1711	
pseudo R^2	0.026		0.068	

$^{***} = \{p < 0.01\}$; $^{**} = \{0.01 \leq p < 0.05\}$; $^{*} = \{0.05 \leq p < 0.10\}$
Standard errors in parentheses. We included dummies for missing observations for *health* and *trust_index* and a dummy for zero income.

Table 5.10: Random effects ordered probit regression: basic and extended model

	Basic		Extended	
hh_linc	0.335***	(0.081)	0.186**	(0.078)
age	−0.027*	(0.014)	−0.021	(0.013)
age^2	0.000	(0.000)	0.000	(0.000)
female	0.161**	(0.065)	0.084	(0.065)
edu_middle	0.022	(0.138)	−0.125	(0.136)
edu_high	0.020	(0.139)	−0.174	(0.137)
urban_high	−0.020	(0.086)	0.015	(0.088)
urban_middle	−0.136	(0.103)	−0.098	(0.099)
occup_pension	0.285**	(0.133)	0.284**	(0.130)
occup_indep	0.159	(0.162)	0.158	(0.159)
occup_nowork	−0.122	(0.092)	−0.004	(0.091)
size_hh	0.084**	(0.038)	0.067*	(0.037)
victsmall_sev	−0.034	(0.111)	0.049	(0.108)
victsmall_notsev	−0.136	(0.095)	−0.136	(0.093)
victserious_sev	0.026	(0.141)	0.097	(0.138)
victserious_notsev	0.043	(0.146)	−0.010	(0.142)
partner	0.580***	(0.101)	0.626***	(0.099)
vict_rate			−4.003**	(1.636)
fear_rate			3.547**	(1.642)
health			0.562***	(0.053)
trust_index			0.062***	(0.008)
social_norm			0.057**	(0.027)
N	1713		1711	
ρ	0.450		0.400	

$^{***} = \{p < 0.01\}$; $^{**} = \{0.01 \leq p < 0.05\}$; $^{*} = \{0.05 \leq p < 0.10\}$
Standard errors in parentheses. We included dummies for missing observations for *health* and *trust_index* and a dummy for zero income.

In Table 5.11 we include measures of victimization that discriminate between single and multiple victimization (using the information on the number of small or serious crimes that respondents were a victim of in the last five years; see Figure 5.2). Higher values indicate that a person has been more often a victim of a certain type of crime. Multiple victimization of a not so severe small crime has a negative but not very strong and marginally significant relation with well-being; the other variables are insignificant.

Table 5.11: Ordered probit regression: basic and extended model including multiple victimization

	Basic		Extended	
hh_linc	0.250***	(0.055)	0.152***	(0.056)
age	−0.015	(0.011)	−0.013	(0.011)
age^2	0.000	(0.000)	0.000	(0.000)
female	0.108**	(0.054)	0.056	(0.055)
edu_middle	0.076	(0.109)	−0.053	(0.110)
edu_high	0.081	(0.109)	−0.094	(0.111)
urban_high	−0.004	(0.058)	0.019	(0.063)
urban_middle	−0.116*	(0.069)	−0.089	(0.070)
occup_pension	0.211**	(0.102)	0.232**	(0.103)
occup_indep	0.104	(0.125)	0.099	(0.126)
occup_nowork	−0.108	(0.072)	0.005	(0.073)
size_hh	0.062**	(0.026)	0.046*	(0.026)
mvictsmall_sev	−0.031	(0.041)	0.015	(0.042)
mvictsmall_notsev	−0.065*	(0.034)	−0.057*	(0.034)
mvictserious_sev	0.013	(0.082)	0.071	(0.082)
mvictserious_notsev	0.010	(0.078)	−0.004	(0.079)
partner	0.436***	(0.073)	0.488***	(0.074)
vict_rate			−3.280***	(1.169)
fear_rate			3.155***	(1.164)
health			0.481***	(0.041)
trust_index			0.046***	(0.006)
social_norm			0.044**	(0.021)
N	1713		1711	
pseudo R^2	0.026		0.068	

$^{***} = \{p < 0.01\}$; $^{**} = \{0.01 \leq p < 0.05\}$; $^{*} = \{0.05 \leq p < 0.10\}$
Standard errors in parentheses. We included dummies for missing observations for *health* and *trust_index* and a dummy for zero income.

Finally, we consider some dynamic effects. Due to the cross-section nature of our data, we cannot consider changes in all(left hand side or right hand side) variables and follow a fixed effects approach like Cornaglia and Leigh (2011). But we are able to use values of the regional variables two years earlier. Moreover, we also know whether people still live at the same address as two years earlier. In our sample, about 8% of the respondents have moved between 2006 and 2008. First, in order to see whether for movers the association with the regional crime rate is different than for non-movers, we included a dummy for movers as well as an interaction term between a dummy for moving in the last two years and the rate of victimization in the area of residence. Both variables are insignificant, and including them hardly changes the other coefficients — see the left hand columns in Table 5.12.

Table 5.12: Ordered probit regression: robustness checks with dynamics

hh_linc	0.154***	(0.056)	0.169***	(0.057)
age	−0.012	(0.011)	−0.017	(0.011)
age^2	0.000	(0.000)	0.000	(0.000)
female	0.055	(0.055)	0.057	(0.058)
edu_middle	−0.060	(0.110)	−0.059	(0.112)
edu_high	−0.103	(0.111)	−0.084	(0.113)
urban_high	0.019	(0.063)	0.050	(0.061)
urban_middle	−0.092	(0.070)	−0.090	(0.072)
occup_pension	0.091	(0.127)	0.126	(0.139)
occup_nowork	0.007	(0.073)	0.012	(0.076)
size_hh	0.048*	(0.026)	0.053*	(0.028)
victsmall_sev	0.025	(0.085)	0.054	(0.090)
victsmall_notsev	−0.126*	(0.074)	−0.152**	(0.077)
victserious_sev	0.129	(0.108)	0.053	(0.114)
victserious_notsev	−0.044	(0.111)	−0.055	(0.116)
partner	0.486***	(0.074)	0.474***	(0.076)
vict_rate	−3.189***	(1.194)		
d_mover	0.377	(0.569)		
d_mover*vict_rate	−1.049	(2.188)		
delta_victrate			−0.438	(0.351)
fear_rate	3.118***	(1.166)		
delta_fearrate			0.091	(0.311)
health	0.484***	(0.041)	0.503***	(0.044)
trust_index	0.046***	(0.006)	0.042***	(0.006)
social_norm	0.045**	(0.021)	0.045**	(0.021)
N	1711		1589	
pseudo R^2	0.068		0.068	

$^{***} = \{p < 0.01\}$; $^{**} = \{0.01 \leq p < 0.05\}$; $^{*} = \{0.05 \leq p < 0.10\}$
Standard errors in parentheses. We included dummies for missing observations for *health* and *trust_index* and a dummy for zero income. *d_mover* = 1 if moved between 2006 and 2008. Delta means % change in respective rate between 2006 and 2008. Dependent variable is *well-being*.

Second, we investigate whether well-being is associated with *changes* in an individual's regional victimization and fear indexes rather than the levels. The right hand columns of Table 5.12 present the results for the non-movers only. We find no significant effect of the changes in the regional variables. Of course it is possible that this is due to the fact that we only distinguish 25 regions, which gives too large regions to capture the probability of victimization and fear of crime in the neighborhood.

5.6 Conclusion

This chapter studies subjective well-being by means of a survey of about 2000 Dutch respondents in 2008, focusing on its association with crime-related measures as well as health, trust, and social norms. The analysis allows us to distinguish a direct association between victimization of crime or the regional crime or fear of crime rate from indirect relations through trust, health, and social norms, which are related to crime-related measures as well as subjective well-being. This approach is different from the usual empirical strategy in the literature on well-being.

Victims in our sample are, as expected, more likely to be male and younger than 55 years. We find that victims have a lower mean score for subjective well-being than non-victims but this difference is not significant. This is confirmed in the regression results: when we control for basic characteristics (age, income, gender, urbanization etcetera), we only find a weak effect of not severe small crimes and no significant effect of more serious crimes. This does not change if we extend the specification with trust, social norms, perceived health, the regional victimization rate, and the regional fear of crime rate. On the other hand, we do find a significantly negative association between well-being and the regional rate of crime but also a somewhat unexpected positive association with an index for fear of crime at the same regional level. Moreover, we find that people who are healthy, have more trust in others, or have higher social norms are significantly happier.

That the relation between victimization and well-being is not a clear or strong one is not new: Møller (2005), Michalos and Zumbo (2000) and Cohen (2008) concluded that crime-related issues (including victimization of violent and property crimes) have very little impact on well-being. They find a significant negative impact but the results in studies that use regression analysis are not robust. There are some limitations regarding how we measured personal victimization that may explain the weak result for personal victimization. First, the personal victimization question may be prone to measurement errors. We use a five year window, which may be too long to capture a strong effect. A shock, typically, mainly affects a person's life immediately after the fact and most psychological problems disappear after a few months (Denkers and Winkel, 1998). Another source of measurement error may come from telescoping as a result of misplacing the timing of victimization. Second, we define two crime types, serious and small crimes, which may be defined too broadly so that our respondents have problems understanding which crimes belong to each category. Third, sorting or endogeneity as a result of unobserved individual characteristics that influence both victimization and well-being might play a role (see Cornaglia and Leigh, 2011). In contrast, the results of Helliwell (2006) and Ravallion and Lokshin (2001) suggest that accounting for potential endogeneity would not change the results significantly.

5.A Variables with explanation

Table A.1: Variables with explanation

(a) binary variables

female	1 if respondent is a woman
edu_middle	1 if respondent's highest education is secondary school
edu_high	1 if — at least vocational school
occup_pension	1 if — is retired or $\geq$ 65 years
occup_indep	1 if — works as independent entrepreneur or in a family firm
occup_nowork	1 if — has no occupation (incl. students)
partner	1 if — lives together with a partner (married or unmarried)
urban_high	1 if — lives in an urbanized area
urban_middle	1 if — in an area with intermediate urban character
victim_small	1 if — was victim of incorrect behavior in the past 5 years
victim_serious	1 if — of a serious crime in the past 5 years
victsmall_sev	1 if — was victim of a small crime in the past 5 years that is perceived severe.
victsmall_notsev	1 if — was victim of a small crime in the past 5 years that is perceived *not* severe.
victserious_sev	1 if — was victim of a serious crime in the past 5 years that is perceived severe.
victserious_notsev	1 if — was victim of a serious crime in the past 5 years that is perceived *not* severe.

Table A.1: Variables with explanation (cont.)

(b) non-binary variables

age	age of respondent (in years)
health	self-assessed health on a scale from 1: poor to 5: excellent
hh_lincome	log of gross monthly household income
mvictsmall_sev	0 if no victim, 1 if — was once victim, 2 if — was 2–5 times a victim, and 3 if more than 5 times victim of a small crime in the past 5 years that is perceived severe.
mvictsmall_notsev	0 if no victim, 1 if — was once victim, 2 if — was 2–5 times a victim, and 3 if more than 5 times victim of a small crime in the past 5 years that is perceived *not* severe.
mvictserious_sev	0 if no victim, 1 if — was once victim, 2 if — was 2–5 times a victim, and 3 if more than 5 times victim of a serious crime in the past 5 years that is perceived severe.
mvictserious_notsev	0 if no victim, 1 if — was once victim, 2 if — was 2–5 times a victim, and 3 if more than 5 times victim of a serious crime in the past 5 years that is perceived *not* severe.
social_norm	average of answers to short questions on severity of 18 small crimes on a scale from 1: not severe at all to 10: very severe
size_hh	number of members in a household
trust1	trust in others on a scale from 1: one cannot be very careful enough to 11: most people can be trusted
trust2	honesty of others on scale from 1: most people try to take advantage of others to 11: most people try to be honest.
trust3	helpfulness of others on a scale from 1: people are selfish to 11: people try to be helpful.
trust_index	degree of trust in other people (from 3: no trust to 33: maximum trust)
fear_rate	rate of people within a region that feel unsafe in 2008
vict_index	severity of crime(s) respondent has been victim of (from 0: not a victim to 20: victim of small and serious crime and both considered very severe)
vict_rate	rate of victimization within a region in 2008
well-being	subjective well-being on a scale from 1: very unhappy to 10: very happy

Bibliography

Abswoude, P. van, P. Tavenier, and M. van der Schee. Keystroke Dynamics and Two Factor Authentication. Master's thesis, University of Amsterdam.

Alesina, A., R. Di Tella, and R. MacCulloch (2004). Inequality and Happiness: Are Europeans and Americans Different? *Journal of Public Economics*, **88**, 2009–2042.

Balestrino, A. (2008). It is a Theft But Not a Crime. *European Journal of Political Economy*, **24**, 455–469.

Bardasi, E. and M. Francesconi (2004). The Impact of Atypical Employment on Individual Wellbeing: Evidence from a Panel of British Workers. *Social Science and Medicine*, **58**, 1671–1688.

Barefoot, J.C., K.E. Maynard, J.C. Beckham, B.H Brummett, K. Hooker, and I.C Siegler (1998). Trust, Health, and Longevity. *Journal of Behavioral Medicine*, **21**, 517–526.

Barnett, T., K. Bass, and G. Brown (1996). Religiosity, Ethical Ideology, and Intentions to Report a Peer's Wrongdoing. *Journal of Business Ethics*, **15**, 1161–1174.

Bartlow, N. and B. Cukic (2006). Evaluating the Reliability of Credential Hardening through Keystroke Dynamics. In *Proceedings of the 17th In-*

ternational Symposium on Software Reliability Engineering, pp. 117–126. Washington DC.

Beattie, J. and G. Loomes (1997). The Impact of Incentives Upon Risky Choice Experiments. *Journal of Risk and Uncertainty*, **14**, 155–168.

Becker, G.S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, **76**, 169–217.

Beek, K.W.H. van, C.C. Koopmans, and B.M.S. van Praag (1997). Shopping at the Labour Market: A Real Tale of Fiction. *European Economic Review*, **41**, 295–317.

Bergadano, F., D. Gunetti, and C. Picardi (2002). User Authentication through Keystroke Dynamics. *ACM Transactions on Information and System Security*, **5**, 367–397.

Bjørnskov, C. (2003). The Happy Few: Cross-Country Evidence on Social Capital and Life Satisfaction. *Kyklos*, **56**, 3–16.

Bjørnskov, C. (2006). The Multiple Facets of Social Capital. *European Journal of Political Economy*, **22**, 22–40.

Blanchflower, D.G. and A.J. Oswald (2004). Well-Being over Time in Britain and the USA. *Journal of Public Economics*, **88**, 1359–1386.

Bleha, S., C. Slivinsky, and B. Hussien (1990). Computer-Access Security Systems Using Keystroke Dynamics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, 1217–1222.

Bolton, R.J. and D.J. Hand (2002). Statistical Fraud Detection: A Review. *Statistical Science*, **17**, 235–249.

Britt, C.L. (2001). Health Consequences of Criminal Victimization. *International Review of Victimology*, **8**, 63–73.

Brown, M. and S.J. Rogers (1993). User Identification via Keystroke Characteristics of Typed Names Using Neural Networks. *International Journal of Man-Machine Studies*, **39**, 999–1014.

Bryan, W. and N. Harter (1899). Studies in the Telegraphic Language. *Psychological Review*, **6**, 345–375.

Camerer, C.F. and R.M. Hogarth (1999). The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty*, **19**, 7–42.

Chang, L. and J.A. Krosnick (2009). National Surveys via RDD Telephone Interviewing versus the Internet: Comparing Sample Representativeness and Response Quality. *Public Opinion Quarterly*, **73**, 641–678.

Cho, S., C. Han, H. Han, and H.-I Kim (2000). Web-Based Keystroke Dynamics Identity Verification Using Neural Network. *Journal of Organizational Computing and Electronic Commerce*, **10**, 295–307.

Clark, A.E. and A.J. Oswald (1994). Unhappiness and Unemployment. *The Economic Journal*, **104**, 648–659.

Clarke, N.L. and S.M. Furnell (2007). Advanced User Authentication for Mobile Devices. *Computers and Security*, **26**, 109–119.

Cohen, M.A. (2008). The Effect of Crime on Life Satisfaction. *Journal of Legal Studies*, **37**, S325–S353.

Cooter, R. (1998). Expressive Law and Economics. *Journal of Legal Studies*, **27**, 585–607.

Cornaglia, F. and A. Leigh (2011). Crime and Mental Wellbeing. CEP Discussion Paper 1049, Center for Economic Performance, London.

Czabański, J. (2008). *Estimates of Cost of Crime: History, Methodologies, and Implications*. Berlin: Springer.

Davies, S. and T. Hinks (2010). Crime and Happiness Amongst Heads of Households in Malawi. *Journal of Happiness Studies*, **11**, 457–476.

Denkers, A.J.M. and F.W. Winkel (1998). Crime Victims' Well-Being and Fear in a Prospective and Longitudinal Study. *International Review of Victimology*, **5**, 141–162.

Deutsch, M. (1958). Trust and Suspicion. *The Journal of Conflict Resolution*, **2**, 265–279.

Di Tella, R. and R.J. MacCulloch (2008). Gross National Happiness as an Answer to the Easterlin Paradox? *Journal of Development Economics*, **86**, 22–42.

Di Tella, R., R.J. MacCulloch, and H. Ñopo (2008). Happiness and Beliefs in Criminal Environments. Working Paper 662, Inter-American Development Bank, Washington DC.

Di Tella, R., R.J. MacCulloch, and A.J. Oswald (2003). The Macroeconomics of Happiness. *The Review of Economics and Statistics*, **85**, 809–827.

Diener, E., E.M. Suh, R.E. Lucas, and H.L. Smith (1999). Subjective Well-Being: Three Decades of Progress. *Psychological Bulletin*, **125**, 276–302.

Dolan, P., G. Loomes, T. Peasgood, and A. Tsuchiya (2005). Estimating the Intangible Victim Costs to Violent Crime. *The British Journal of Criminology*, **45**, 958–976.

Dolan, P. and S.C. Moore (2007). From Preferences to Experiences: Valuing the Intangible Victim Costs of Crime. *International Review of Victimology*, **14**, 265–280.

Dolan, P., T. Peasgood, and M. White (2008). Do We Really Know What Makes Us Happy? A Review of the Economic Literature on the Factors associated with Subjective Well-Being. *Journal of Economic Psychology*, **29**, 94–122.

Douhou, S. and J.R. Magnus (2009). The Reliability of User Authentication Through Keystroke Dynamics. *Statistica Neerlandica*, **63**, 432–449.

Douhou, S., J.R. Magnus, and A.H.O. van Soest (2011a). Peer Reporting and the Perception of Fairness. CentER Discussion Paper 2011–068, CentER, Tilburg.

Douhou, S., J.R. Magnus, and A.H.O. van Soest (2011b). The Perception of Small Crime. *European Journal of Political Economy*, **27**, 749–763.

Duin, R.P.W. and D.M.J. Tax (2005). Statistical Pattern Recognition. In C.H. Chen and P.S.P Wang (Eds.), *Handbook of Pattern Recognition and Computer Vision*, pp. 3–24. Singapore: World Scientific.

Dustmann, C. and F. Fasani (2011). Crime and Mental Health in the UK. `http://dipeco.economia.unimib.it/web/seminari/fasani.pdf`.

Easterlin, R.A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In P.A. David and M.W. Reder (Eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*, pp. 89–125. New York: Academic Press.

Ehrlich, I. (1973). Participation in Illegitimate Activities: A Theoretical and Empirical Investigation. *Journal of Political Economy*, **81**, 521–565.

Eide, E. (2001). Economics of Criminal Behavior. In B. Bouckaert and G. De Geest (Eds.), *Encyclopedia of Law and Economics: The Economics of Crime and Litigation*, pp. 345–389. Northampton: Edward Elgar.

Elster, J. (1989). Social Norms and Economic Theory. *The Journal of Economic Perspectives*, **3**, 99–117.

Evans, S.S. and J.E. Scott (1984). The Seriousness of Crime Cross-Culturally: The Impact of Religiosity. *Criminology*, **22**, 39–59.

Finch, J. (1987). The Vignette Technique in Survey Research. *Sociology*, **21**, 105–114.

Freeman, R.B. (1999). The Economics of Crime. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, pp. 3529–3571. Amsterdam: Elsevier.

Frey, B.S. and A. Stutzer (2002). What Can Economists Learn from Happiness Research? *Journal of Economic Literature*, **40**, 402–435.

Gaines, R.S., W. Lisowski, S.J. Press, and N. Shapiro (1980). Authentication by Keystroke Timing: Some Preliminary Results. RAND Publication Series R-2526-NSF, RAND Corporation, Santa Monica.

Gibbons, S. (2004). The Costs of Urban Property Crime. *The Economic Journal*, **114**, F441–F463.

Gladwell, M. (2005). *Blink: The Power of Thinking Without Thinking*. New York: Little, Brown, and Co.

Glaeser, E.L. and B. Sacerdote (1999). Why Is There More Crime in Cities? *Journal of Political Economy*, **107**, 225–258.

Glaeser, E.L., B. Sacerdote, and J.A. Scheinkman (1996). Crime and Social Interactions. *Quarterly Journal of Economics*, **111**, 507–548.

Gneezy, U. (2005). Deception: The Role of Consequences. *The American Economic Review*, **95**, 384–394.

Graaf, G. de (2010). A Report on Reporting: Why Peers Report Integrity and Law Violations in Public Organizations. *Public Administration Review*, **70**, 767–779.

Greenberg, J. (1990). Employee Theft as a Reaction to Underpayment Inequity: The Hidden Cost of Pay Cuts. *Journal of Applied Psychology*, **75**, 561–568.

Greenberg, J. (2002). Who Stole the Money, and When? Individual and Situational Determinants of Employee Theft. *Organizational Behavior and Human Decision Processes*, **89**, 985–1003.

Greenberg, J. and K.S. Scott (1996). Why Do Workers Bite the Hands that Feed Them? Employee Theft as a Social Exchange Process. In B.M. Staw and L.L. Cummings (Eds.), *Research in organizational behavior: Vol. 18. An annual series of analytical essays and critical reviews*, pp. 111–156. Greenwich: JAI Press Inc.

Gunetti, D. and C. Picardi (2005). Keystroke Analysis of Free Text. *ACM Transactions on Information and System Security*, **8**, 312–347.

Gutiérrez, F.J., M.M. Lerma-Rascón, L.R. Salgado-Garza, and F.J. Cantú (2002). Biometrics and Data Mining: Comparison of Data Mining-based Keystroke Dynamics Methods for Identity Verification. In *Proceedings of* *the Second Mexican International Conference on Artificial Intelligence*, pp. 460–469. Mérida.

Guven, A. and I. Sogukpinar (2003). Understanding Users' Keystroke Patterns for Computer Access Security. *Computers and Security*, **22**, 695–706.

Halman, L. and R. Luijkx (2008). What Happened to Dutch Values? Investigating General and Differential Trends in Values in the Netherlands. In T. Petterson and Y. Esmer (Eds.), *Changing Values, Persisting Cultures: Case Studies in Value Change*, pp. 175–198. Leiden: Koninklijke Brill.

Hanson, R.F., G.K. Sawyer, A.M. Begle, and G.S. Hubel (2010). The Impact of Crime Victimization on Quality of Life. *Journal of Traumatic Stress*, **23**, 189–197.

Helliwell, J.F. (2003). How's Life? Combining Individual and National Variables to Explain Subjective Well-Being. *Economic Modelling*, **20**, 331–360.

Helliwell, J.F. (2006). Well-Being, Social Capital and Public Policy: What's New? *The Economic Journal*, **116**, C34–C45.

Herzog, S. (2003). Does the Ethnicity of Offenders in Crime Scenarios Affect Public Perceptions of Crime Seriousness? A Randomized Survey Experiment in Israel. *Social Forces*, **82**, 757–781.

Herzog, S. and S. Oreg (2008). Chivalry and the Moderating Effect of Ambivalent Sexism: Individual Differences in Crime Seriousness Judgments. *Law and Society Review*, **42**, 45–74.

Hoquet, S., J.-Y. Ramel, and H. Cardot (2005). Fusion of Methods for Keystroke Dynamic Authentication. In *Proceedings of the 4th IEEE Workshop on Automatic Identification Advanced Technologies*, pp. 224–229. Washington DC.

Houser, D., S. Vetter, and J. Winter (2011). Fairness and Cheating. Working paper, ICES George Mason University, Fairfax.

Inglehart, R. (1996). The Diminishing Utility of Economic Growth: From Maximizing Security toward Maximizing Subjective Well-Being. *Critical Review*, **10**, 509–531.

Isenring, G.L. (2008). Perception of Seriousness and Concern about White-Collar Crime: Some Results of an Opinion Survey Among Swiss Banks. *European Journal of Criminal Policy and Research*, **14**, 371–389.

Jiang, C.-H., S. Shieh, and J.-C. Liu (2007). Keystroke Statistical Learning Model for Web Authentication. In *Proceedings of the 2nd ACM Symposium on Information, Computer, and Communications Security*, pp. 359–361. New York.

Jones, G.E. and M.J. Kavanagh (1996). An Experimental Examination of the Effects of Individual and Situational Factors on Unethical Behavioral Intentions in the Workplace. *Journal of Business Ethics*, **15**, 511–523.

Joyce, R. and G. Gupta (1990). Identity Authentication Based on Keystroke Latencies. *Communications of the ACM*, **33**, 168–176.

Kacholia, V. and S. Pandit (2003). Biometric Authentication Using Random Distributions (BioART). In *Proceedings of the 15th Annual Canadian Security Symposium*, pp. 1–8. Ottawa.

Kang, P., S. Park, S.-S. Hwang, H.-J. Lee, and S. Cho (2008). Improvement of Keystroke Data Quality Through Artificial Rhythms and Cues. *Computers and Security*, **27**, 3–11.

Keizer, K., S. Lindenberg, and L. Steg (2008). The Spreading of Disorder. *Science*, **322**, 1681–1685.

King, G. (1997). The Effects of Interpersonal Closeness and Issue Seriousness on Blowing the Whistle. *Journal of Business Communication*, **34**, 419–436.

King, G. and A. Hermodson (2000). Peer Reporting of Coworker Wrongdoing: A Qualitative Analysis of Observer Attitudes in the Decision to

Report versus Not Report Unethical Behavior. *Journal of Applied Communication Research*, **28**, 309–329.

King, G., C.J.L Murray, J.A. Salomon, and A. Tandon (2004). Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review*, **98**, 191–207.

Kingdon, G.G. and J. Knight (2003). Well-Being Poverty versus Income Poverty and Capabilities Poverty? Working Paper 199, Centre for the Study of African Economies, Oxford.

Koss, M.P., P.G. Koss, and W.J. Woodruff (1991). Deleterious Effects of Criminal Victimization on Women's Health and Medical Utilization. *Archives of Internal Medicine*, **151**, 342–347.

Koss, M.P., W.J. Woodruff, and P.G. Koss (1990). Relation of Criminal Victimization to Health Perceptions Among Women Medical Patients. *Journal of Consulting and Clinical Psychology*, **58**, 147–152.

Kube, S. and C. Traxler (2011). The Interaction of Legal and Social Norm Enforcement. *Journal of Public Economic Theory*, **13**, 639–660.

Kwak, N. and J. Oh (2009). Feature Extraction for One-Class Classification Problems: Enhancements to Biased Discriminant Analysis. *Pattern Recognition*, **42**, 17–26.

Kwan, Y.K., L.L. Chiu, W.C. Ip, and P. Kwan (2002). Perceived Crime Seriousness: Consensus and Disparity. *Journal of Criminal Justice*, **30**, 623–632.

Lee, H.-J. and S. Cho (2007). Retraining a Keystroke Dynamics-Based Authenticator with Impostor Patterns. *Computers and Security*, **26**, 300–310.

Leggett, J., G. Williams, M. Usnick, and M. Longnecker (1991). Dynamic Identity Verification via Keystroke Characteristics. *International Journal of Man-Machine Studies*, **35**, 859–870.

Lehmann, E.L. (1966). Some Concepts of Dependence. *Annals of Mathematical Statistics*, **37**, 1137–1153.

Lipton, D.L. and H.K.T. Wong (1985). Modern Trends in Authentication. *ACM SIGSAC Review*, **3**, 36–42.

Loog, M. and R.P.W. Duin (2004). Linear Dimensionality Reduction via a Heteroscedastic Extension of LDA: The Chernoff Criterion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26**, 732–739.

Mesmer-Magnus, J.R. and C. Viswesvaran (2005). Whistleblowing in Organizations: An Examination of Correlates of Whistleblowing Intentions, Actions, and Retaliation. *Journal of Business Ethics*, **62**, 277–297.

Michalos, A.C. and B.D. Zumbo (2000). Criminal Victimization and the Quality of Life. *Social Indicators Research*, **50**, 245–295.

Miethe, T.D. (1982). Public Consensus on Crime Seriousness: Normative Structure or Methodological Artifact? *American Society of Criminology*, **20**, 515–526.

Møller, V. (2005). Resilient or Resigned? Criminal Victimisation and Quality of Life in South Africa. *Social Indicators Research*, **72**, 263–317.

Monrose, F., M.K. Reiter, and S. Wetzel (2002). Password Hardening Based on Keystroke Dynamics. *International Journal of Information Security*, **1**, 69–83.

Monrose, F. and A.D. Rubin (1997). Authentication via Keystroke Dynamics. In *Proceedings of the 4th ACM Conference on Computer and Communications Security*, pp. 48–56. Zürich.

Monrose, F. and A.D. Rubin (2000). Keystroke Dynamics as a Biometric for Authentication. *Future Generation Computing Systems*, **16**, 351–359.

Myers, S.L. (1982). Crime in Urban Areas: New Evidence and Results. *Journal of Urban Economics*, **11**, 148–158.

Near, J.P. and M.P. Miceli (1985). Organizational Dissidence: The Case of Whistle-Blowing. *Journal of Business Ethics*, **4**, 1–16.

Obaidat, M.S. and B. Sadoun (1997). Verification of Computer Users Using Keystroke Dynamics. *IEEE Transactions on Systems, Man, and Cybernetics*, **27**, 261–269.

O'Connell, M. and A. Whelan (1996). Taking Wrongs Seriously: Public Perceptions of Crime Seriousness. *The British Journal of Criminology*, **36**, 299–318.

O'Fallon, M.J. and K.D. Butterfield (2005). A Review of the Empirical Ethical Decision-Making Literature: 1996–2003. *Journal of Business Ethics*, **59**, 375–413.

Orviska, M. and J. Hudson (2002). Tax Evasion, Civic Duty, and the Law Abiding Citizen. *European Journal of Political Economy*, **19**, 83–102.

Oudejans, M. and C.M. Vis (2008). Slachtoffers van (Poging tot) Oplichting. conducted for WODC EWB, CentERdata, Tilburg.

Payne, B.K., R.R. Gainey, R.A. Triplett, and M.J.E. Danner (2004). What Drives Punitive Believes? Demographic Characteristics and Justifications for Sentencing. *Journal of Criminal Justice*, **32**, 195–206.

Peacock, A., X. Ke, and M. Wilkerson (2004). Typing Patterns: A Key to User Identification. *IEEE Security and Privacy*, **2**, 40–47.

Pease, K. (1988). Judgements of Crime Seriousness: Findings from the 1984 British Crime Survey. Home Office Research and Planning Unit Paper 44, Great Britain Home Office, London.

Philipson, T.J. and R.A. Posner (1996). The Economic Epidemiology of Crime. *Journal of Law and Economics*, **39**, 405–433.

Piquero, N.L., S. Carmichael, and A.R. Piquero (2008). Assessing the Perceived Seriousness of White-Collar and Street Crimes. *Crime and Delinquency*, **54**, 291–312.

Poortinga, W. (2006). Social Capital: An Individual or Collective Resource for Health? *Social Science and Medicine*, **62**, 292–302.

Powdthavee, N. (2005). Unhappiness and Crime: Evidence from South Africa. *Economica*, **72**, 531–547.

Praag, B.M.S. van and B.E. Baarsma (2005). Using Happiness Surveys to Value Intangibles: The Case of Airport Noise. *The Economic Journal*, **115**, 224–246.

Putnam, R. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.

Ramchand, R., J.M. MacDonald, A. Haviland, and A.R. Morral (2009). A Developmental Approach for Measuring the Severity of Crimes. *Journal of Quantitative Criminology*, **25**, 129–153.

Ravallion, M. and M. Lokshin (2001). Identifying Welfare Effects from Subjective Questions. *Economica*, **68**, 335–357.

Riedel, M. (1975). Perceived Circumstances, Inferences of Intent and Judgments of Offense Seriousness. *Journal of Criminal Law and Criminology*, **66**, 201–208.

Roodman, D. (2009). Estimating Fully Observed Recursive Mixed-Process Models with CMP. CGD Working Paper 168, Center for Global Development, Washington DC.

Rose, A.M. and A.E. Prell (1955). Does the Punishment Fit the Crime? A Study in Social Valuation. *The American Journal of Sociology*, **61**, 247–259.

Rose, R. (2000). How Much Does Social Capital add to Individual Health? A Survey Study of Russians. *Social Science and Medicine*, **51**, 1421–1435.

Rosenmerkel, S.P. (2001). Wrongfulness and Harmfulness as Components of Seriousness of White-Collar Offenses. *Journal of Contemporary Criminal Justice*, **17**, 308–327.

Rossi, P.H., R.A. Berk, and A. Campbell (1997). Just Punishments: Guideline Sentences and Normative Consensus. *Journal of Quantitative Criminology*, **13**, 267–290.

Rossi, P.H., J.E. Simpson, and J.L. Miller (1985). Beyond Crime Seriousness: Fitting the Punishment to the Crime. *Journal of Quantitative Criminology*, **1**, 59–90.

Rossi, P.H., E. Waite, C.E. Bose, and R.E. Berk (1974). The Seriousness of Crimes: Normative Structure and Individual Differences. *American Sociological Review*, **39**, 224–237.

Sah, R.K. (1991). Social Osmosis and Patterns of Crime. *Journal of Political Economy*, **99**, 1272–1295.

Schlüter, A. and B. Vollan (2011). Morals as an Incentive? A Field Study on Honour Based Flower Picking. *European Review of Agricultural Economics*, **38**, 79–97.

Schonlau, M., W. DuMounchel, W.-H. Ju, A.F. Karr, M. Theus, and Y. Vardi (2001). Computer Intrusion: Detecting Masquerades. *Statistical Science*, **16**, 58–74.

Schrager, L. Shill and J.F. Short (1980). How Serious a Crime? Perceptions of Organizational and Common Crimes. In G. Geis and E. Stotland (Eds.), *White-Collar Crime: Theory and Research*, pp. 14–31. Beverly Hills: Sage.

Sellin, T. and M.E. Wolfgang (1964). *The Measurement of Delinquency*. New York: John Wiley and Sons.

Shorack, G.R. and J.A. Wellner (1986). *Empirical Processes with Applications to Statistics*. New York: John Wiley and Sons.

Sims, R.L. and J.P. Keenan (1998). Predictors of External Whistleblowing: Organizational and Intrapersonal Variables. *Journal of Business Ethics*, **17**, 411–421.

Song, D., P. Venable, and A. Perrig (1997). User Recognition by Keystroke Latency Pattern Analysis. `http://citeseer.ist.psu.edu/song97user.html`.

Statistics Netherlands (2008). *Jaarboek Onderwijs in Cijfers 2008*. Voorburg/Heerlen: Statistics Netherlands.

Stutzer, A. and R. Lalive (2004). The Role of Social Work Norms in Job Searching and Subjective Well-Being. *Journal of the European Economic Association*, **2**, 696–719.

Stylianou, S. (2003). Measuring Crime Seriousness Perceptions: What Have We Learned and What Else Do We Want To Know. *Journal of Criminal Justice*, **31**, 37–56.

Sumner, L.W. (1996). *Welfare, Happiness and Ethics*. Oxford: Clarendon Press.

Traxler, C. (2010). Social Norms and Conditional Cooperative Taxpayers. *European Journal of Political Economy*, **26**, 89–103.

Traxler, C. and J.K. Winter (2012). Survey Evidence on Conditional Norm Enforcement. *European Journal of Political Economy*, **28**, 390–398.

Trevino, L. and B. Victor (1992). Peer Reporting of Unethical Behavior: A Social Context Perspective. *The Academy of Management Journal*, **35**, 38–64.

Umphress, D. and G. Williams (1985). Identity Verification Through Keyboard Characteristics. *International Journal of Man-Machine Studies*, **23**, 263–273.

Victor, B., L. Trevino, and D.L. Shapiro (1993). Peer Reporting of Unethical Behavior: The Influence of Justice Evaluations and Social Context Factors. *Journal of Business Ethics*, **12**, 253–263.

Wärneryd, K.E. and B. Walerud (1982). Taxes and Economic Behavior: Some Interview Data on Tax Evasion in Sweden. *Journal of Economic Psychology*, **2**, 187–211.

Warr, M. (1984). Fear of Victimization: Why Are Women and the Elderly More Afraid? *Social Science Quarterly*, **65**, 681–702.

Warr, M. (1989). What is the Perceived Seriousness of Crimes? *Criminology*, **27**, 795–821.

Weibull, J.W. and E. Villa (2005). Crime, Punishment and Social Norms. SSE/EFI Working Paper 610, Stockholm School of Economics, Stockholm.

Winkelmann, L. and R. Winkelmann (1998). Why Are the Unemployed So Unhappy? Evidence from Panel Data. *Economica*, **65**, 1–15.

Winkelmann, R. (2005). Subjective Well-Being and the Family: Results from an Ordered Probit Model with Multiple Random Effects. *Empirical Economics*, **30**, 749–761.

Yu, E. and S. Cho (2004). Keystroke Dynamics Identity Verification: Its Problems and Practical Solutions. *Computers and Security*, **23**, 428–440.

Zeng, Z.H., Y. Fu, G.I. Roisman, Z. Wen, Y.X. Hu, and T.S. Huang (2006). One-Class Classification for Spontaneous Facial Expression Analysis. In *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, pp. 281–286. Southampton.

CentER dissertation series

CentER for Economic Research, Tilburg University, The Netherlands

No.	Author	Title	ISBN	Published
247	John Kleppe	Modelling Interactive Behaviour, and Solution Concepts	978 90 5668 248 4	January 2010
248	Elleke Janssen	Inventory Control in Case of Unknown Demand and Control Parameters	978 90 5668 249 1	May 2010
249	Jun Zhou	Access to Justice: An Economic Approach	978 90 5668 251 4	April 2010
250	Henry van der Wiel	Competition and Innovation: Together a Tricky Rollercoaster for Productivity	978 90 5668 250 7	April 2010
251	Andrea Krajina	An M-Estimator of Multivariate Tail Dependence	978 90 5668 252 1	April 2010
252	Lenny Visser	Thresholds in Logistics Collaboration Decisions: A Study in the Chemical Industry	978 90 8891 172 9	June 2010
253	Dirk Broeders	Essays on the Valuation of Discretionary Liabilities and Pension Fund Investment Policy	978 90 5668 254 5	June 2010
254	Barbara Flügge	IS Standards in Designing Business-to-Government Collaborations – the Case of Customs	978 90 5668 253 8	June 2010

No.	Author	Title	ISBN	Published
255	Gerwald van Gulick	Game Theory and Applications in Finance	978 90 5668 255 2	June 2010
256	Renxiang Dai	Essays on Pension Finance and Dynamic Asset Allocation	978 90 5668 256 9	June 2010
257	Maria Gantner	Some Nonparametric Diagnostic Statistical Procedures and their Asymptotic Behavior	978 90 5668 257 6	September 2010
258	Marcel Hiel	An Adaptive Service Oriented Architecture – Automatically solving Interoperability Problems	978 90 5668 258 3	September 2010
259	Alma Timmers	The Perceived Cultural Changes and the Changes in Identification of the Employees during a Merger between Two Airlines	978 90 5668 259 0	August 2010
260	Femke van Horen	Breaking the Mould on Copycats: What makes Product Imitation Strategies Successful?	978 90 5668 260 6	September 2010
261	Owen Powell	Essays on Experimental Bubble Markets	978 90 5668 261 3	September 2010
262	Vasilios Andrikopoulos	A Theory and Model for the Evolution of Software Services	978 90 5668 262 0	October 2010
263	Carlos Jorge Da Silva Lourenço	Consumer Models of Store Price Image Formation and Store Choice	978 90 5668 263 7	November 2010
264	Hsing-Er Lin	Effects of Strategy, Context, and Antecedents and Capabilities on Outcomes of Ambidexterity – A Multiple Country Case Study of the US, China and Taiwan	978 90 5668 264 4	December 2010

No.	Author	Title	ISBN	Published
265	Jeffrey Powell	The Limits of Economic Self-Interest: The Case of Open Source Software	978 90 5668 265 1	December 2010
266	Galla Salganik	Essays on Investment Flows of Hedge Fund and Mutual Fund Investors	978 90 5668 266 8	December 2010
267	Cécile Fruteau	Biological Markets in the Everyday Lives of Mangabeys and Vervets: An Observational and Experimental Study	978 90 5668 267 5	December 2010
268	Yang Zhao	The Role of Directors' Professional and Social Networks in CEO Compensation and the Managerial Labour Market	978 90 5668 268 2	December 2010
269	Mieszko Mazur	Essays on Managerial Remuneration, Organizational Structure and Non-Cash Divestitures	978 90 5668 269 9	December 2010
270	Ralph Stevens	Longevity Risk in Life Insurance Products	978 90 5668 270 5	February 2011
271	Cristian Dobre	Semidefinite Programming Approaches for Structured Combinatorial Optimization Problems	978 90 5668 271 2	March 2011
272	Kenan Kalayci	Essays in Behavioral Industrial Organization	978 90 5668 272 9	February 2011
273	Néomie Raassens	The Performance Implications of Outsourcing	978 90 5668 273 6	March 2011

No.	Author	Title	ISBN	Published
274	Muhammad Ather Elahi	Essays on Financial Fragility	978 90 5668 274 3	March 2011
275	Jan Stoop	Laboratory and Field Experiments on Social Dilemmas	978 90 5668 275 0	March 2011
276	Maria Cristina Majo	A Microeconometric Analysis of Health Care Utilization in Europe	978 90 5668 276 7	February 2011
277	Jérémie Lefebvre	Essays on the Regulation and Microstructure of Equity Markets	978 90 5668 277 4	March 2011
278	Willem Muhren	Foundations of Sensemaking Support Systems for Humanitarian Crisis Response	978 90 5668 278 1	March 2011
279	Mary Pieterse-Bloem	The Effect of EMU on Bond market Integration and Investor Portfolio Allocations	978 90 5668 279 8	May 2011
280	Chris Müris	Panel Data Econometrics and Climate Change	978 90 5668 280 4	April 2011
281	Martin Knaup	Market-Based Measures of Bank Risk and Bank Aggressiveness	978 90 5668 281 1	April 2011
282	Thijs van der Heijden	Duration Models, Heterogeneous Beliefs, and Optimal Trading	978 90 5668 282 8	May 2011
283	Titus Galama	A Theory of Socioeconomic Disparities in Health	978 90 5668 283 5	June 2011

No.	Author	Title	ISBN	Published
284	Hana Vonková	The Use of Subjective Survey Data: Anchoring Vignettes and Stated Preference Methods	978 90 5668 284 2	May 2011
285	Frans Stel	Improving the Performance of Co-Innovation Alliances: Cooperating Effectively with New Business Partners	978 90 5668 285 9	July 2011
286	Eric Engesaeth	Managerial Compensation Contracting	978 90 5668 286 6	June 2011
287	David Kroon	The Post-Merger Integration Phase of Organizations: A Longitudinal Examination of Unresolved Issues of Justice and Identity	978 90 5668 287 3	May 2011
288	Christian Bogmans	Essays on International Trade and the Environment	978 90 5668 288 0	June 2011
289	Kim Peijnenburg	Consumption, Savings, and Investments over the Life Cycle	978 90 5668 289 7	May 2011
290	Youtha Cuypers	The Determinants and Performance Implications of Change in Inter-Organizational Relations	978 90 5668 290 3	June 2011
291	Marta Serra Garcia	Communication, Lending Relationships and Collateral	978 90 5668 291 0	June 2011
292	Marike Knoef	Essays on Labor Force Participation, Aging, Income and Health	978 90 5668 292 7	September 2011
293	Christophe Spaenjers	Essays in Alternative Investments	978 90 5668 293 4	September 2011

No.	Author	Title	ISBN	Published
294	Moazzam Farooq	Essays on Financial Intermediation and Markets	978 90 5668 294 1	September 2011
295	Jan van Tongeren	From National Accounting to the Design, Compilation, and Use of Bayesian Policy Analysis Frameworks	978 90 5668 295 8	October 2011
296	Lisanne Sanders	Annuity Market Imperfections	978 90 5668 296 5	October 2011
297	Miguel Atanásio Lopes Carvalho	Essays in Behavioral Microeconomic Theory	978 90 5668 297 2	September 2011
298	Marco Della Seta	Essays in Corporate Financing and Investment under Uncertainty	978 90 5668 298 9	October 2011
299	Roel Mehlkopf	Risk Sharing with the Unborn	978 90 5668 299 6	October 2011
300	Roy Lindelauf	Design and Analysis of Covert Networks, Affiliations and Projects	978 90 5668 300 9	October 2011
301	Viswanadha Reddy	Essays on Dynamic Games	978 90 5668 301 6	November 2011
302	Pedro Duarte Bom	The Macroeconomics of Fiscal Policy and Public Capital	978 90 5668 302 3	November 2011
303	Daniël Smit	Freedom of Investment between EU and non-EU Member States and its Impact on Corporate Income Tax Systems within the European Union	978 90 5668 303 0	December 2011

No.	Author	Title	ISBN	Published
304	Juan Miguel Londoño Yarce	Essays in Asset Pricing	978 90 5668 304 7	December 2011
305	Edwin Lohmann	Joint Decision Making and Cooperative Solutions	978 90 5668 305 4	January 2012
306	Verena Hagspiel	Flexibility in Technology Choice: A Real Options Approach	978 90 5668 306 1	December 2011
307	Fangfang Tan	Behavioral Heterogeneity in Economic Institutions: An Experimental Approach	978 90 5668 307 8	January 2012
308	Luc Bissonnette	Essays on Subjective Expectations and Stated Preferences	978 90 5668 308 5	January 2012
309	Unnati Saha	Econometric Models of Child Mortality Dynamics in Rural Bangladesh	978 90 5668 309 2	February 2012
310	Anne ter Braak	A New Era in Retail: Private-Label Production by National-Brand Manufacturers and Premium-Quality Private Labels	978 90 5668 310 8	February 2012
311	Emiel Maasland	Essays in Auction Theory	978 90 5668 311 5	February 2012
312	John Glen	Essays on the UK Residential Property Market	978 90 5668 312 2	March 2012
313	Amal Elgammal	Towards a Comprehensive Framework for Business Process Compliance	978 90 5668 313 9	April 2012

No.	Author	Title	ISBN	Published
314	David Hollanders	The Effect of Aging on Pensions	978 90 5668 315 3	June 2012
315	Salima Douhou	Essays on (Small) Crime: Perception, Social Norms, Happiness, and Prevention	978 90 5668 316 0	June 2012