Collegio Carlo Alberto

Optimal Criminal Behavior and the Disutility of Jail: Theory and Evidence On Bank Robberies

Giovanni Mastrobuoni

No. 220

November 2011

Carlo Alberto Notebooks

www.carloalberto.org/working_papers

Optimal Criminal Behavior and the Disutility of Jail: Theory and Evidence On Bank Robberies*

Giovanni Mastrobuoni†

First version March 2009, this version November 2011‡

Abstract

Based on unique data on individual bank robberies perpetrated in Italy between 2005 and 2007, this paper estimates the distribution of criminals' disutility of jail. The identification rests on the money versus risk trade-off criminals face when deciding whether to stay an additional minute robbing the bank. When robbers are successful in robbing a bank and the observed duration is assumed to be the optimal one, the disutility of jail represents the only unknown determinant of that optimal duration. One can thus solve for the disutility as a function of the expected marginal haul, average haul, and hazard rate of arrest.

Keywords: Crime, Deterrence, Sentence Enhancements, Bank Robberies, Value of Freedom, Disutility of Jail
JEL classification codes: K40, K42, H11

---

*I had many useful conversations about this paper with Alex Tetenov. I would also like to thank two anonymous referees, Gani Aldashev, Phil Cook, Andrea Ichino, Jens Ludwig, Dean Lueck, Costas Meghir, Steven Raphael, Ricardo Reis, Vincenzo Verardi and all other seminar participants at Collegio Carlo Alberto, NBER Law and Economics Summer Institute, Universitat Autonoma de Barcelona, University of Milano Cattolica, University of Namur, Universitat Pompeu Fabra, Osaka University, 2010 American Law and Economics Association meetings, 1st Bonn and Paris workshop on the Economics of Crime, and AEA meeting in Atlanta. Martino Bernardi and Vipin Veetil provided excellent research assistance.

†Collegio Carlo Alberto and CeRP, giovanni.mastrobuoni@carloalberto.org

"Every second past two minutes increased the odds that a bank robber would be caught. A professional would leave a bank when the clock struck two whether he had the money or not. Lynn Phelps knew these guys were amateurs, dicking around in the bank for nine minutes." (Crais, 2007)

1 Introduction

This paper uses unique data on individual Italian bank robberies organized between 2005 and 2007 to identify criminals' *individual* disutility of jail. The identification rests on a model of crime where the criminal, after selecting a bank, a weapon, a masquerade, and a team, chooses the optimal duration of the robbery. In line with the opening quote from Robert Craig's fiction "The Two Minute Rule," bank robbers choose the optimal duration of the bank robbery given the expected benefits (more loot) and expected costs (more risk) of staying an additional minute inside the bank. The model's first order conditions are used to solve for the only unobserved variable: the disutility of jail time.

Since the disutility of jail depends on the expected sentence length–a longer jail time leads to larger losses in utility– this paper identifies the individual responsiveness to sentencing.[1] Results show that the distribution of disutility of jail time is positively skewed and resembles an earnings distribution. This heterogeneity in the criminals' "fear of jail" might depend on how much they discount the future (DiIulio, 1996), but is also likely to depend on their opportunity cost of spending jail time. Changes in the disutility of jail are predicted to lead to significant changes in criminal behavior, and these changes are larger among criminals with a higher opportunity cost of spending jail time. These more "able" criminals are the ones that during robberies are masqueraded, use firearms, and work in groups. Interestingly, these are the same *modi operandi* (modes of operation) that lead to statutory sentence enhancements. I'll show that Italian judges do not follow these prescriptions close enough. Only the use of firearms leads to significant sentence enhancements.

Apart from reconciling statutory and actual sentence enhancements, harshening the rather mild sanctions (the average sentence for a bank robbery is 3.3 years in Italy while its 11.4 years in the US) would be another way to reduce Italy's dramatic number of bank robberies, especially the more profitable ones. Moreover, I show that the harshening would first drive the most able offenders out of the bank robbery business.

[1]The disutility might also be used to evaluate more comprehensively the cost and benefits of various aspects of the criminal justice system, for example it is part of the social cost of incarceration (Barbarino and Mastrobuoni, 2008).

This finding that more "able" criminals respond more strongly to general deterrence might explain why in the US, where jail sentences for robbing a bank are on average 400 percent higher the pool of robbers is considered to be more amateurish than in Italy (Weisel, 2007).

These results are robust to the inclusion of unobserved ability or, equivalently, heterogeneity in expectations with respect to, both, the haul and the risk, and to the inclusion of measurement error in the duration of the bank robberies. The results also show that large degrees of risk aversion (log utility) are clearly rejected by the data. Interestingly, there are no empirical studies that I am aware of that use observational data to try to estimate the criminals' degree of risk aversion.

There is only one other paper, Abrams and Rohlfs (2010), that estimates the average disutility of jail, which the authors call the *Value of Freedom*. Based on data on posting bails the authors' estimate is around $4,000 per year; they explain this low figure by saying that "(t)his seemingly low estimate may result in part because they pertain to a particularly poor segment of the population. Credit constraints may also affect the estimate." This paper goes beyond estimating the average disutility of jail, backing out, under some parametric assumptions, its distribution.

This paper is also related to the vast amount of papers that have tried to find evidence of deterrence. A rational model of crime predicts that criminals commit an additional crime whenever the expected marginal utility that they derive from the crime is larger than the expected marginal sanction (Becker, 1968, Ehrlich, 1973, Freeman, 1999). In its simplest version individuals gain from successful crimes with some probability $1 - p$ and risk being apprehended and spending S years in jail with probability p. The individual commits a crime whenever the marginal benefits outweigh the marginal costs. Even though this model is simple and intuitive it has been difficult to estimate. Data are typically aggregated across space and time, which makes it difficult to measure legal and illegal earnings (Vicusi, 1986b). Measurement errors have plagued the measurement of expected benefits and simultaneity issues (policy makers increase police enforcement and the severity of sanctions when crime levels are high) have made the estimation of the deterrence effect of the probability of apprehension p difficult. The disutility of jail $U(S)$, which can also be interpreted as the opportunity cost of jail, is not observed and makes the estimation still more challenging. Moreover, extensions of the model that would increase its realism–such as additional allocations of time, the effect of crime or apprehension in one period on future legitimate and criminal earnings, the risk that a criminal is victimized by other criminals, and the degree of social stigma associated with crime–complicate the

estimation even further.

Some studies have tried to find evidence on deterrence using individual level data on perceived deterrence but such data is usually based on prison surveys (Polich et al., 1980) or on other self-reported crime data (Grogger, 1998, Glaeser and Sacerdote, 1999).[2] In both, surveys and self-reports, crime activities might be subject to untruthful reporting or at least to underreporting (Vicusi, 1986a). Kessler and Levitt (1999) use the introduction of sentence enhancement while Helland and Tabarrok (2007) use a quasi-randomization of sentence enhancements to isolate deterrence and find strong evidence of it. Lee and McCrary (2005), instead, find very little evidence of deterrence among juvenile criminals who move to the adult sanctioning system: their criminal behavior changes very little upon turning 18. Drago et al. (2009) use an Italian quasi-experimental setting and find evidence of deterrence. All these studies estimate *average* deterrence effects. This paper goes a step further, identifying individual responsiveness to sanctions.

In spirit this paper is also related to the vast literature that tries to estimate the value of life based on trade offs between fatality risk and different kinds of returns, for example wage premia in the labor market (Thaler and Rosen, 1976, Viscusi, 1993), or the saving of time when driving (Ashenfelter and Greenstone, 2004).

Section 2 describes the data. A unique feature of the data is that it tells us many minutes the robbery lasted. The duration represents the robbers' control variable in the model outlined in Section 3, which is later estimated in Section4.

2 "The Italian Job"

According to the Uniform Crime Statistics each year in the U.S. there are around 10,000 bank robberies, representing more than 10 percent of all commercial robberies, with an average loss of 4,000 dollars (Weisel, 2007). Relative to its size, Italy faces a far greater problem. Each year there are more bank robberies in Italy than in the rest of Europe put together: approximately 3,000. Data from the European Banking Federation reveal that Italy is followed by Canada and Germany, which have around 800 robberies per year, and by Spain with 500 (Table 1). The U.S., which is not part of the Federation, has more than 5 times the population of Italy but just 3 times as many bank robberies (Weisel, 2007).

Low probabilities of apprehension, large cash holdings, but also mild sentencing, and the banks' fear that more stringent security devices would lead to a loss of clients are

[2]Nagin (1998) and Cameron (1988) survey the hundreds of papers written on deterrence.

believed to be the main drivers of Italy's high number of bank robberies. And the trend over time is not wholly encouraging. Figure 1 shows the average haul (right axis) and the number of bank robberies (left axis) between 1990 and 2003. While the average haul went down, the number of bank robberies went from around 1,500 in the early 90s to almost double that number 10 years later.

Perceived costs of robbing banks depend on the probability of apprehension and on the expected sanctions. More than 90 percent of Italian bank robberies end up without an arrest, while in the U.S. 33 percent of bank robbers are arrested on the same day they commit the robbery. Moreover, US federal guidelines impose sentences of *at least* 20 years (plus 5 years when a weapon is used), while in Italy the sentence length ranges between 3 and 10 years depending on the severity of the crime. The range becomes 4.5 to 20 years when at least one of the following conditions is satisfied (art. 628 of the penal code): a weapon is used, the robber is masked or he is not alone, violence is used to incapacitate a victim, the robber belongs to an organized crime association.

The expected costs of robbing a bank are, therefore, considerably lower in the Italy than in the US. What about the expected benefits? Robbing a bank seems to pay. The average haul is 20,000 euro (in the US it is approximately 6,000 euro). This leads to a direct cost for society of more than 57 million euro a year. But the indirect cost is even larger. A survey of 21,000 retail bank branches representing 65 percent of all Italian branches shows that in 2006 banks spent an average of 10,700 euro per branch to prevent bank robberies (a total of more than 300 million euro (OSSIF, 2006)). Each branch spent an additional 4,900 euro to prevent thefts and 6,300 euro to protect financial couriers. Therefore, the total amount spent by banks in 2006 to prevent thefts and robberies was more than 700 million euro. This might in part explain why Italian banks charge on average the largest account management fees in Europe: 90 euro against a European average of just 14 euro (European Commission, 2007). Moreover, Miller-Burke et al. (1999) show that in the U.S. most employees have multiple negative health consequences from experiencing a bank robbery while at work, including anxiety and post-traumatic stress disorder. This is unlikely to be very different in Italy and generates an additional cost.

Despite these frightening numbers, there is, to the best of my knowledge, almost no empirical research in economics and very little research in criminology that has tried to study bank robberies using robbery-level data. One reason for this is certainly the lack of data. Several studies describe in great detail robberies (Cook, 2009, 1990, 1987, 1986, 1985) and bank robberies in particular (Federal Bureau of Investigation, 2007, Weisel, 2007, Baumer and Carrington, 1986), but only one study–Hannan (1982)–tries to test

deterrence explicitly using data on bank robberies and banks' security devices. The major shortcoming of Hannan (1982) is that the adoption of new security devices depends on past robberies, which might explain why the author finds no significant effects of the presence of security devices on robberies.

I have been granted access to a unique data set: the universe of individual bank robberies perpetrated in Italy between 2005 and 2007. The data are divided into two parts, robbery-level data and branch-level data, that are merged together. Each year branch managers are required to update the characteristics of their branch (security devices, number of employers, etc). Moreover, after each robbery branch managers are required to fill out a form describing the facts (i.e number of bank robbers, haul, weapons, technique, etc.). The initial number of robberies is 6,434 but 1,215 are excluded because of missing information on either the robbery or the characteristics of the branch. Managers also have to record the exact duration of the robbery in minutes. All bank branches have surveillance cameras that can be used to reconstruct the exact timing. Nevertheless, there is evidence of heaping in the duration of the robberies. Figure 2 plots the distribution and the cumulative distribution of the durations between 1 and 46 minutes. The 46th minute stands for robberies that last more than 45 minutes. There are 185 of them out of 5,219, or 3.54 percent, while almost 90 percent of robberies last at most 9 minutes. The spikes at multiples of 5 show the heaping. For example, there are no observations between 30 and 40 minutes other than at 30, 35, and 40 minutes. For three reasons the results are more reliable when the analysis is restricted to bank robberies that last less than 10 minutes: i) it minimizes measurement error bias that is due to heaping; ii) because 90 percent of robberies last less than 10 minutes and sufficient data is needed to estimate hazard rates; iii) because the model outlined in Section 3, and the assumed smoothness of the objective function are more likely to apply to shorter robberies. But as a robustness check all the analysis is also performed using robberies that last up to 30 minutes (after 30 minutes too few observations are left). Since heaping introduces measurement error, in Section 4.3.4 I address how measurement error biases the estimation.

Table 3 presents the distribution of durations below 30 minutes separated into successful (no arrest) and unsuccessful (arrest) robberies. At time 0 the sample starts with 4,972 robberies that last less than 30 minutes. 297 last just one minute. Of these 24 lead to an arrest and 273 don't, and are labeled as successful, even if the robbers walk out of the bank empty-handed. After the first minute 4675 robberies are left, of which 71 lead to an arrest and 1,049 terminate without an arrest, and so on.

The summary statistics in Table 4 show that between 2005 and 2007 only 6.33 percent

of bank robbers were arrested after robberies that lasted less than 10 minutes.[3] The typical robbery lasts around 3.2 minutes and leads to a haul of approximately 14,000 euro. Given that more than half of all bank robberies involve two or more perpetrators the average haul per criminal is smaller or equal to approximately 8,000 euros. Only 14 percent of bank robberies involve firearms, as judges sanction their use with increased punishments. 43 percent of all bank robbers mask their face when robbing a bank. 21 percent of bank robberies happen in central Italy, 28 percent in the South and the rest in the North.[4] When compared to the distribution in the population of branches, bank robbers are more likely to choose banks that have on average smaller amounts of cash, or banks that are located in less populous areas.

The data set is rich with information about the security devices installed in the bank. I summarize this information by counting the number of different devices that each bank has, and compute how many characteristics these devices have on average for each bank. For example, 92 percent of the banks have a security entrance but the characteristics differ widely. Some have metal detectors, some have a double door where people can be trapped, some have a biometric sensor, etc., while some entrances might display all these characteristics. Robbed banks tend to have more security devices installed than the average bank (7.2 versus 6.7), and these devices tend to have more characteristics per device. The main reason for this is that banks tend to install new devices after they experience a bank robbery. The majority of these devices are not visible to the criminal (like automatic banknote distributors, banknote spotters, time-delayers, banknote tracing devices, vaults, and alarm systems) while 33 percent are clearly visible (like metal detectors, vault's time-locks, and protected teller's post). Since visible and invisible devices might have a different impact on the robbery, I will control for the fraction of invisible devices. The last 4 columns of Table 4 allow a comparison between the summary statistics of robberies that last more or less than the median, which is equal to 3 minutes. The average duration of robberies is 2.44 minutes for those that last less than 3 minutes and 4.93 minutes for those the last more than 3 minutes. This difference translates into slightly larger probabilities of arrest 6.28 vs. 6.43 percent, but considerably larger hauls, 11,559 versus 18,469 euro. These differences can in part be explained by differences in the *modus operandi*. Longer robberies are more likely to be operated by teams (75 versus

[3]Fifty-nine percent of these arrests happen during the bank robbery, while the rest happens afterwards. All the results are qualitatively similar when I exclude the robberies where the arrests do not happen immediately.

[4]The following central regions separate the southern regions from the Northern ones: Lazio, Marche, Toscana, Molise, and Umbria.

62 percent), and in longer operations robbers are more likely to be using a firearm (16 versus 12 percent). Given that the *modus operandi* is likely to influence not only the duration but also the probability of success and the expected haul, it is important to control for it when I model the bank robbers' decision about the duration of the bank robbery. The other observable characteristics of branches show only minor differences based on the duration of the robberies.

The data that were provided by the Italian Banking Association do not contain any information about the robbers.[5] In order to have some information about typical sentences I also collected judiciary level data on 95 bank robbers that were sentenced to jail in the judicial district of Turin, a city in Northern Italy (more on this in Section 4.3.1). Unfortunately the judiciary level data does not provide enough detail about the robberies to link them to the robbery data.

3 A Continuous Time Version of Becker's Model of Crime

Bank robbers face an obvious trade-off: the longer they stay inside the bank the more money they are able to collect, but the risk of getting caught goes up as well. In this section I model this trade-off in order to identify the criminal's disutility of jail. Conditional on having chosen to rob a bank the criminal's expected utility $V(t, a)$ is a function of the duration of the bank robbery, and of his own ability a, which, once the robber starts a robbery, he cannot modify.[6]

$$\begin{aligned} V(t,a) &= [1 - P(T < t|a)]E[U[Y(t,a)]] - P(T < t|a)D(a,w) \\ &= [1 - F(t|a)]E[U[Y(t,a)]] - F(t|a)D(a,w), \end{aligned} \tag{1}$$

where $P(T < t|a) = F(t|a)$ represents the probability of apprehension before time t. The random variable T defines the two states of the world, arrest $T < t$ and no arrest $T \geq t$), and is influenced by the ability of the robbers. $E[U[Y(t, a)]]$ is the expected utility from a haul equal to Y, which also depends on the duration of the robbery.

$D(a, w)$ represents the disutility from jail, which is unobserved to the econometrician

[5]But it should be noted that such information would come from arrested robbers and thus likely be biased because of selection.

[6]The model assumes rational and forward-looking behavior which is more likely to be valid for experienced robbers. But more than two thirds of sentenced robbers are recidivists, with some level of experience.

and depends on the robber's criminal ability a and on other factors w, like legitimate earnings, family composition, fear of physical harm, etc. The reason criminal ability determines the robber's disutility is that the opportunity cost of ending up in jail raises with ability.

The optimal duration of a bank robbery t^* is determined by a criminal with ability a by solving:

$$-F'(t^*)[E[U[Y(t^*,a)]]+D(a)]+[1-F(t^*|a)]E[U'[Y(t^*,a)]]=0, \tag{2}$$

where $U'=\frac{\partial U}{\partial t}$. Notice that allowing robbers to be arrested with some exogenous probability q, no matter the duration of the bank robbery, means that the expected utility would be $W(t,a,q)=qD+(1-q)V(t,a)$ and one would still get the same first order condition shown in Equation refeq:FOC. Following this, all the results are robust to arrests that happen independently on t and are not recorded by the Italian Banking Association.

Solving the first order condition for the disutility of jail D gives

$$\begin{aligned} D(a,w) &= \frac{1-F(t^*|a)}{F'(t^*|a)}E[U'[Y(t^*,a)]]-E[U[Y(t^*,a)]] \\ &= \frac{1}{\lambda(t^*|a)}E[U'[Y(t^*,a)]]-E[U[Y(t^*,a)]]. \end{aligned} \tag{3}$$

Thus conditional on ability a the chosen duration of robbers t^* identifies their disutility $D(a,w)$. I assume that the utility function is linear $U(\cdot)=Y(\cdot)$, which implies risk neutrality. Using this assumption one does not need to specify an initial level of wealth (Block and Heineke, 1975). Section 4.5 shows how the results change when one departs from risk neutrality.

There are two main measurement issues related to Equation 3: i) individual expectations are unobserved, and ii) even if we knew the functional form of the expected utility and of the hazard rate, the ability would still be unobserved.

There are two ways to measure expectations and ability, and each one has its own advantages and disadvantages, parametrically and semi-parametrically.

3.1 Parametric Estimates

Experienced robbers of high ability should know how to best perform a bank robbery. They know which banks are more vulnerable and less risky; they know which weapon they should use, and whether it pays to work in a group; they understand that using a

masquerade might reduce their likelihood of getting caught. All these choices, let's call them *modus operandi* x, are likely to carry a signal about the ability of a robber, or of a group of robbers. Allowing the ability to minimize risk a_r to differ from the ability to maximize the haul a_h these signals can be weighted using a linear function:

$$a_r = x'\beta_r + u_r \tag{4}$$

$$a_h = x'\beta_h + u_h \quad . \tag{5}$$

The random variables u_i measure the part of the ability that is not captured by $x'\beta_i$ ($i = h, r$), or simply some incidents that happen during the bank robbery and are observable to the robbers but not the econometrician. Let me repeat that robbers choose which bank to target, and how to carry out the robbery, in other words, x is endogenous. Harding (1990), for example, interviewing almost 500 robbers finds that most of them choose whether to use a gun rationally, considering the benefits (improvement in outcomes) and costs (increase in sanctions). But x is predetermined once the robbers enter the bank, and once x is given the bank robber has to choose t^*.

Assuming that i) criminals have expectations that are group-specific, ii) that there are no unobservable abilities or unobservable incidents, $u_i = 0$ ($i = h, r$), the estimates of $\lambda(t^*|x)$ and $E(U'[Y(t^*, x)])$ measure the *perceived* hazard rate and the *perceived* marginal utility of individuals with characteristics x. Estimates of $\lambda(t^*|x)$, $E[U'[Y(t^*, x)]]$ and $E[U[Y(t^*, x)]]$ provide the distribution of $D(x, w)$, the disutility of jail, a decisive component of criminal behavior. In this setup, based on Equation 5, the characteristics of the robbers and the banks they decide to target signal their ability. And these signals should be used to set sentence enhancements that target specific robbers, as prescribed in Italy's art. 628 of the penal code. The precision of the estimates depends on how precisely bank managers measure the duration of bank robberies t, and Section 4.3.4 deals with this issue.

What if criminals' individual expectations differ from the conditional ones, the "grouped" ones, or $u_i \neq 0$ ($i = h, r$)? In this case one solution would be to use repeated robberies of individual robbers. Under the strong assumption of no learning across robberies (fixed unobserved ability and no unobserved incidents), one could difference out u. In order to identify robbers one would have to use data from surveillance cameras, which are not yet available because judicial data would introduce selection bias: later it will be shown that less able robbers are more likely to be arrested. In Section 4.4 I show that under some parametric assumptions one can estimate the joint distribution of the two unob-

served abilities, use simulation methods to reproduce the unobserved heterogeneity, and thus determine how robust the findings are when we depart from the assumption of no unobserved ability $u_i = 0$ $(i = h, r)$. The main intuition is that unobserved heterogeneity in the expected marginal haul is like a random coefficient on the duration of the bank robbery. Forcing the coefficients to be constant moves the the random coefficient into the error term. But unlike the idiosyncratic part of the error term this additional part is heteroscedastic and depends on time, which identifies its variance.

The other way to control for individual heterogeneity is to base the estimates on actual realizations. These will contain the individual specific ability, but also incidents that are unobserved to both the econometrician and the robbers, leading to overly noisy estimates.

3.2 Semi-parametric Estimates

In the previous section the assumption was that individual expectations that depend on ability can be approximated expressing the haul (or the utility) and the risk of apprehension as a parametric function of the *modus operandi*. A robber knows his marginal benefits and marginal costs of staying an additional minute inside the bank even before entering the bank, and does't get any additional information about those margins while in action. While it might be difficult to continuously re-optimize based on what happens during a robbery, and while it might take time to interpret the signals one gets during a robbery, the assumption that robbers stick to their initial choice might be too strong.

Over time the robber might get some new information about the speed at which he is collecting the haul. A way to measure the marginal haul and the total haul, taking this additional information into account, is to use the actual realization Y to compute the expected marginal haul $E[Y'(t, a)]$ and the expected total haul $E[Y(t, a)]$. Unlike for the haul, there are no clear signals that change risk perceptions over time. Moreover, realized risk does not change continuously (one is either apprehended or not), meaning that one cannot use realizations to approximate perceptions. For this reason one still has to specify a parametric function to model risk.

The advantage of this semi-parametric method—"semi" because the hazard rate is still going to be parameterized—is that the actual realization is directly linked to individual ability a, addressing the concern that in the previous approach t might be correlated with unobserved ability. Since the individual expectation is equal to the realization plus the individual forecast error $E[Y(t^*, a)] = Y + \xi$, the realization is a good approximation of the expectations, as long as the robber's forecast error is small. In the parametric approach these unexpected incidents enter the error term and are averaged out, here, instead, they

add noise to the distribution of the disutility of jail, leading to an overly heterogeneous distribution.

But these very different and less parametric estimates, shown in Section 4.4, provide a robustness check for the more parametric ones.

4 Empirical Analysis of Preferences and Strategies of Bank Robbers

Solving the model of optimal duration for the unobserved part D using a linear utility ($U = Y$), the disutility of incarceration depends on the marginal haul $E(Y'(t))$, on the average haul $E(Y(t))$, and on the hazard rate of apprehension $\lambda(t)$. The next step is to devise an empirical strategy to estimate these functions. Let me start with the average and the marginal haul.

4.1 The Average and the Marginal Haul

Figure 3 shows a locally smoothed regression with optimal bandwidth of the haul as a function of time (Cleveland, 1979). The average haul appears to be linear in time, which is consistent with the typical technology used to rob a bank: i) enter the bank and walk to the teller, ii) ask the teller for the money, typically the teller's direct cash holdings, iii) collect and store the cash. Of all these actions the last is probably the most time consuming, and is likely to produce constant marginal cash returns with respect to time.

For these reasons I will model the haul using a linear regression, clearly a good approximation of the more flexible conditional mean. Moreover, a liner regression allows me to estimate group specific marginal effects without suffering from the curse of dimensionality, typical of more non-parametric methods. Using a linear model and allowing the slope of the haul with respect to t to depend on x the vector of the *modus operandi* and of the branch characteristics, the estimating equation is:

$$y_i = \alpha + \beta'_x x_i + \beta_{t^*} t^*_i + \beta'_{t^*x} x_i t^*_i + \epsilon_i \tag{6}$$

where $x_i = (m_i \ z_i)$. Linearity also has the advantage of delivering both $\widehat{E}[Y'(t^*_i, x_i)] = \beta_{t^*} t^*_i + \beta'_{t^*x} x_i$ and $\widehat{E}[Y(t^*_i, x_i)] = \widehat{y}_i$ at once. Notice that the purpose of this equation is to provide the best approximation to the individual expectation of the haul. Given that the technology appears to be linear, it is reasonable to assume that robbers have expectations

about the haul that change linearly with the time as well. But the coefficients on the regressors cannot be given any causal interpretation. Indeed, as outlined in the model, they are supposed to capture as much selection on ability as possible. As specified before, an alternative way is to measure the expectations using actual realizations (for successful robbers), $\widehat{E}[Y(t_i^*)] = y_i$ and $\widehat{E}[Y'(t_i^*)] = y_i/t_i^*$, where the last equation assumes again a linear technology.

Table 5 presents the estimates of Eq. 6. Column 1 shows that when I do not control for any other characteristics of either the bank or the bank robbery, each additional minute spent robbing a bank is associated with a 1,000 euro increase in the haul but, as we will see later in the hazard models, it is also associated with an increases in the probability of apprehension. In column 2 I allow the conditional mean $E(Y(t))$ but not the marginal effect $E(Y'(t))$ to depend on the characteristics of the bank robbery. Using firearms is associated with larger hauls (+4,000 euro), and so is being disguised (+1,400 euro), which is probably a signal of ability and professionalism. Operating in groups, instead, seems to be associated with lower the per-capita hauls. In column 3, the same variables that turned out to be significant, are interacted with the duration of the bank robbery, allowing for differential slopes.

Bank robberies in South and Central Italy have average hauls that are approximately 1,500 euro larger than in the North. In isolated banks, in smaller banks, that is banks with less than 5 employees, and in banks with lower amounts of cash hauls tend to be smaller. Security devices seem to pay off. Security devices are negatively correlated with the haul (-248 euro), and so are the average number of characteristic that these security devices have (-2,500 euro). A higher fraction of invisible security devices is also associated with smaller expected hauls. Banks that are guarded are subject to lower hauls, but the difference is not significantly different from zero. The slope does depend significantly on firearms (+1,800 euro), on the average number of characteristics per security device (-1400 euro), and small cash holdings (-900 euro).

Even with the interactions all slopes stay positive, though having a firearm seems to be the only variable that is associated with large increases. Columns 4 and 5 replicate columns 2 and 3 using the sample of robberies that last up to 30 minutes. Given that durations that are larger than 10 are more subject to measurement error from heaping (see Figure 2), it is not surprising that the coefficient on duration becomes smaller. Another difference is that adding robberies that last longer the effect of firearms interacted with duration becomes smaller while its direct effect becomes larger. All other coefficients change very little.

Columns 3 and 5 show the specification that I use to predict the haul and the marginal haul per minute. Notice that column 1 shows an inconsistency.[7] Setting all durations equal to 0, the regression in column 1 predicts an average haul of around 4,400 euro (the constant term), instead of zero, which is what the model of optimal duration would predict. But this is not the case when the complete model in column 3 is used, suggesting that the estimate in column 1 suffers from omitted variable bias: unconditionally, more able robbers take less time to rob a bank. Setting all the durations to zero in the richer model, that better controls for ability, predicts an *average* haul of -305 euro, which is not far from zero.

Given that the estimated disutility of jail is a function of $\widehat{\beta}_{t^*}$ and $\widehat{\beta}_{t^*w}$ together with $\widehat{\lambda}(t)$, increasing the number of interactions is going to increase the set of estimated disutilities. In contrast, if all robbers shared the same expected haul, expected marginal haul, and the same hazard function there would be just one estimated disutility of jail.

4.2 The Hazard Rate of Arrest

But the benefits are only part of the story. Criminals are sometimes arrested, and might serve prison time. Figure 4 shows the estimated unconditional hazard rate, $\widehat{\lambda}(t)$, using the exponential hazard model and Cox's proportional hazard model, using the 9 minutes (left panel) and the 30 minutes (right panel) sample.[8] The reason I focus on the exponential model is to avoid "aiming at a moving target." If the estimated hazard rates differ across time, due to selection on ability, it is impossible to pinpoint the criminals' expected marginal cost for each additional minute spent inside the bank. Cox's estimates are subject to selection, but are shown to compare with the more constrained estimates based on the exponential model. While I focus on the exponential model and the non-parametric Cox model, all hazard models lead to similar results.

Table 6 shows how the same regressors that I used for $E(Y(t))$ influence $\lambda(t)$ based on both the exponential and Cox's proportional hazard model. In the Cox model the coefficients do not depend on the baseline hazard, but the results are quite similar when a constant baseline hazard is used. In the first column of each model I control for the characteristics of the robbery, while the last columns additionally control for the characteristics of the bank. Focusing on the comprehensive regression, criminals who use firearms are less likely to get arrested, and so are robbers who work in groups. This is probably be-

[7] I thank an anonymous referee for pointing this out.

[8] In Table 3 and in the hazard models robberies that end without an arrest are treated as censored. Notice that the purpose is again to get the best predictor of the hazard rate and not to infer causality.

cause robbers who work in groups are likely to monitor the streets and realize possible dangers. As before, some of the effects might be driven by selection. For example, the number of security devices has a puzzling negative effect. This is probably because more able robbers are more likely to target more "challenging" banks, but are also more likely to be successful. The geographic region does not influence the hazard, while smaller and more isolated banks tend to be less risky. Conditional on the other covariates whether the bank has a guard or not does not seem to matter. Results based on the 30 minute sample are very similar to the 9 minute one (Columns 5 to 8).

An exponential hazard model with random unobserved heterogeneity component that is distributed gamma is rejected by the data in favor of no heterogeneity. This means that allowing for random effects leaves the coefficients on the *modus operandi* basically unchanged. Nevertheless, the unobserved component might still influence the disutility of jail and later in Section 4.4 I allow the disutility to contain heterogeneity in ability with respect to, both, risk and haul.

4.3 Estimating the Disutility of Jail

Based on Eq. 3 estimates of $E(Y(t,x))$, $E(Y'(t,x))$, and $\lambda(t,x)$ determine the disutility of jail, D. In particular, using Equation 3 and the exact functional form used to estimate costs and benefits (including Equation 6):

$$\begin{aligned}\widehat{D}(t_i^*, x_i) &= \frac{1}{e^{\widehat{\mu}+\widehat{\gamma}_h' x_i}}(\widehat{\beta}_{t^*} + \widehat{\beta}_{t^*x}' x_{1i}) - (\widehat{\alpha} + \widehat{\beta}_x' x_i + \widehat{\beta}_{t^*} t_i^* + \widehat{\beta}_{t^*x}' x_{1i} t_i^*) \\ &= \frac{1}{e^{\widehat{\mu}+\widehat{\gamma}_h' x_i}}\widehat{\delta}_i - (\widehat{\alpha} + \widehat{\beta}_x' x_i + \widehat{\delta}_i t^*), \qquad (7)\end{aligned}$$

where x_{1i} is the subset of regressors that are interacted with the duration of the robbery to estimate the haul and δ_i represents the individual marginal haul. Whenever the Cox hazard model is used the hazard rate depends also on t^*.

Alternatively, when the actual realizations of the haul are used to approximate the criminals' expectations the disutilities are equal to:

$$\widetilde{D}(t_i^*, x_i, y_i) = \frac{1}{e^{\widehat{\mu}+\widehat{\gamma}_h' x_i}} \frac{y_i}{t_i^*} - y_i, . \qquad (8)$$

The realized haul is used to measure the expected one, and the average haul per minute is used to measure the expected marginal one. The robber's forecast errors might introduce additional noise, especially if t_i^* is small and is measured with some error (since we are

dividing by it). It is plausible to think that the robbers enters the bank with some expectations that are based on his prior experience and that only as time passes he learns whether his day has been a lucky one or not. In such a learning model the predictions and thus the estimated disutilities lie in between the parametric and semi-parametric ones. At the beginning of the robbery criminals' expectations depend on their initial prior (based on the parametric model), but as time passes they update their expectations based on the amount of haul they collect. Assuming that the predicted parametric estimate represents the robbers' prediction for the first minute, and is later treated as if it was an initial observation, as time passes, more and more weight (in proportion to time) is added to the actual realization. With this, so called "least squares learning algorithm,[9]" the disutility becomes

$$\overline{D}(t_i^*, x_i, y_i) = \frac{1}{t_i^*}\widehat{D}(t_i^*, x_i) + \left(1 - \frac{1}{t_i^*}\right)\widetilde{D}(t_i^*, x_i, y_i), , \tag{9}$$

meaning that the semi-parametric estimate receives a weight of 1/2 after the first minute, 2/3 after the second and so on. The estimates of $\widetilde{D}$ and $\overline{D}$ are later shown in Section 4.4 which is fully devoted to the issue of unobserved ability. In that Section I also show that under some reasonable assumptions, one can estimate the distribution of unobserved ability and adjust the disutilities accordingly.

Before displaying the distributions of D, notice that this *total* disutility is going to depend on the number of years robbers expect to spend in jail when arrested. It is difficult to tell how criminals discount jail time, and the data do not allow me to estimate such a function. The simplest scenario is that robbers do not discount future jail time at all, and that the total disutility D is simply equal to the yearly times the expected years of jail time, $D = d \times S$. If robbers discount their future disutility of jail at rate δ, then $D = \sum_{t=0}^{S-1} \delta^t d = d\frac{1-\delta^S}{1-\delta}$

In Italy there are no official statistics on prison time served by convicted bank robbers. In order to compute d, the "yearly" disutility of jail, I collected data on sentences related to bank robberies. The data refer to 96 bank robbers convicted in the Piedmont region. Because of lack of information about the targeted branches the corresponding 323 bank robberies, committed between 1993 and 2007, cannot be linked to the robberies outlined in the previous data.

[9]These learning algorithms have been shown to be good approximations of more complicated Bayesian learning algorithms (Cogley and Sargent, 2008).

4.3.1 The Expected Sentence Length

Table 7 shows the summary statistics for the sample of 323 bank robberies attributed to 96 different bank robbers who were sentenced to jail in the Piedmont region, located in Northern Italy, between 2005 and 2007. This means that in our sample each robber has been convicted based on an average of 3.4 bank robberies. The bank robbers are on average 35 years old, most are Italian (92 percent), and despite the convictions coming from a Northern region, 35 percent were born in the south of Italy. 67 percent of the robbers are recidivists and 34 percent plea bargain. The other variables vary by robbery. In 22.5 percent of the cases robbers use firearms (versus 13.7 percent from the Italian Banking Association data), in 57.2 percent they wear a mask (versus 42.7 percent) and in 68.9 percent they work in teams (versus 66.3 percent). 4 percent of the time the robber uses hostages. The average total haul is 12,374 euro, slightly lower than the total haul based on the banking data. While the *modus operandi* of robbers that were sentenced in Piedmont are on average slightly different than in the country-wide robbery data of the Italian Banking Association, the criminal law and, thus, the determinants of the sentence length should be the same across the country.

The average sentence length is 3.4 years in prison. Data on sentence durations allows me to model the log-sentence length based on the same *modus operandi* variables observed for the bank robberies and to impute the variation in the log-Disutility of jail, D, that is driven by the variation in the sentence length, S, $\log(D) = \log(d) + \log(S)$. Thus $\log(D) - \log(S) = \log(d)$ represents the log-Disutility for each year in jail.[10]

In order to determine the way the *modus operandi* shapes the expected sentence length, I estimate the log-sentence length on whether the robber used firearms, was masked, or worked in a group. Estimates are shown in Table 8. Based on Column 1, using a firearm increases the sentence by approximately 50 percent (by less once I control for recidivism, hostages, plead bargain, year, total number of robberies committed, total haul). Wearing a mask and working in groups has a smaller effect on the sentence. Working in groups increases the sentence length by 20-25 percent, and being masked by less than 10 percent but without being statistically different from zero. Only the use of firearms leads to strong and significant sentence enhancements. This is likely to explain why so many robbers choose to work in groups and to wear a mask, while so few use a firearm.

[10]Notice that I used a discount factor $\delta = 1$, otherwise $log(D) = \log(d) + \log(\frac{1-\delta^S}{1-\delta})$.

4.3.2 The Total and the Yearly Disutility of Jail

Figure 5 shows for those criminals who were not arrested, and whose choice of t was more likely to be the optimally chosen one, the distribution of the parametric total disutility of jail between the 5th and the 95th percentile. The yearly figures are estimated dividing the total disutility by the predicted sentence length based on the regression shown in Column 3 of Table 8. The same distributions but for the sample of robberies that last up to 30 minutes is shown in Figure 6. An interesting feature of all the estimated distributions is the shape, which resembles an earnings distribution. Since the value of staying out of jail is likely to depend on the robbers' earnings potential, it follows that these earnings are distributed like legitimate earnings. It is worth stressing that nothing in the model prevents the shape of the distribution from taking any other form or generating negative "values of freedom." Indeed, for 10 percent of the robberies the model predicts negative disutilities of ending up in jail. This is entirely driven by those criminals who rob banks with a large average number of characteristics per security device. These criminals have such small marginal hauls that the disutility ends up being negative. These negative values of freedom can clearly be driven by factors that the model does not control for, like unobserved heterogeneity in expectations, or heterogeneity in risk aversion. Robbers that targeted banks with a large average number of characteristics per security device might have been unable to predict such small marginal hauls. For the 30 minute sample the coefficient on the average number of characteristics per security device is considerably closer to zero and only a very small fraction of disutilities are negative. Later, in Section 4.4, I allow the disutilities to depend on unobserved heterogeneity in expectations about risks and benefits, while in Section 4.5 I allow the utility to be different from linear.

The kernel densities show that dividing the total disutility by the expected sentence length reduces the heterogeneity in disutility. According to Table 9, not only the variance but also the coefficient of variation gets smaller when controlling for the expected sentence length. Since expected sentences are likely to be measured with some noise (it is hard to know what robbers really expect the jail sentence to be), the true expected sentences might explain an even larger share of the variation. Table 9 also shows that the distribution is highly right-skewed. As a consequence, the median is small compared to the mean: 44,000 against 71,000 euro for the exponential model with the cutoff at 9 minutes, and 37,000 against 49,000 with the cutoff at 30 minutes. The corresponding figures for the yearly disutility are 15,000 and 20,000 euro (12,500 and 14,500 euro for the 30 minutes sample). These figures are implicitly assuming that robbers do not discount time. If they did, the yearly figures would be relatively larger by $\log S - \log(\frac{1-\delta^S}{1-\delta})$, which for an average

sentence of 3.4 years, would be less than a quarter with a discount factor of 80 percent, and around 12 percent with a discount factor of 90 percent. Unobserved heterogeneity in discount rates might thus drive some of the heterogeneity in D.

4.3.3 Disutility of Jail: Ability vs. Deterrence

Robbers with different values of freedom target different banks, and use different *modus operandi*. In order to describe this selection, I compute the derivative of the disutility with respect to the same variables that are related to the haul and to the risk of arrest. Given that D differs across individuals, so will its derivatives.

Table 10 shows the derivative of $\log D(t, x)$ with respect to duration t, and *modus operandi* x, which includes the bank characteristics. The Table shows the average derivative, its standard deviation, and the 5th and 95th percentile.

The sanctioning rules (judges adjust sentences proportionally to the aggravation of the robbery) suggest to use the log value of freedom instead of the level.[11] The observable characteristics of banks and bank robberies change the (log) value of freedom the way one would expect, given the sanctioning rules set by the penal code. Art 628 of the penal code sanctions masked robberies, robberies perpetrated by more than one criminal, and robberies where firearms are used more than "simple" robberies (*rapina semplice*). These deterrence effects are coherent with the sign of the changes shown in Table 10. The use of firearms leads to an *increase* in the disutility of jail of about 178 percent (the standard deviation of the change is 63 percent). Using masks and operating in a group also leads to a sizable and usually significant *increase* in disutility (65 to 80 percent). All these changes are larger than zero for the top 95 percent of the distribution. But these increases are considerably larger than the increase in the sentence length that one predicts based on judiciary level data (Table 8), suggesting that criminals that use firearms, work in groups, and mask themselves not only take longer sentences into account, thus increasing the *total* disutility, but are also of higher ability. The heterogeneity in ability is clearly visible when I derive the disutility with respect to variables that do not influence the sentence length.

Not surprisingly, robbers who operate against banks with little cash holdings are of substantially lower ability. Those that choose banks with less than 5 employees tend to be of higher ability, mainly because robberies in smaller banks are clearly less risky. Bank employees need to be monitored for the duration of the robbery, therefore the greater the number of employees the riskier the robbery becomes. Security devices, instead, generate an ambiguous selection. While only the more able criminals select banks with

[11] Using the disutility of jail in levels gives very similar results.

more security devices, the same is not true for the average number of characteristics. The fraction of visible devices and whether the bank has a guard or not do not significantly alter the selection of criminals. But not all of these results are significant when using the standard deviation of the changes to evaluate the significance. The model based on the Cox proportional model gives very similar results.

The last four columns of Table 10 replicate the changes for the sample of robberies that last up to 30 minutes. Consistent with the differences shown in the average haul regressions, the use of firearms leads to smaller changes (-98 percent). All other changes tend to be similar.

4.3.4 How Does Measurement Error in Duration Change the Distribution of D?

While measurement error in the duration is going to have no effect on a constant baseline hazard estimate, it will bias the estimated marginal haul downwards. Some simulations that I performed show that while rounding a duration, measured in seconds, to the nearest or to the smallest minute has almost no effect on the coefficient of duration (chosen to have the same level of significance as in the actual data), rounding the duration randomly to one of the two nearest minutes induces a larger bias (-13 percent). The largest bias (-16 percent) arises when 10 percent of the durations are randomly set to be equal to 5 minutes. The relative bias of size m is going to induce a change in D that is equal to:

$$m\frac{\partial D_i}{\partial \log \beta_i} = m\frac{\beta_i(1 - t_i^*\lambda(t_i^*, x_i))}{\lambda(t_i^*, x_i)} \tag{10}$$

where $\beta_i = \beta_{t^*}t_i^* + \beta'_{t^*x}x_i t_i^*$ represents the individual slope with respect to t. Since $1 - t_i^*\lambda(t_i^*, x_i)$ is generally positive, the bias reduces the estimated D. This can clearly be seen in Figure 7, where I plot the density of D, assuming three different biases: a 5, 10, and 20 percent attenuation bias of the slope. Dealing with the attenuation bias reduces the fraction of negative disutilities of jail from 10.3 to 8.8 percent. The median and the mean are clearly more sensitive to the measurement error. A 10 percent correction almost doubles the median (from 44,000 to 77,000 euro) and the mean (from 71,000 to 133,000 euro). Adding another 10 percent correction increases the median and the mean by a relatively smaller amounts (77,000 to 108,000 euro and 133,000 to 195,000 euro). Similar changes apply to the larger sample of robberies that last up to 30 minutes. It is interesting to note that the measurement error reduces the differences between the distributions estimated using the cutoffs of 9 and 30 minutes. Since the sensitivity to the

measurement error is quite high, the policy simulations of the following subsections are performed allowing again for different degrees of measurement error.

4.3.5 Deterrence and Heterogeneity in Deterrence

The model allows me to answer the following question: How much would policy makers have to increase the disutility of jail (the sentences) to drive the number of bank robberies to zero? In terms of the model, one needs to determine the level of disutility that corresponds to an optimal duration that is equal to zero:

$$D(0,x) = \frac{1}{\lambda(0|x)} E(Y'|0,x) - E(Y|0,x), \tag{11}$$

or, using Equation 7

$$D(0,x) = \frac{1}{e^{\mu+\gamma_h' x_i}} (\beta_{t^*} + \beta_{t^* x}' x_{1i}) - (\alpha + \beta_x' x_i), \tag{12}$$

$\log D(0,x) - \log D(t^*,x)$ represents the percentage increase in disutility needed for robbers that use a *modus operandi* x, in t^* minutes to drive the duration to 0.[12]

Given that the only way to increase the disutility of jail is by increasing the expected jail time, deterrence is going to depend on the robbers' discount factors. What matters is how an increase in (relative) jail time translates into an increase in (relative) disutility. If robbers do not discount future jail time than $\frac{\partial \log(D)}{\partial \log(S)} = 1$, while if they do $\frac{\partial \log(D)}{\partial \log(S)} = \frac{-S\delta^S}{1-\delta^S} \log \delta$. The difference between these two elasticities measures how discounting attenuates deterrence. With a discount factor of 80 percent attenuation at the average sentence of 3.4 years is equal to 34 percent, while with a discount factor of 90 percent the attenuation is equal to 17 percent.

With this in mind, Table 12 shows the distribution of the changes, assuming no discounting. Unlike the distribution of the disutility of jail these elasticities are less sensitive to measurement error. The 5th percentile shows that without correcting for measurement error, in order to drive 5 percent of the sample to a duration of zero one needs a 3 percent increase in the total disutility of jail, or equivalently the same increase in sentence length. Controlling for measurement error, the change in penalty needed is almost unchanged. In order to reduce the bank robberies by a quarter, the penalties would have to increase by between 6 and 9 percent, depending on the degree of the bias. To curb robberies by one-half, penalties would have to increase by between 11 and 17 percent. In order to

[12] When the exponential model is used the hazard rate does not depend on t and $\lambda(0|x) = \lambda(x)$.

almost eliminate bank robberies (-95 percent), the sanctions would have to increase by 77 percent in the absence of measurement error and by 48 percent if the measurement bias was equal to 10 percent. Overall, the estimated model predicts criminal behavior to be highly responsive to changes in the sanctioning system, though discounting would attenuate this responsiveness according to the function discussed in the previous paragraph. Notice also, that given the assumption of risk neutrality, robbers would be equally responsive to changes in the likelihood of arrest. The lower panel of Table 12 shows that the responsiveness is even larger when looking at the sample of robberies that last up to 30 minutes.

The data allow us to go even further and explore which robbers are more likely to respond to an increase in sanctions. Table 13 shows the mean for the *modus operandi* variables and for the variables describing the banks for values above and below the median percentage increase in disutility needed for robbers to drive the duration to 0. Values below the median signal high responsiveness to sanctions (the corresponding average log change in disutility is 9 percent), values above the median signal low responsiveness to sanctions (the corresponding average log change in disutility is 44 percent). A pretty clear picture emerges from the table, for both samples. Robbers with higher disutilities of jail (133,000 versus 31,000 euro), more likely professional robbers, are also more responsive to sanctions. In particular, essentially all robbers that use firearms belong to the high responsiveness category. Masked robbers are also considerably more likely to be highly responsive (64 versus 27 percent). This means that harshening sanctions would mostly deter those robbers that are responsible for the largest losses. The amateur robbers would most likely keep on trying to rob banks. It is worth noticing that in the US, where sanctions are definitely more severe, bank robberies are believed to be mostly the work of amateurs (Weisel, 2007, Department of Justice, 2003). A signal of amateurism is that in the US 80 percent of the incidents involved only one offender (Department of Justice, 2003), against just one third in Italy.

Only in terms of disguises and the use of firearms do US robbers appear equally or even more professional. About 40 percent of US bank robbers wear masks (Weisel, 2007), the same as in Italy. And the use of firearms in holdups is clearly more widespread in the US than in Italy, 30 percent against 14 percent, though this might also be driven by their relative abundance of weapons on the market.

4.4 Unobserved heterogeneity

All the analyses have relied on the assumption that the estimated $\lambda(t^*|x)$, $E[Y'(t^*, x)]$ and $E[Y(t^*, x)]$ capture the robbers perceived cost and benefits. But the regressors x might not capture the entire ability of robbers. As mentioned earlier, in the absence of repeated observations one can i) use simulation methods to assess the importance of unobserved ability, or ii) use actual realizations (which contain all the individual ability plus some noise). In the first method I feed the estimates of $\lambda(t^*|x)$, $E[Y'(t^*, x)]$ and $E[Y(t^*, x)]$ with the unobserved components of ability, the errors u_i $(i = h, r)$.

Let us start with the parametric estimates by explicitly allowing Equation 7 to contain individual heterogeneity. A reasonable and general way to introduce unobserved heterogeneity is by allowing the hazard rate and the marginal haul to vary across individuals, even after conditioning on x. This is like saying that everything else equal, some individuals expect to gather more money per minute and to risk less for each minute spent inside the bank, or

$$\widehat{D}(t_i^*, x_i,) = \frac{1}{e^{\mu_i + \gamma' x_i}}(\beta_{i,t^*} + \beta'_{i,t^*x} x_{1i}) - (\alpha + \beta'_x x_i + \beta_{i,t^*} t_i^* + \beta'_{i,t^*x} x_{1i} t_i^*), \tag{13}$$

Notice that μ and the βs that measure the marginal haul are now indexed i. Rewriting the equation in terms of the initial μ and β and assuming that the individual heterogeneity with respect to β is constant across the different x's

$$\widehat{D}(t_i^*, x_i,) = \frac{1}{\varphi_i e^{\mu + \gamma' x_i}} \phi_i (\beta_{t^*} + \beta'_{t^*x} x_{1i}) - (\alpha + \beta'_x x_i + \phi_i \beta_{t^*} t_i^* + \phi_i \beta'_{t^*x} x_{1i} t_i^*), \tag{14}$$

where $\phi_i = \frac{\beta_i}{\beta}$ and $\varphi_i = e^{\mu_i}/e^{\mu}$. ϕ_i and φ_i represent the individual marginal effects relative to the average ones. In order to simulate ϕ_i and φ_i we need to specify their joint distribution, which requires several premisses:

No correlation between the observed and the unobserved heterogeneity: If the unobserved heterogeneity is correlated with the *modus operandi* x the estimated coefficients would capture such correlation, much in the same way a coefficient that is biased because of omitted variable bias captures the correlation between the omitted variable and the dependent variable. For this reason, I assume the unobserved heterogeneity and the observed one to be uncorrelated with each other, much in the same way in a linear regression predicted values are orthogonal to predicted error terms.

Shape: Given the shape of the distribution of the disutility of jail it is reasonable to assume that the joint distribution of ϕ_i and φ_i is hump-shaped, meaning that larger

deviations from the mean are less likely than smaller ones. An obvious and convenient candidate is the bivariate log-normal distribution. The distribution restrict the individual *relative* differences with respect to the average (ϕ_i and φ_i) to be larger than 0.

Mean: Section 4.3.4 discussed the possibility that the marginal effects are systematically biased downwards. Here, instead, the marginal effects (marginal haul and the hazard rate) are assumed to be unbiased, meaning that $E(\phi_i) = E(\varphi_i) = 1$.

Variance: The best way to derive the variance of ϕ_i is by rewriting the haul regression with the individual heterogeneity (random coefficients) as

$$y_i = \alpha + \beta'_x x_i + \beta_{t^*} t_i^* + \beta'_{t^*x} x_i t_i^* + u_i \ , \tag{15}$$

where $u_i = (\phi_i - 1)(\beta_{t^*} t_i^* + \beta'_{t^*x} x_{1i} t_i^*) + \epsilon_i = (\phi_i - 1)\delta_i t_i^* + \epsilon_i$. ϵ_i is an idiosyncratic error term that does not depend on ϕ_i, while δ_i measures, as in Equation 7, the individual observed marginal haul. The error term u_i has mean zero but its variance conditional on δ_i and t_i^* is

$$E(u_i^2) = E[(\phi_i - 1)^2](\delta_i t_i^*)^2 + E(\epsilon_i)^2 \ . \tag{16}$$

By analogy one can estimate the variance of ϕ regressing the squared residuals u_i^2 on the squared of the product between the marginal haul and the duration of the robbery $(\delta_i t_i^*)^2$, or

$$u_i^2 = \sigma_\phi^2 (\delta_i t_i^*)^2 + \theta + \varepsilon_i \ . \tag{17}$$

Notice that this procedure resembles the Breusch and Pagan test for heteroskedasticity and random coefficient variation (Breusch and Pagan, 1979) with the only difference that the evidence of heterogeneity is used to estimate the variance of its unobserved component $\sigma_\phi^2 = E[(\phi_i - 1)^2]$.

Running a quantile regression that is robust to outliers in the squared residuals of the haul the estimated variance is equal to 0.46, with a t-statistic of 50. As for the variance of φ, despite Section 4.2's rejection of a random component to the hazard in one specification I allow φ_i to be random, different from ϕ, but with the same variance.

Correlation: ϕ and φ are likely to be correlated with each other. One would expect the correlation between φ and ϕ to be negative if the abilities are positively correlated with each other.

In order to draw ϕ_i and φ_i from a joint distribution one needs to know that correlation.

While one cannot directly measure the correlation between the unobserved components, the correlation between the observable ones ($\widehat{\lambda}(t_i^*|x_i)$ and $\widehat{\delta_i}$) is easily available. The correlation is equal to -20.64 percent for the "9 minutes" sample and -20.37 percent for the "30 minutes" one. Notice that sharing the same regressors does not mean that the two functions need to be "mechanically" correlated in a particular way. The low correlation means that minimizing risk sometimes goes against maximizing the money. There can be trade-offs, like choosing to be partnered by someone that secures the escape on the outside of the bank: the partner lowers the risk of the robbery but pretends a cut, lowering the per-capita haul. Unless unobserved signals of ability involve very different trade-offs than the observed ones, one can use the correlation of -20 percent to draw the unobserved abilities from the joint distribution of ϕ_i and φ_i.

For the sake of brevity I perform the following robustness checks on just the robberies that last up to 9 minutes, but, similar to what happened to all previous analyses, all the following findings are robust to extending the sample to robberies that last up to 30 minutes. Figure 8 shows the distribution of ϕ (the one of φ is essentially the same). The estimated variance of 0.46 translates into considerable unobserved heterogeneity: ϕs close to 0 and larger than 2 are quite likely. The upper panel in Figure 9 shows the distribution of parametric estimate of D with partial (left) and full unobserved heterogeneity (right). The relative heterogeneity in the hazard rates, even if it is centered around one, tends to move the distribution to the right. The estimates of D, based on actual realization, are shown in the lower left panel. A clear advantage of using actual realizations is that they predict only positive disutilities. But they tend to be larger than the parametric ones, and with heavy right tails. The reason is that Equation 8 predicts very large marginal hauls every time a short duration is linked to a large haul. Such outliers were in part averaged out in the parametric estimates. Since robbers are likely to learn their productivity over time the lower right panel shows an estimate of D based on the learning model shown in Equation 9.

Table 14 shows several statistics of the distribution of disutility with and without unobserved heterogeneity. Allowing for just heterogeneity in ϕ and not in φ keeps the average disutility almost unchanged but increases the standard deviation by almost 25 percent. Adding the heterogeneity in φ moves considerable mass to the tails of the distribution increasing the mean by 50 percent and the standard deviation by 260 percent! The reason is that whenever φ is close to zero the model predicts very large disutilities to make up for the relatively short duration of the robberies. The median, instead, is more robust. The semi-parametric estimates suffer from a similar problem. The average

disutility is equal to 170,000 euro but the median is only 88,000 euro. The learning model predicts disutilities that are similar to the one obtained using full parametric estimates with full heterogeneity. But these similarities end when one looks at the responsiveness to changes in D that are shown in the lower part of Table 14.

The additional parametric heterogeneity in D lowers the estimated responsiveness to changes in D, especially at the upper part of the distribution of "unresponsiveness." For example, while to reduce the number of robberies by 50 percent one needs a 17 percent increase in D when $\phi = \varphi = 1$ (no unobserved ability), a 21 percent increase when $\phi \neq 1$ and $\varphi = 1$, and again a 16 percent increase when $\phi \neq 1$ and $\varphi \neq 1$, reductions of 90 percent demand increases by 77, 127, and 130 percent.[13]

In the semi-parametric estimates the marginal haul is inversely related to the duration t^*. This increases the robbers estimated responsiveness to changes in disutilities, especially for reductions that are above 50 percent (P50). The estimates tend to be too low. For example, a 6 percent increase in disutility is predicted to lead to a 50 percent reduction in robberies (9 percent in the learning model). But while the estimated disutilities and the estimated responsiveness tend to differ based on whether one uses the parametric or the semi-parametric estimates, Table 15 shows that the robbers' profile by how they responsive they are to incentives are the same as before. The most "responsive" robbers are still those that have higher disutilities of jail o matter how one models unobserved heterogeneity. And these robbers are more likely to use firearms, and more likely to work in a group and disguise their appearance.

On the one hand, the shape of the distribution of disutility and the profile of robbers that tend to be more responsive to sanctions is reasonably robust to the inclusion of parametric or semi-parametric unobserved heterogeneity in expectations and ability. On the other hand, the average disutility and the changes in disutility needed to reduce the number of robberies tend to differ depending on whether one uses the parametric or the semi-parametric method.

4.5 Risk Averse, Risk Neutral, or Risk Lover?

Up until now I assumed risk neutrality. There is some evidence using experimental data that criminals that are in prison are indeed at most risk preferrers Block and Gerety (1995). Allowing for a constant relative risk parameter, using an isoelastic utility function, is straightforward if one assumes a given level of initial wealth. I assume an initial level

[13]Notice that since responsiveness can only be estimated for positive disutilities some differences are driven by the different sample selections shown in the upper part of Column 1.

of zero wealth. This is like saying that criminals rob banks every time they really need the money. Instead of using the haul Y_i as the dependent variable I simply assume a constant relative risk aversion and use $Y_i^{1-r}/(1-r)$, where r is the coefficient of relative risk aversion. Table 16 shows the corresponding regression functions. The relative size of most coefficients is preserved.

As before, the predicted values represent the expected utilities, while the coefficients on duration represent the marginal utilities. Using the exponential hazard estimates from before one can compute the disutility of jail for the different risk preferences. The distributions are shown in Figure 10. One striking feature is that log utility, that corresponds to an isoelastic utility with risk aversion parameter of 1, generates a distribution of disutilies with 80 percent of values being negative (see the first column of Table 17). Notice that adding some initial wealth to the utility would only increase the marginal cost of arrest and thus increase the fraction of negative disutilities. The distribution of durations with large risk aversion is clearly rejected by the data. A risk aversion parameter of 0.5 instead generates a distribution which is not very different from the one based on risk neutrality.[14] The data would also be consistent with risk loving preferences. Table 17 shows that the relative changes in utility needed to drive the durations to 0 are generally decreasing in the risk aversion parameter r. The only exception to this rule is the log utility, but the exception is driven by the very selected sample of individuals with positive disutilities (20 percent of the total). Overall the distribution of the changes needed to drive durations to zero do not vary by much, casting doubt on the possibility that heterogeneity in risk aversion and not in disutilities is the main driver of the difference in observed durations.

5 Conclusions

During bank robberies both the probability of apprehension and the average haul increase over time. At the margin this trade-off depends on: i) the criminal's expected haul at time t, ii) its expected increase between t and t+1; iii) the hazard rate of arrest, and iv) the criminals' disutility of ending up in jail, which in part depends on deterrence. Unique data on 5,000 Italian bank robberies – representing 57 percent of all European bank robberies, with information on the observed duration among successful robberies allow me to identify and then analyze the individual disutility of jail.

[14]Estimates that are based on experimental data are typically close close to 0.5, while those estimated using financial data are larger (Gollier, 2004, Kocherlakota, 1996, Holt and Laury, 2002, Dohmen et al., 2010).

The vast majority of criminals face relatively low disutilities of jail, while a few face very high ones. The shape of the distribution resembles the shape of an earnings distribution.

Simulating relative changes in deterrence, suggests that deterrence effects are high, and that the most responsive robbers to deterrence are the more able ones, those that have a higher disutility of ending up in jail because their opportunity cost of prison time is higher.

These criminals tend to rob banks using firearms, being disguised, and working in teams. They are also more likely to target the right banks, those with higher cash holdings but fewer employees. This differential deterrence potential, coupled with considerably harsher sanctions (the average prison sentence is 137 months in the US and just 40 months in Italy), is likely to explain why nowadays, unlike Italian bank robbers, US ones "are clearly amateurs and not bank robbery specialists." (Department of Justice, 2003). The results are robust to the inclusion of unobserved ability and heterogeneity in expectations. The hypothesis that these robbers have large risk aversion is rejected by the data.

The relatively large number of bank robberies, together with evidence of possibly large deterrence effects, suggests that in Italy prison sentences might be too low. But it appears that not the legislator, but rather the judges should reconsider their customs. Actual sentences are often below the minimum ones set by law. And despite well designed sentence enhancements that penalize *modus operandi* that are linked to high ability robbers judges, with the exception of the use of firearms, tend to neglect them.

References

David S. Abrams and Chris Rohlfs. Optimal Bail and the Value of Freedom: Evidence from the Philadelphia Bail Experiment. *Economic Inquiry*, April 2010.

Orley Ashenfelter and Michael Greenstone. Using Mandated Speed Limits to Measure the Value of a Statistical Life. *Journal of Political Economy*, 112(1):226–267, 2004. ISSN 0022-3808.

Alessandro Barbarino and Giovanni Mastrobuoni. The Incapacitation Effect of Incarceration: Evidence from Several Italian Collective Pardons. Carlo Alberto Notebooks 999, Collegio Carlo Alberto, 2008.

Terry Baumer and Michael O. Carrington. The Robbery of Financial Institutions. Executive summary, National Institute of Justice, U.S. Department of Justice, 1986.

Gary S. Becker. Crime and Punishment: An Economic Approach. *The Journal of Political Economy*, 76(2):169–217, 1968.

Michael K. Block and Vernon E. Gerety. Some Experimental Evidence on Differences between Student and Prisoner Reactions to Monetary Penalties and Risk. *The Journal of Legal Studies*, 24(1):123–138, 1995. ISSN 00472530.

Michael K. Block and John M. Heineke. A Labor Theoretic Analysis of the Criminal Choice. *The American Economic Review*, 65(3):314–325, 1975. ISSN 00028282.

Trevor S. Breusch and Adrian R. Pagan. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5):1287–94, September 1979.

Samuel Cameron. The Economics of Crime Deterrence: A Survey of Theory and Evidence. *Kyklos*, 41(2):301–323, 1988.

William S. Cleveland. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, 74(368):829–836, 1979.

T. Cogley and T.J. Sargent. Anticipated utility and rational expectations as approximations of bayesian decision making. *International Economic Review*, 49(1):185–221, 2008.

Philip J. Cook. Is robbery becoming more violent? an analysis of robbery murder trends since 1968. *The Journal of Criminal Law and Criminology*, 76(2):480–489, 1985. ISSN 00914169.

Philip J. Cook. The Relationship between Victim Resistance and Injury in Noncommercial Robbery. *The Journal of Legal Studies*, 15(2):405–416, 1986. ISSN 00472530.

Philip J. Cook. Robbery Violence. *The Journal of Criminal Law and Criminology*, 78(2): 357–376, 1987. ISSN 00914169.

Philip J. Cook. Robbery in the United States: An Analysis of Recent Trends and Patterns. *Violence: Patterns, Causes, Public Policy. New York: Harcourt Brace Jovanovich*, 1990.

Philip J. Cook. Robbery. In Michael Tonry, editor, *Handbook on Crime and Public Policy*. Oxford University Press, 2009.

Robert Crais. *The Two Minute Rule*. Pocket Star, 2007. ISBN 1416514961.

Federal Bureau of Investigation Department of Justice, editor. *Crime in the United States, 2002*, chapter Special Reports: Bank Robbery in the United States, pages 301–312. Department of Justice, Federal Bureau of Investigation, 2003.

John J. Jr. DiIulio. Help Wanted: Economists, Crime and Public Policy. *The Journal of Economic Perspectives*, 10(1):3–24, 1996. ISSN 08953309.

Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde. Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3):1238–60, June 2010.

Francesco Drago, Roberto Galbiati, and Pietro Vertova. The Deterrent Effects of Prison: Evidence from a Natural Experiment. *Journal of Political Economy*, 117(2):257–280, April 2009.

Isaac Ehrlich. Participation in illegitimate activities: A theoretical and empirical investigation. *The Journal of Political Economy*, pages 521–565, 1973. ISSN 0022-3808.

European Commission. Report on the Retail Banking Sector Inquiry. Commission staff working document, European Commission Directorate-General for Competition, Jan 2007.

Federal Bureau of Investigation. Headline Archives. Bank Robberies. Technical report, FBI, May 2007.

Richard B. Freeman. The Economics of Crime. *Handbook of Labor Economics*, 3:3529–3571, 1999.

Edward L. Glaeser and Bruce Sacerdote. Why is There More Crime in Cities? *Journal of Political Economy*, 107(S6):225–258, 1999.

Christian Gollier. *The economics of risk and time*. The MIT Press, 2004.

Jeff Grogger. Market Wages and Youth Crime. *Journal of Labor Economics*, 16(4): 756–791, 1998.

Timothy H. Hannan. Bank Robberies and Bank Security Precautions. *The Journal of Legal Studies*, 11(1):83–92, 1982.

Richard W. Harding. Rational-Choice Gun Use in Armed Robbery: The Likely Deterrent Effect on Gun Use of Mandatory Additional Imprisonment. *Criminal Law Forum*, 1 (3):427–450, 1990. ISSN 1046-8374.

Eric Helland and Alexander Tabarrok. Does Three Strikes Deter?: A Nonparametric Estimation. *Journal of Human Resources*, 42(2), 2007.

Charles A. Holt and Susan K. Laury. Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655, December 2002.

Daniel P. Kessler and Steven D. Levitt. Using Sentence Enhancements to Distinguish between Deterrence and Incapacitation. *Journal of Law & Economics*, 42(1):343–63, April 1999.

Narayana R. Kocherlakota. The equity premium: It's still a puzzle. *Journal of Economic Literature*, 34(1):pp. 42–71, 1996. ISSN 00220515.

David S. Lee and Justin McCrary. Crime, Punishment, and Myopia. NBER Working Papers 11491, National Bureau of Economic Research, Inc, July 2005.

Jude Miller-Burke, Mark Attridge, and Peter M. Fass. Impact of Traumatic Events and Organizational Response: A Study of Bank Robberies. *Journal of Occupational & Environmental Medicine*, 41(2):73–83, February 1999.

Daniel S. Nagin. Criminal Deterrence Research at the Outset of the Twenty-First Century. *Crime and Justice*, 23:1–42, 1998.

OSSIF. Rapporto sulle Spese del Settore Bancario per la Sicurezza Anticrimine nel 2006. Technical report, OSSIF, Associazione Bancaria Italiana, 2006.

Suzanne Polich, Mark A. Peterson, and Harriet B. Braiker. *Doing Crime: A Survey of California Prison Inmates*. Rand Corporation, 1980.

Richard Thaler and Sherwin Rosen. The value of saving a life: Evidence from the labor market. In *Household Production and Consumption*, NBER Chapters, pages 265–302. National Bureau of Economic Research, Inc, September 1976.

W. Kip Vicusi. *Inner-City Black Youth Unemployment*, chapter Market Incentives For Criminal Behavior. Univerisy of Chicago Press, Chicago, 1986a.

W. Kip Vicusi. The Risks and Rewards of Criminal Activity: A Comprehensive Test of Criminal Deterrence. *Journal of Labor Economics*, 4(3):317–340, July 1986b.

W Kip Viscusi. The value of risks to life and health. *Journal of Economic Literature*, 31 (4):1912–46, December 1993.

Deborah Lamm Weisel. Bank Robbery. Problem-Oriented Guides for Police, Problem-Specific Guides Series 48, Community Oriented Policing Service, U.S. Department of Justice, March 2007.

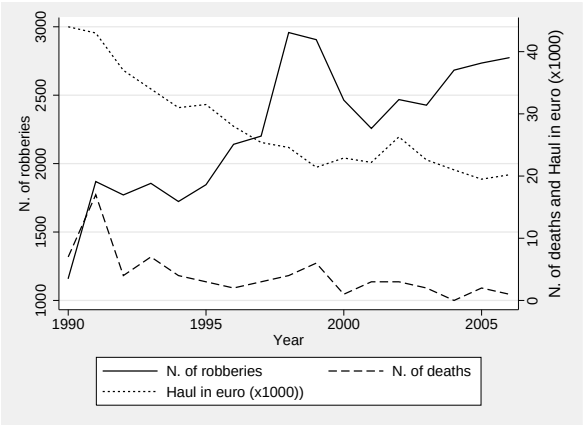


Figure 1: Number of Italian Bank Robberies, Average haul, and of the Number of Casualties

Notes: This figure shows the total number of Italian Bank Robberies (left axis), the average haul in euro (times 1,000) and of the number of casualties (both on the right axis) between 1990 and 2006.

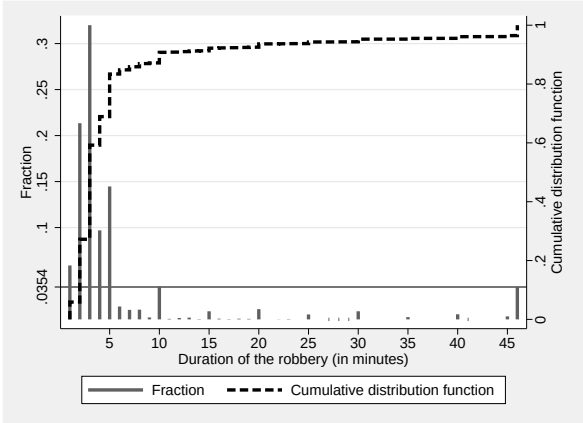
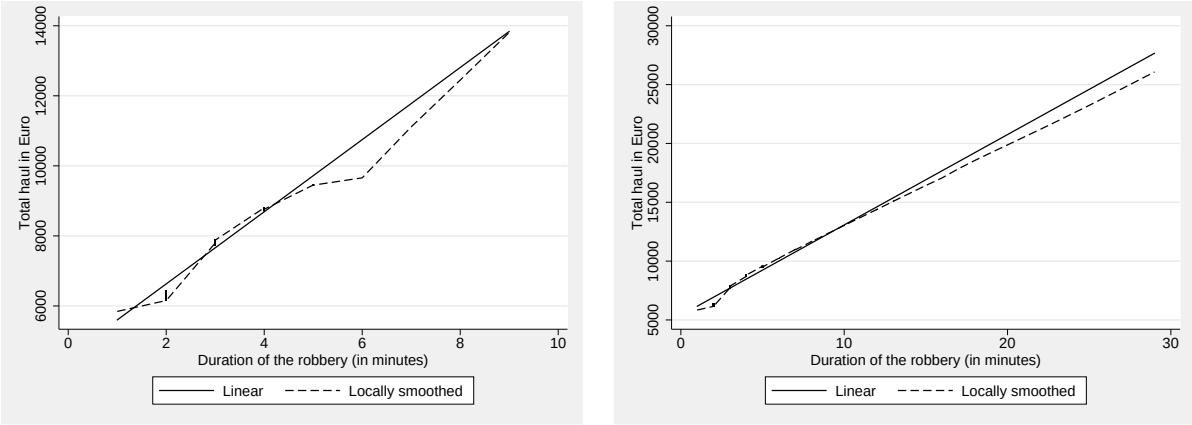


Figure 2: The Distribution of Duration t

Notes: The spikes indicate the distribution of duration (on the left axis) while the dashed line indicates its cumulative distribution (on the right axis). Minute 46 stands for all the durations that lasted more than 45 minutes.

Figure 3: The Average Haul



Notes: The solid line represents the linear regression, the dashed one a locally smoothed regression (Cleveland, 1979). The left panel shows the regression lines with duration truncated at 9 minutes, the right one truncates the duration at 30 minutes.

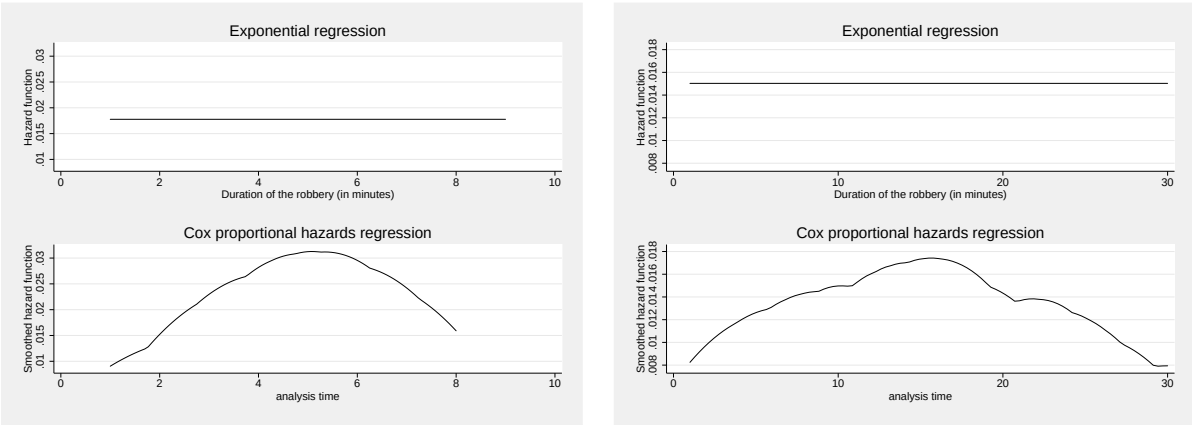


Figure 4: The Estimated Hazard Rate

Notes: The Cox proportional hazard is estimated applying an Epanechnikov kernel smooth with optimal bandwidth on the estimated increments of the cumulative hazards. The left panel shows the estimates with duration truncated at 9 minutes, the right one truncates the duration at 30 minutes.

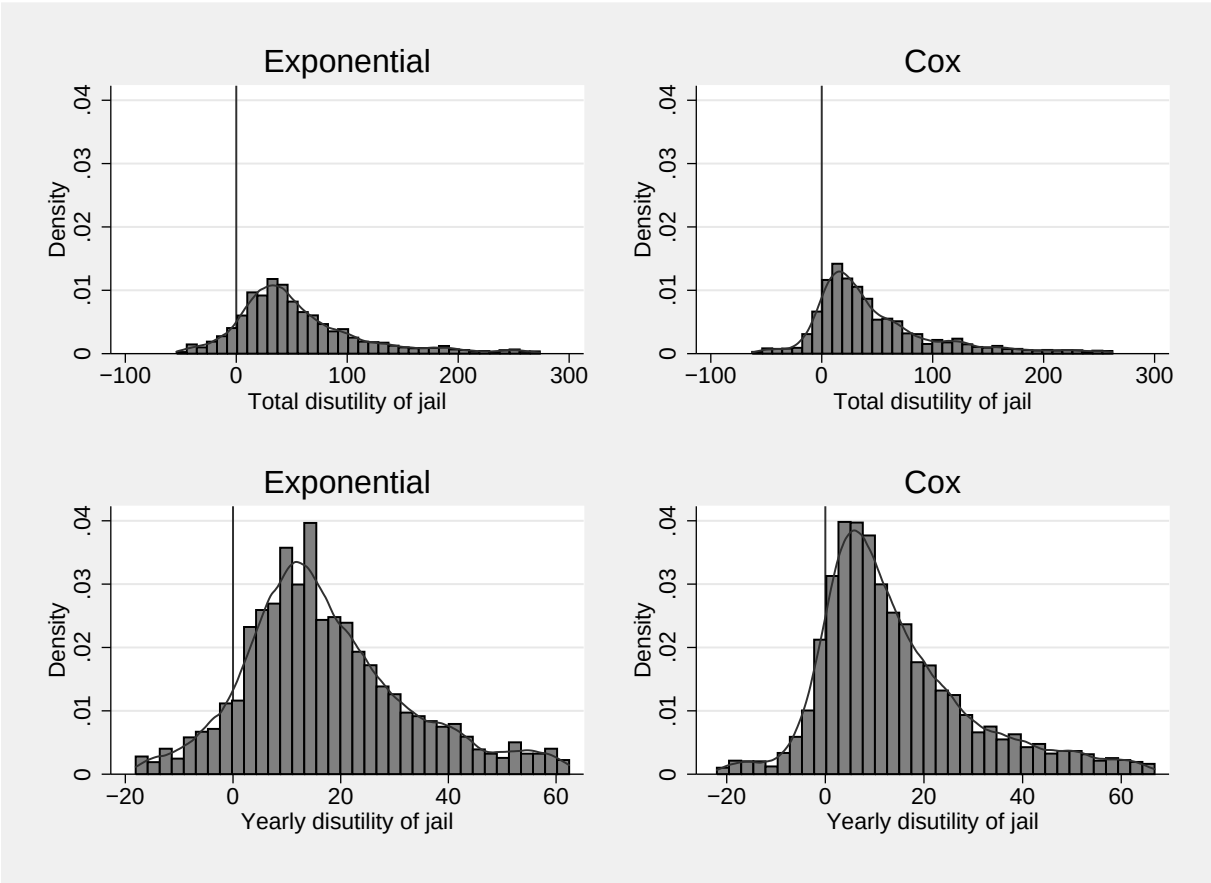


Figure 5: The Distribution of the Conditional Value of Freedom ($t \leq 9$ minutes)

Notes: The upper panels show the distribution of the total disutility of jail, the lower ones the corresponding yearly figures assuming a discount factor of one (between the first and the 95th percentile). These estimates are based on successful robberies that lasted less than 9 minutes.

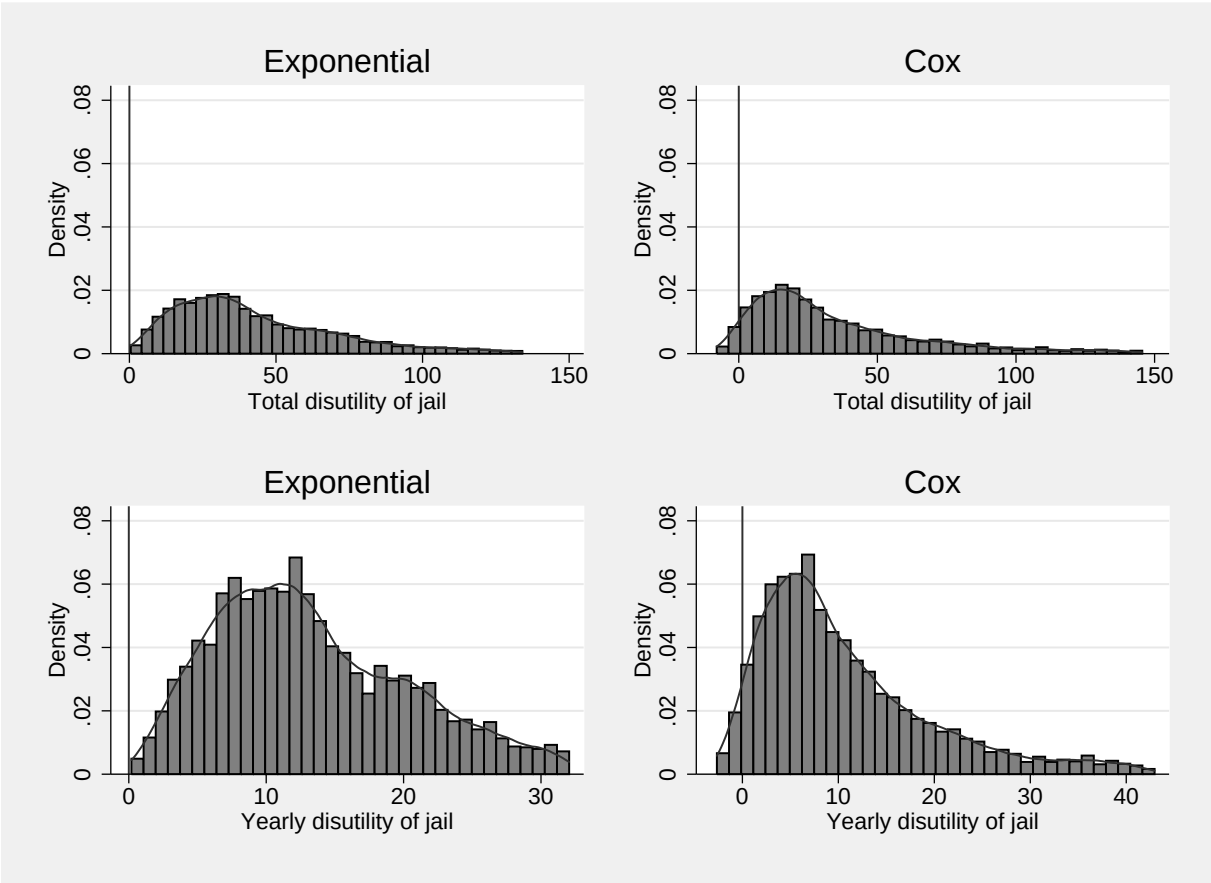
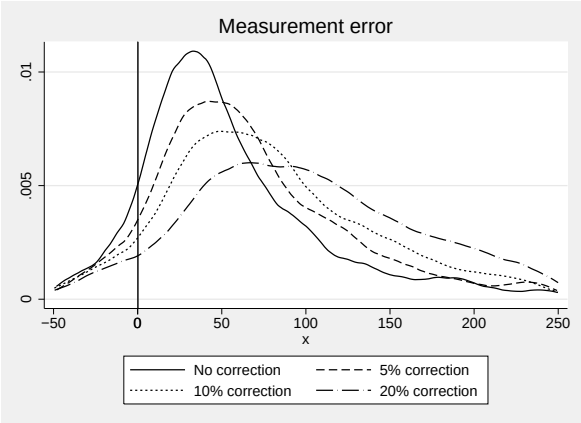
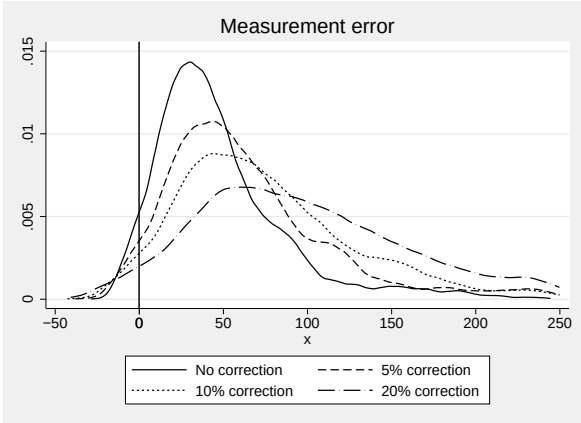


Figure 6: The Distribution of the Conditional Value of Freedom ($t \leq 30$ minutes)

Notes: The upper panels show the distribution of the total disutility of jail, the lower ones the corresponding yearly figures assuming a discount factor of one (between the first and the 95th percentile). These estimates are based on successful robberies only that lasted less than 30 minutes.



$t \leq 9$ minutes



$t < 30$ minutes

Figure 7: The Distribution of the Conditional Value of Freedom Depending on the Measurement Error

Notes: The relative bias correction of size m changes D according to $m\frac{\partial D_i}{\partial \log \beta_i} = m\frac{\beta_i(1-t_i^*\lambda(t_i^*,x_i))}{\lambda(t_i^*,x_i)}$, where $\beta_i = \beta_{t^*}t_i^* + \beta'_{t^*w}w_it_i^*$ represents the individual slope with respect to t.

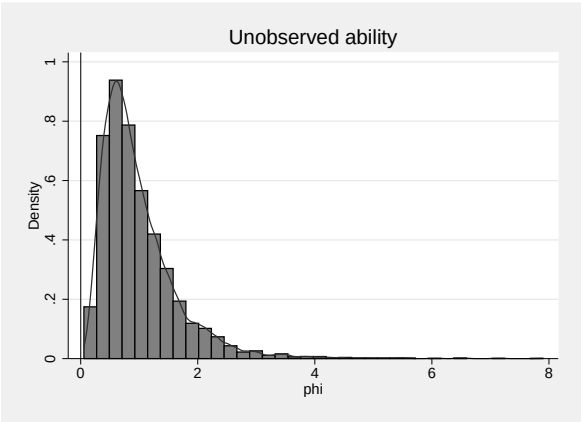


Figure 8: The Distribution of Unobserved Ability

Notes: The figure shows $f(\phi)$ but φ is drawn from the same distribution, and the two unobserved abilities are negatively correlated with each other (-20 percent).

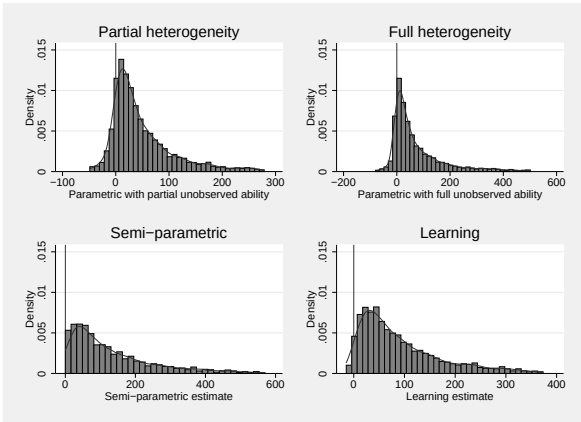


Figure 9: The Distribution of the Conditional Value of Freedom ($t \leq 9$ minutes)

Notes: The upper panels show the distribution of the total disutility of jail, the lower ones the corresponding yearly figures assuming a discount factor of one (between the first and the 95th percentile). These estimates are based on successful robberies that lasted less than 9 minutes.

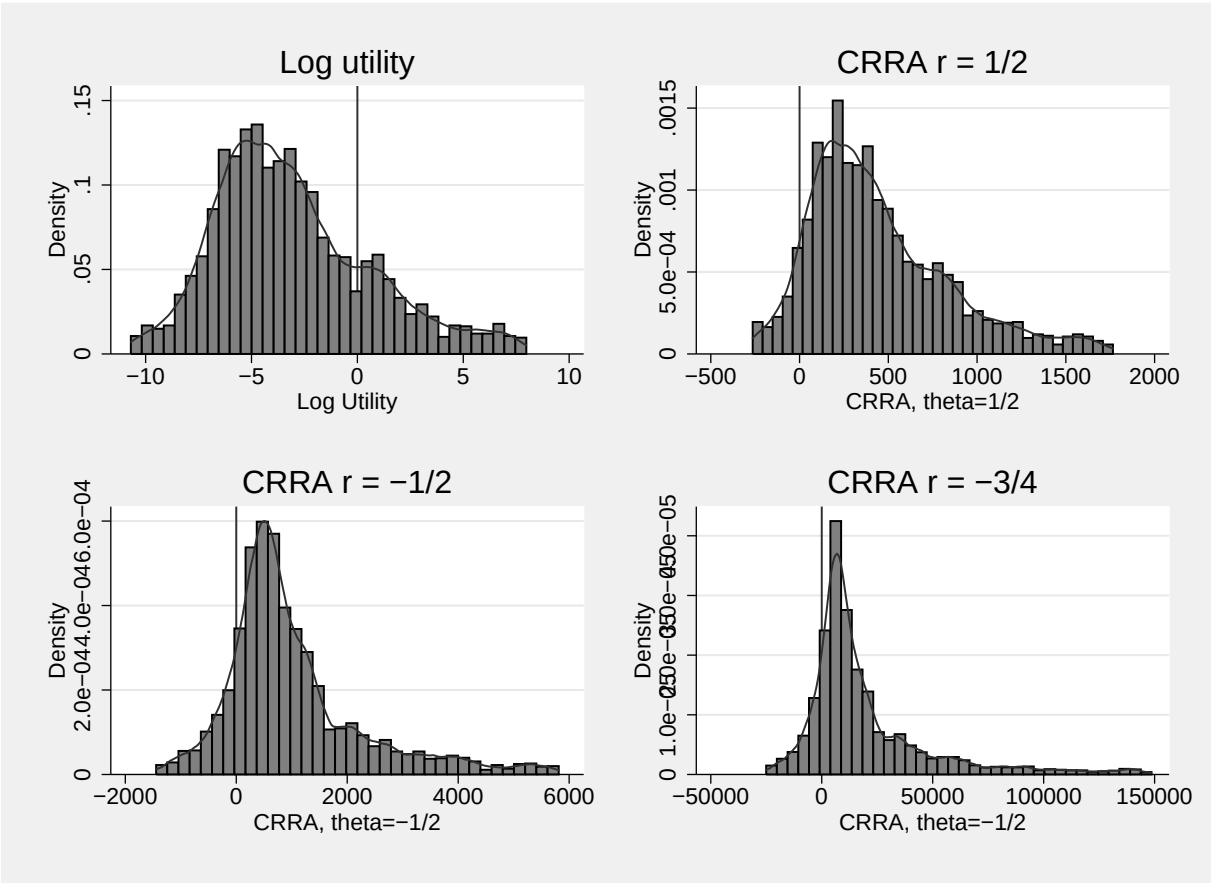


Figure 10: The Distribution of the Conditional Value of Freedom by Risk Preference

Notes: Each plot show the distribution of the total disutility of jail for different degrees of risk aversion (between the first and the 95th percentile).

Table 1: Number of Bank Robberies across the World

	Total Robberies	R. per Branch (in %)		Total Robberies	R. per Branch
Andorra	0	0	Japan	133.29	0.98
Australia	119	2.54	Liechtenstein	0	0
Belgium	117.43	1.37	Lithuania	12.29	1.79
Bulgaria	1	0.32	Luxembourg	2.14	0.71
Canada	827.71	14.1	Malta	0.71	0.7
Croatia	27.43	2.45	Monaco	0	0
Cyprus	6.57	0.91	New Zealand	25.14	2.18
Czech Republic	66.29	4.08	Norway	11.86	0.96
Denmark	160.14	7.91	Poland	72.71	0.61
Estonia	1.71	0.69	Portugal	97.29	1.78
Finland	8.71	0.53	Slovak Republic	13.57	1.16
France	639.29	2.28	Slovenia	11.57	1
Germany	837.71	1.96	Spain	523.43	1.36
Greece	143.57	3.68	Sweden	38.86	2
Hungary	33.29	1.03	Switzerland	16.29	0.43
Iceland	2.71	1.66	The Netherlands	77.14	2.41
Ireland	64.57	5.22	Turkey	83.86	1.22
Italy	2770.86	8.67	UK	191.86	1.74

Source: European Banking Federation. "Total Robberies" are the average yearly number of robberies from 2000 to 2006.

Table 2: Probability of Success and Average Haul (in €) in Different Countries

	P(success)	Av. Haul		P(success)	Av. Haul
Australia	0.56	14227	Italy	0.9	20183
Belgium	0.57	47434	Japan	0.29	.
Bulgaria	1	12880	Lithuania	0.55	63545
Canada	0.97	3011	Norway	0.5	807
Croatia	0.87	25592	Poland	0.71	5502
Cyprus	1	35548	Portugal	0.89	8643
Czech Republic	0.75	11053	Slovak Republic	0.86	11200
Denmark	0.93	22023	Slovenia	0.7	2591
Estonia	1	4470	Spain	0.92	16065
Finland	0.67	804795	Sweden	0.71	18608
France	0.79	14331	Switzerland	0.65	90065
Germany	0.76	32417	The Netherlands	0.41	60380
Greece	0.89	29307	Turkey	0.73	4848
Hungary	0.5	17003	UK	0.6	32827
Ireland	0.82	8626			

Source: European Banking Federation for the year 2006.

Table 3: "Life table" of bank robberies

Time	Number surviving to time $t-1$	Arrested	Successful	Total
		between $t-1$ and t		
1	4972	24	273	297
2	4675	71	1049	1120
3	3555	99	1572	1671
4	1884	31	477	508
5	1376	53	702	755
6	621	4	71	75
7	546	5	50	55
8	491	1	55	56
9	435	0	12	12
10	423	20	169	189
11	234	0	4	4
12	230	0	9	9
13	221	2	9	11
14	210	0	3	3
15	207	7	41	48
16	159	1	4	5
17	154	1	2	3
18	151	5	0	5
19	146	0	4	4
20	142	9	49	58
22	84	0	2	2
23	82	0	3	3
25	79	0	29	29
27	50	0	1	1
28	49	0	1	1
29	48	0	1	1
30	47	3	44	47

Notes: This table shows the distribution of successful and unsuccessful bank robberies that last at most half an hour.

Table 4: Summary statistics

Sample	Whole		duration $\leq$ median (3 min)		duration $>$ median	
	Mean	SD	Mean	SD	Mean	SD
Arrested	6.33%	24.35%	6.28%	24.27%	6.43%	24.54%
Duration of the robbery (*in minutes*)	3.24	1.40	2.44	0.66	4.93	1.01
Total haul	13,778	24,291	11,559	14,959	18,469	36,505
Haul	7,879	11,772	7,025	8,736	9,684	16,294
Firearms	0.14	0.34	0.12	0.33	0.16	0.37
Two robbers	0.52	0.50	0.51	0.50	0.56	0.50
Three or more robbers	0.14	0.35	0.11	0.32	0.19	0.40
Masked robbers	0.43	0.50	0.43	0.50	0.43	0.50
Center Italy	0.21	0.41	0.20	0.40	0.22	0.42
South Italy	0.28	0.45	0.27	0.45	0.30	0.46
Guarded	0.08	0.27	0.07	0.26	0.09	0.28
Isolated branch	0.25	0.43	0.25	0.43	0.25	0.43
Bank with little cash	0.63	0.48	0.62	0.48	0.65	0.48
Bank with less than 5 employees	0.51	0.50	0.50	0.50	0.53	0.50
Number of Security Devices	5.62	1.17	5.62	1.16	5.62	1.20
Average Number of Characteristics per	1.26	0.38	1.27	0.39	1.24	0.36
% of invisible devices	0.67	0.16	0.68	0.16	0.66	0.15
N.obs.	4,549		3,088		1,461	

Notes: This table shows the summary statistics for the sample of bank robberies that last less than 10 minutes. The last four columns split the sample depending on whether the duration is above or below 3 minutes.

Table 5: Linear Regressions of the Per-Capita Haul

	(1)	(2)	(3)	(4)	(5)
			Haul		
Durations		$t < 10$ minutes		$t \leq 30$ minutes	
Duration of the robbery (in minutes)	1,062.13***	1,079.91***	3,151.39***	741.10***	1,227.89***
	(187.58)	(186.81)	(843.98)	(70.62)	(243.24)
Firearms		4,020.51***	-2,040.58	4,195.42***	3,317.52***
		(835.82)	(1,752.23)	(793.43)	(973.17)
Two robbers		-2,598.85***	-2,611.24***	-2,496.26***	-2,492.60***
		(373.28)	(370.23)	(388.13)	(383.68)
Three or more robbers		-3,010.80***	-3,025.82***	-3,225.07***	-3,229.65***
		(676.84)	(664.33)	(668.64)	(664.61)
Masked robbers		1,364.66***	1,304.85***	1,865.47***	1,807.49***
		(365.44)	(368.74)	(387.07)	(386.61)
Center Italy		1,600.39***	1,519.19***	1,647.11***	1,580.43***
		(398.44)	(398.24)	(413.86)	(411.58)
South Italy		1,650.65***	1,644.49***	2,302.38***	2,232.74***
		(478.67)	(474.08)	(523.12)	(522.23)
Isolated branch		-378.98	-415.72	-323.20	-350.71
		(353.66)	(352.96)	(369.60)	(364.26)
Bank with little cash		-1,336.68***	1,446.88	-1,495.81***	-183.54
		(425.89)	(1,353.89)	(450.88)	(602.56)
Bank with less than 5 employees		-368.96	-381.05	-1,018.54**	-1,034.68**
		(382.61)	(380.81)	(411.44)	(414.15)
Number of Security Devices		-248.98**	-285.25**	-387.19**	-412.80***
		(123.36)	(123.90)	(154.29)	(153.66)
Average Number of Characteristics		-2,492.52***	1,620.68	-2,925.49***	-1,828.12***
per Security Device		(386.25)	(1,046.48)	(409.76)	(591.49)
% of invisible devices		-1,955.32**	-2,233.70**	-1,911.84*	-2,030.25**
		(957.68)	(969.39)	(1,021.55)	(1,021.78)
Guarded		-345.05	-557.58	-350.92	-350.59
		(800.82)	(796.12)	(810.04)	(810.44)
Interaction					
Firearms			1,795.36***		162.74
			(669.29)		(148.75)
Bank with little cash			-879.13*		-301.39**
			(477.00)		(147.01)
Average Number of Characteristics			-1,374.38***		-275.43**
per Security Device			(356.06)		(139.42)
Constant	4,435.89***	yes	yes	yes	yes
	(538.55)				
Observations	4549	4549	4549	4972	4972
R-squared	0.016	0.058	0.070	0.100	0.103

Notes: The average haul is modeled as a linear function of the duration and of the *modus operandi*. In columns 3 the duration of the bank robbery is interacted with all the variables that in a first step (not shown) had a significant coefficient. Columns 4 and 5 replicate columns 2 and 3 but using robberies that last up to 30 minutes and not just 9. Robust standard errors in parentheses: : *** p<0.01, ** p<0.05, * p<0.1

Table 6: Hazard Models of Arrest

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$t < 10$ minutes				$t \leq 30$ minutes			
	Exponential		Cox Proportional		Exponential		Cox Proportional	
Firearms	-0.38*	-0.42**	-0.40**	-0.44**	-0.61***	-0.63***	-0.63***	-0.67***
	(0.20)	(0.21)	(0.20)	(0.21)	(0.178)	(0.177)	(0.180)	(0.180)
Two robbers	-0.48***	-0.47***	-0.55***	-0.54***	-0.59***	-0.57***	-0.65***	-0.64***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.116)	(0.116)	(0.118)	(0.119)
Three or more robbers	-0.42**	-0.48**	-0.54***	-0.60***	-0.73***	-0.76***	-0.82***	-0.87***
	(0.19)	(0.19)	(0.19)	(0.19)	(0.163)	(0.168)	(0.167)	(0.172)
Masked robbers	-0.63***	-0.65***	-0.61***	-0.64***	-0.58***	-0.59***	-0.56***	-0.57***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.119)	(0.120)	(0.120)	(0.122)
Center Italy		-0.19		-0.19		-0.10		-0.11
		(0.17)		(0.17)		(0.148)		(0.150)
South Italy		0.07		0.02		-0.04		-0.07
		(0.14)		(0.14)		(0.129)		(0.130)
Isolated branch		-0.04		-0.04		-0.13		-0.12
		(0.15)		(0.15)		(0.134)		(0.135)
Bank with little cash		-0.03		-0.03		0.06		0.06
		(0.13)		(0.13)		(0.111)		(0.112)
Bank with less than 5 employees		-0.36***		-0.40***		-0.32***		-0.35***
		(0.12)		(0.12)		(0.109)		(0.111)
Number of Security Devices		-0.11**		-0.11**		-0.09**		-0.08*
		(0.05)		(0.05)		(0.044)		(0.044)
Average Number of Characteristics		-0.04		0.02		0.07		0.11
per Security Device		(0.16)		(0.16)		(0.139)		(0.142)
% of invisible devices		-0.11		-0.10		-0.14		-0.14
		(0.36)		(0.36)		(0.330)		(0.333)
Guarded		0.12		0.19		0.12		0.16
		(0.23)		(0.23)		(0.214)		(0.217)
N.obs.	4,549	4,549	4,549	4,549	4,972	4,972	4,972	4,972

Notes: This table shows the estimated coefficients of an exponential (columns 1-3) and a Cox proportional (columns 4-6) hazard model of arrest. Robust standard errors in parentheses: : *** p<0.01, ** p<0.05, * p<0.1

Table 7: Summary statistics from trials related to bank robberies

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
	Characteristics of bank robbers				
Age	35.691	10.194	18	65	94
Foreigner	0.083	0.278	0	1	96
Southern	0.344	0.477	0	1	96
Number of robberies	3.365	3.369	1	15	96
Recidivist	0.667	0.474	0	1	96
Plea bargain	0.344	0.477	0	1	96
Total sentence	3.452	1.647	1.333	12.667	94
	Characteristics of robberies				
Firearms	0.22	0.415	0	1	323
Masked	0.570	0.496	0	1	323
Group robbery	0.687	0.464	0	1	323
Hostages	0.04	0.197	0	1	323
Total haul	12.417	21.667	0	145	323
Year	2004.898	1.474	1993	2007	322

Notes: These data are based on trials against 95 bank robbers, involved in a total of 323 bank robberies organized between 1997 and 2007, that were held in the judicial district of Piedmont.

Table 8: Determinants of the Sentence Length

	(1)	(2)	(3)	(4)
		log-Sentence		
Firearms	0.50***	0.36**	0.39***	0.28***
	(0.09)	(0.16)	(0.10)	(0.09)
Masked	0.10	0.08	0.07	0.03
	(0.09)	(0.08)	(0.08)	(0.08)
Group robbery	0.25***	0.14	0.20**	0.09
	(0.09)	(0.11)	(0.08)	(0.08)
Number of robberies		0.02		0.03**
		(0.01)		(0.02)
Recidivist		-0.03		-0.03
		(0.11)		(0.08)
Hostages		0.05		-0.10
		(0.10)		(0.18)
Total haul		0.00*		0.00
		(0.00)		(0.00)
Plea bargain		-0.21*		-0.27***
		(0.12)		(0.08)
Year	-0.00	-0.01		-0.02
	(0.03)	(0.02)		(0.02)
Observations	316	316	95	94
R-squared	0.331	0.431	0.197	0.361

Notes: These regressions are based on trials against 95 bank robbers, involved in a total of 323 bank robberies organized between 1997 and 2007, that were held in the judicial district of Piedmont. Robust standard errors in parentheses: : *** p<0.01, ** p<0.05, * p<0.1

Table 9: Conditional Heterogeneity in D (in 1,000s of €)

	% Negative	Mean	St. Dev.	C. Var.	P10	P25	P50	P75	P90
				Sample with $t \leq 9$, $N = 4,054$					
				Total disutility					
Parametric, exponential	0.103	71.41	93.08	1.30	-0.23	20.18	44.83	87.59	179.09
Parametric, Cox	0.117	66.68	128.51	1.93	-2.40	11.46	32.65	76.00	162.82
				Yearly disutility					
Parametric, exponential	0.103	20.14	21.28	1.06	-0.08	7.51	15.34	27.72	46.77
Parametric, Cox	0.117	18.75	30.81	1.64	-0.88	4.12	11.27	23.72	46.93
				Sample with $t \leq 30$, $N = 4,370$					
				Total disutility					
Parametric, exponential	0.002	49.36	40.81	0.83	13.08	22.87	37.41	63.31	97.49
Parametric, Cox	0.045	47.61	85.00	1.79	4.11	12.56	26.14	53.21	98.72
				Yearly disutility					
Parametric, exponential	0.002	14.53	9.05	0.62	5.02	8.06	12.50	19.30	26.50
Parametric, Cox	0.045	14.11	21.17	1.50	1.48	4.26	8.61	16.25	29.41

Notes: This table shows the mean, the standard deviation, the compensating variation, and the 10th, 25th, 50th, 75th, and 90th percentile of the disutility of jail using exponential and Cox proportional hazard rates. The yearly figures are estimated dividing the yearly figures by the predicted sentence length based on the regression shown in Column 3 of Table 8.

Table 10: log-Value of Freedom changes

	$t \leq 9$				$t \leq 30$			
	Average	SD	P5	P95	Average	SD	P5	P95
Firearms	1.78	0.63	1.17	2.92	0.98	0.17	0.83	1.25
Two robbers	0.68	0.37	0.49	1.06	0.85	0.28	0.63	1.25
Three or more robbers	0.64	0.28	0.50	0.94	0.99	0.23	0.82	1.38
Masked robbers	0.77	0.22	0.67	0.96	0.69	0.13	0.60	0.90
Central Italy	0.18	0.06	0.16	0.21	0.07	0.05	0.02	0.10
South Italy	-0.15	0.18	-0.31	-0.08	-0.04	0.15	-0.17	0.03
Isolated branch	0.07	0.10	0.05	0.12	0.18	0.06	0.14	0.27
Bank with little cash	-0.90	0.57	-1.91	-0.27	-0.62	0.25	-1.00	-0.42
Bank with less than 5 employees	0.46	0.21	0.37	0.68	0.47	0.19	0.36	0.73
Number of Security Devices	0.22	0.26	0.12	0.50	0.21	0.23	0.11	0.50
Average Number of Characteristics	-0.78	0.26	-1.12	-0.33	-0.42	0.09	-0.56	-0.32
% of invisible devices	0.22	0.19	0.13	0.44	0.26	0.16	0.16	0.50
Guarded	-0.14	0.08	-0.16	-0.12	-0.15	0.07	-0.21	-0.13

Notes: This table shows the log-changes in the disutility of jail that correspond to a unitary change in the use of firearms, etc, together with the standard deviation, the 5th and the 95th percentile.

Table 11: Measurement Error in Duration t and the Disutility D

	% Negative	Mean	St. Dev.	C. Var.	P10	P25	P50	P75	P90
				Sample with $t \leq 9$ min., $N = 4,054$					
No correction	0.103	71.41	93.08	1.30	-0.23	20.18	44.83	87.59	179.09
5% correction	0.095	102.44	143.57	1.40	1.52	29.06	62.31	116.44	240.81
10% correction	0.092	133.47	196.39	1.47	3.28	37.64	77.56	144.88	315.67
20% correction	0.088	195.52	304.15	1.56	7.35	54.40	108.67	200.34	445.74
				Sample with $t \leq 30$ min., $N = 4,370$					
No correction	0.002	49.36	40.81	0.83	13.08	22.87	37.41	63.31	97.49
5% correction	0.001	80.99	95.57	1.18	20.31	32.83	54.57	90.80	152.30
10% correction	0.001	112.63	158.77	1.41	26.34	42.34	70.12	118.73	212.87
20% correction	0.000	175.90	288.42	1.64	38.66	61.05	101.26	171.35	342.47

Notes: This table shows how a 5, 10, and 20 percent correction in the marginal haul due to classical measurement error in the duration of the bank robbery influences the estimated disutility of jail D.

Table 12: Change in $\log D$ that Corresponds to $t^* = 0$

	Mean	St. Dev.	P5	P25	P50	P75	P95
	Sample with $t \leq 9$ min., $N = 4,054$						
No correction	0.27	0.37	0.03	0.09	0.17	0.31	0.77
5% correction	0.22	0.31	0.03	0.08	0.14	0.25	0.59
10% correction	0.19	0.27	0.03	0.07	0.13	0.21	0.48
20% correction	0.15	0.20	0.03	0.06	0.11	0.18	0.38
	Sample with $t \leq 30$ min., $N = 4,370$						
No correction	0.15	0.18	0.03	0.06	0.10	0.18	0.45
5% correction	0.12	0.13	0.02	0.05	0.09	0.14	0.33
10% correction	0.11	0.11	0.02	0.05	0.08	0.13	0.28
20% correction	0.10	0.11	0.02	0.04	0.07	0.11	0.24

Notes: This table shows the mean, the standard deviation, and the 5th, 25th, 50th, 75th, and 95th percentile of $\Delta \log D$ that is needed to drive the optimal duration of the bank robbery to 0.

Table 13: High and Low Responsiveness

	Average characteristic			
	$t \leq 9$ min.		$t \leq 30$ min.	
Responsiveness	*high*	*low*	*high*	*low*
$\Delta \log D$	0.09	0.44	0.06	0.25
Disutility	133.22	31.08	69.95	28.99
Firearms	0.31	0.00	0.25	0.07
Two robbers	0.58	0.48	0.62	0.43
Three or more robbers	0.16	0.12	0.19	0.13
Masked robbers	0.64	0.27	0.64	0.26
Central Italy	0.28	0.17	0.26	0.16
South Italy	0.22	0.31	0.26	0.31
Isolated branch	0.26	0.25	0.28	0.21
Bank with little cash	0.49	0.74	0.54	0.72
Bank with less than 5 employees	0.51	0.52	0.54	0.49
Number of Security Devices	5.75	5.59	5.76	5.51
Average Number of Characteristics	1.13	1.22	1.20	1.32
% of invisible devices	0.70	0.66	0.69	0.66
Guarded	0.09	0.08	0.09	0.07

Notes: This table shows the average X's according to whether the $\Delta \log D$ needed to drive the optimal duration of the bank robbery to 0 is above or below the median. A lower change corresponds to a higher *responsiveness*.

Table 14: Unobserved Ability and the Disutility D

	% Negative	Mean	St. Dev.	P10	P25	P50	P75	P90
				Disutilities D				
$\phi = \varphi = 1$	0.103	71.41	93.08	-16.16	20.18	44.83	87.59	274.15
$\phi \sim logN(\cdot,\cdot), \varphi = 1$	0.131	66.72	118.38	-13.64	9.78	31.95	78.47	264.13
$\phi, \varphi \sim logN(\cdot,\cdot)$	0.152	115.69	304.97	-14.78	8.89	39.17	115.66	466.80
Semi-parametric	0.000	170.76	314.05	0.00	34.65	88.28	192.81	577.62
Learning	0.033	121.11	186.81	3.49	32.07	70.53	143.53	383.55
				Change in $\log D$ that Corresponds to $t^* = 0$				
$\phi = \varphi = 1$	-	0.27	0.37	0.03	0.09	0.17	0.31	0.77
$\phi \sim logN(\cdot,\cdot), \varphi = 1$	-	0.38	0.52	0.03	0.09	0.21	0.45	1.27
$\phi, \varphi \sim logN(\cdot,\cdot)$	-	0.34	0.54	0.02	0.06	0.16	0.40	1.30
Semi-parametric	-	0.07	0.04	0.02	0.04	0.06	0.08	0.15
Learning	-	0.14	0.22	0.03	0.06	0.09	0.16	0.39

Notes: This table shows the mean, the standard deviation, and the 5th, 25th, 50th, 75th, and 95th percentile of the disutility of jail as well as of $\Delta \log D$ that is needed to drive the optimal duration of the bank robbery to 0 when unobserved heterogeneity is added to the estimation. The estimates allow for unobserved ability in the marginal haul and in the hazard. The first column shows the fraction of disutilities that are negative.

Table 15: High and Low Responsiveness

	Average characteristic					
	$\phi \sim logN(\cdot,\cdot)$		Semi-parametric		Learning	
Responsiveness	*high*	*low*	*high*	*low*	*high*	*low*
$\Delta \log D$	0.12	0.58	0.04	0.10	0.06	0.23
Disutility	139.68	26.29	274.58	94.90	205.79	46.47
Firearms	0.28	0.04	0.19	0.08	0.25	0.04
Two robbers	0.56	0.53	0.62	0.45	0.60	0.47
Three or more robbers	0.17	0.12	0.14	0.13	0.16	0.13
Masked robbers	0.59	0.35	0.70	0.17	0.69	0.20
Central Italy	0.27	0.19	0.29	0.14	0.30	0.13
South Italy	0.23	0.30	0.23	0.32	0.24	0.32
Isolated branch	0.27	0.24	0.29	0.22	0.28	0.22
Bank with little cash	0.54	0.68	0.65	0.60	0.60	0.66
Bank with less than 5 employees	0.50	0.52	0.58	0.45	0.54	0.48
Number of Security Devices	5.70	5.65	5.75	5.48	5.73	5.55
Average Number of Characteristics	1.14	1.20	1.26	1.26	1.18	1.29
% of invisible devices	0.69	0.67	0.67	0.67	0.68	0.67
Guarded	0.08	0.08	0.07	0.07	0.08	0.08

Notes: This table shows the average X's according to whether the $\Delta \log D$ needed to drive the optimal duration of the bank robbery to 0 is above or below the median. A lower change corresponds to a higher *responsiveness*.

Table 16: Expected Utilities With Different Risk Preferences

	(1)	(2)	(3)	(4)
			Haul	
Utility:	Log	$r = -1/2$	$r = 1/2$	$r = -3/4$
Duration of the robbery (in minutes)	0.26***	24.52***	63.97***	1,023.87**
	(0.05)	(4.62)	(23.49)	(413.84)
Firearms	0.08	-7.47	-62.78	-1,104.58
	(0.14)	(12.97)	(40.40)	(687.16)
Two robbers	-0.36***	-24.41***	-33.48***	-415.03***
	(0.04)	(3.13)	(7.43)	(118.12)
Three or more robbers	-0.42***	-29.30***	-36.09**	-402.92
	(0.06)	(4.80)	(16.82)	(298.81)
Masked robbers	0.17***	11.44***	17.86**	222.44
	(0.04)	(2.85)	(8.64)	(148.93)
Center Italy	0.16***	14.00***	18.98***	219.35**
	(0.05)	(3.62)	(7.03)	(106.35)
South Italy	0.08	6.33*	38.13***	610.40***
	(0.05)	(3.53)	(11.28)	(193.48)
Isolated branch	-0.01	-0.92	-7.92	-109.47
	(0.04)	(3.06)	(7.23)	(118.38)
Bank with little cash	0.12	7.16	46.33	865.61
	(0.10)	(8.30)	(35.98)	(627.80)
Bank with less than 5 employees	-0.19***	-6.37**	0.62	53.50
	(0.04)	(2.90)	(9.01)	(155.52)
Number of Security Devices	-0.05***	-3.42***	-3.19	-37.92
	(0.02)	(1.20)	(2.17)	(33.15)
Average Number of Characteristics	-0.05	0.53	52.50**	917.49**
per Security Device	(0.12)	(8.29)	(25.49)	(436.09)
% of invisible devices	-0.64***	-42.11***	-10.55	-31.51
	(0.12)	(8.65)	(19.02)	(302.25)
Guarded	-0.22***	-16.32***	-0.24	4.71
	(0.08)	(5.89)	(18.36)	(309.44)
Interaction				
Firearms	0.06	9.09**	42.68**	696.34**
	(0.04)	(4.12)	(17.16)	(300.17)
Bank with little cash	-0.07**	-6.27**	-22.33*	-394.98*
	(0.03)	(2.61)	(13.17)	(232.61)
Average Number of Characteristics	-0.11***	-9.27***	-28.86***	-461.04***
per Security Device	(0.03)	(2.57)	(9.19)	(160.42)
Observations	4155	4549	4549	4549
R-squared	0.067	0.075	0.050	0.042

Notes: The utility is modeled as a linear function of the duration and of the *modus operandi*. Column 1 assumes a logarithmic utility of the per-capita haul, column 2, 3, and 4 a constant relative risk aversion utility with risk aversion parameter 1/2, -1/2 (risk loving), and -3/4 (very risk loving) parameter. The "risk loving" utilities are divided by 10,000. All the columns use the same regressors as column 3 in Table 5. *** p<0.01, ** p<0.05, * p<0.1

Table 17: Change in $\log D$ that Corresponds to $t^* = 0$

	Negative D	Mean	St. Dev.	P5	P25	P50	P75	P95
Log utility	0.714	0.41	0.52	0.05	0.11	0.22	0.50	1.25
CRRA, $r = 1/2$	0.078	0.27	0.36	0.04	0.09	0.16	0.30	0.81
Linear utility	0.096	0.27	0.37	0.03	0.09	0.17	0.31	0.77
CRRA, $r = -1/2$	0.128	0.31	0.47	0.03	0.09	0.19	0.35	0.94
CRRA, $r = -3/4$	0.142	0.33	0.43	0.03	0.09	0.21	0.39	1.08

Notes: This table shows the mean, the standard deviation, and the 5th, 25th, 50th, 75th, and 95th percentile of $\Delta \log D$ that is needed to drive the optimal duration of the bank robbery to 0.