INDIVIDUAL PERCEPTIONS OF THE
CRIMINAL JUSTICE SYSTEM

Lance Lochner

Individual Perceptions of the Criminal Justice System
Lance Lochner
NBER Working Paper No. 9474
January 2003, Revised June 2006
JEL No. K4, D8

ABSTRACT

This paper empirically examines perceptions of the criminal justice system held by young males using longitudinal survey data from the recent National Longitudinal Survey of Youth 1997 Cohort and the National Youth Survey. While beliefs about the probability of an arrest are positively correlated with local official arrest rates, they are largely idiosyncratic and unresponsive to information about the arrests of other random individuals and local neighborhood conditions. There is little support, therefore, for the `broken windows' theory of Wilson and Kelling (1982). Yet, perceptions do respond to changes in an individual's own criminal and arrest history. Young males who engage in crime but are not arrested revise their perceived probability of arrest downward, while those who are arrested revise their probability upwards. Beliefs respond similarly to changes in a sibling's criminal and arrest history. The perceived probability of arrest is then linked to subsequent criminal behavior. Cross-sectionally, youth with a lower perceived probability of arrest are significantly more likely to engage in crime during subsequent periods. Following an arrest, individuals commit less crime, consistent with deterrence theory and the fact that their perceived probability of arrest increases.

Lance Lochner
Department of Economics
University of Rochester
Rochester, NY 14627
and NBER
lance@troi.cc.rochester.edu

Individual Perceptions of the Criminal Justice System[*]

Lance Lochner
University of Western Ontario
NBER

May, 2005

Abstract

This paper empirically examines perceptions of the criminal justice system held by young males using longitudinal survey data from the recent National Longitudinal Survey of Youth 1997 Cohort and the National Youth Survey. While beliefs about the probability of an arrest are positively correlated with local official arrest rates, they are largely idiosyncratic and unresponsive to information about the arrests of other random individuals and local neighborhood conditions. There is little support, therefore, for the 'broken windows' theory of Wilson and Kelling (1982). Yet, perceptions do respond to changes in an individual's own criminal and arrest history. Young males who engage in crime but are not arrested revise their perceived probability of arrest downward, while those who are arrested revise their probability upwards. The perceived probability also declines when a sibling engages in crime, but it does not appear to rise when that sibling is arrested. Additional estimates suggest that beliefs about the probability of arrest deter crime. This is true cross-sectionally (i.e. individuals with a higher perceived probability of arrest commit less crime) and when following an individual over time (accounting for unobserved individual heterogeneity).

1   Introduction

The economics literature on crime implicitly assumes that individuals are well-informed about arrest and conviction rates (as well as sentencing policies) and, therefore, respond immediately to any changes in the criminal justice system. Many criminologists have questioned this assumption based on studies of perceived certainty and severity of punishment. Empirical studies of deterrence theory are, therefore, based on two distinct approaches: one which estimates the effect of *actual* or official measures of certainty and severity of punishment (or observed measures related to certainty and severity of punishment, e.g. the number of police in an area) on crime; and another which measures the effect

of *perceived* certainty or severity of punishment on crime. Most studies of the former type conclude that deterrence effects are important while studies of the latter type provide mixed results.[1]

It is worth noting that many of the perceptions-based studies do not necessarily measure the effects of beliefs on actual criminal outcomes. Instead, they often measure differences in intentions to commit crime or the likelihood that someone will engage in crime given a set of circumstances surrounding an opportunity.[2] Unfortunately, the vast differences in design across studies – studies vary substantially in sample population, types of criminal activity examined, measures of crime (actual behavior, intentions, or likelihood), and the nature of questions about perceptions – and the diverse findings make it difficult to draw any strong conclusions from perceptual studies about the effect of a change in the perceived probability of arrest on actual criminal behavior. More importantly, Nagin (1998) recently notes that "While great effort has been committed to analyzing the links between sanction risk perceptions and behavior, comparatively little attention has been given to examining the origins of risk perceptions and their connection to actual sanction policy."[3]

Understanding the evolution of beliefs is important for analyzing a number of social and economic decisions. Given our context, it is useful to focus on the role of beliefs in theories of crime. Sah (1991) provides a theoretical analysis of crime based on a model in which individual beliefs about the probability of punishment are determined by the number of people they observe committing crime and their arrest rates. His theory suggests interesting dynamic responses to changes in criminal enforcement policy as well as levels of segregation. This paper outlines a complementary framework for analyzing how an individual's own crime and arrest history affects his beliefs and how those beliefs affect behavior.[4] Individuals with similar tastes and initial beliefs may follow different crime paths over their lives if they are arrested at different rates (or even arrested at different points in their criminal careers). In Sah's model and the framework discussed in this paper, there are delayed responses in criminal activity when official arrest rates increase. Furthermore, even a temporary increase in arrest rates can have long-term impacts on crime rates. The significance of these results depends on the relevance of and information used in belief updating. While a few empirical studies[5] have found that

---

[1]Studies using actual police, arrest, or punishment measures include Blumestein, et al. (1978), Cameron (1988), Ehrlich (1973,1981), Grogger (1991), Levitt (1997, 1998a, 1998b), Myers (1983), Tauchen, Witte, and Griesinger (1994), Trumbull (1989), Waldfogel (1993), and Witte (1980). Deterrence studies using perceived measures of certainty or severity include Tittle (1977), Paternoster, et al, (1983), Piliavin, et al., (1986), Viscusi (1986), Klepper and Nagin (1989), Grasmick and Bursick (1990), Schneider and Ervin (1990), Bachman, Paternoster, and Ward (1992), Nagin (1993), Paternoster and Simpson (1996), and Pogarsky and Piquero (2003).

[2]See, e.g. Tittle (1977), Klepper and Nagin (1989), Grasmick and Bursick (1990), Bachman, Paternoster, and Ward (1992), Nagin (1993), Paternoster and Simpson (1996), and Pogarsky and Piquero (2003).

[3]Exceptions include Richards and Tittle (1981), Saltzman, et al. (1982), Paternoster, et al. (1983, 1985), Piliavin, et al. (1986), and Horney and Marshall (1992).

[4]This framework is developed more formally and fully analyzed in Lochner (2004).

[5]Taking a VAR approach to estimating the relationship between crime, arrests, and the business cycle, Corman, Joyce,

2

time patterns in crime and arrests are consistent with information transmission and belief updating among criminals, this paper directly examines the empirical importance of individual (and sibling) crime and arrest histories as well as alternative sources of information in determining beliefs about the probability of arrest.

The 'broken windows' theory of Wilson and Kelling (1982) suggests that individuals are more likely to engage in crime in neighborhoods exhibiting decay (i.e. broken windows or abandoned buildings), because they believe they are less likely to be arrested or interfered with. Understanding the information used in generating beliefs and how perceptions influence behavior is central to this theory. In the empirical analysis below, we explore the relationship between neighborhood decay and perceptions among young males.

The economics literature has recently begun to analyze how the evolution of beliefs over time can affect aggregate outcomes. In special environments, the information cascade literature (e.g. Banerjee, 1992, Bikhchandani, Hirshleifer, and Welsh, 1992) has shown that the aggregation of individual decisions can lead to informational cascades and conformity when individuals possess idiosyncratic information and gather information from others. Furthermore, Heavner and Lochner (2002) show that policies like anti-gang initiatives or mentor programs will have heterogeneous impacts on neighborhoods that differ in the current level of gang and criminal activity. More generally, the way in which individuals acquire information and develop expectations is important in determining outcomes and policy effects in any environment; yet, little is actually known about these processes.[6]

This paper makes two contributions to the perceptions-based deterrence literature using self-reported beliefs about the probability of (one's own) arrest from two sources of longitudinal data (the National Longitudinal Survey of Youth 1997 Cohort, NLSY97, and the National Youth Survey, NYS). First, this paper examines which factors influence individual perceptions about the probability of arrest. In particular, we study (i) the correlation between various demographic characteristics (e.g. race and ethnicity) and beliefs about the probability of arrest, (ii) the correlation between local official arrest rates and perceptions, (iii) the role of neighborhood conditions in determining beliefs, and (iv) the role of new information and belief updating in determining perceptions about the probability of arrest. Regarding the last point, this paper shows that individuals update their beliefs in rational ways. Those who engage in crime while avoiding arrest reduce their perceived probability of arrest,

---

and Lovitch (1987) find empirical evidence for both delayed effects of an increase in arrests on crime and for long-term effects of a temporary increase in arrests. Ayres and Levitt (1998) find evidence consistent with learning among auto thieves when Lojack (a new technology allowing police to locate stolen vehicles equipped with the system) is introduced to some cities.

[6]See Manski (1992) for a clear discussion about the importance of understanding expectations formation in studying schooling decisions. Anderson and Holt (1997) study information cascades using an experimental approach.

while those who are arrested increase their perceived probability.[7] Beliefs also respond to changes in the criminal histories of their siblings but not to information about the arrest outcomes of their siblings or other random persons.

The second main contribution of the paper is an empirical analysis of the relationship between the perceived probability of arrest and the subsequent commission of crime. By examining the relationship about the perceived probability of arrest for different crimes and the probability of actual commission of those crimes in large random national samples of young males, the results can be interpreted as actual deterrence effects that might be expected when beliefs are changed through policy intervention on a broad scale. Consistent with deterrence theory, estimates suggest that individuals with a higher perceived probability of arrest are less likely to commit crime in the future. Differences in behavior across individuals are, therefore, not only due to differences in tastes for crime, criminal returns, or opportunity costs. They may also be related to differences in beliefs about the criminal justice system.

To account for unobserved heterogeneity in tastes for crime, we take two approaches using the NLSY97 data. First, we consider instrumental variable estimation in a fixed effects environment using a linear probability model. These estimates provide some evidence that beliefs about the probability of arrest for auto theft affect subsequent engagement in theft. Second, we combine our models of belief updating and criminal behavior to structurally estimate both problems simultaneously (along with the true probability of arrest) in the presence of unobserved heterogeneity in tastes for crime and in the perceived probability of arrest. Once measurement error in beliefs is specifically modelled and taken into account, the structural estimates suggest strong deterrent effects based on the perceived probability of arrest.

After a brief discussion of the main issues involved in studying the evolution of beliefs about the probability of arrest and criminal behavior in Section 2, this paper empirically examines these issues using data from the NLSY97 and NYS. Section 3 summarizes the data on criminal participation and perceptions in these data, exploring how beliefs vary among young males. The role of belief updating is examined in Section 4, and the influence of beliefs about the probability of arrest on criminal activity is discussed in Section 5. Section 6 provides a structural estimation analysis of belief updating, crime, and arrests that is based on the model sketched in Section 2. Section 7 concludes.

---

[7]Early studies by criminologists that explore the link between perceptions and crime report that individuals engaged in crime tend to lower their perceived probability of arrest, referring to these effects as 'experiential effects' (Minor and Harry, 1982, Paternoster, et al., 1983, Piliavin, et al., 1986, Saltzman, et al., 1982). Horney and Marshall (1992) estimate a positive correlation between arrests per crime in recent years and current beliefs, but they cannot determine whether that correlation is due to belief updating or a strong correlation between initial (and, perhaps, stable) beliefs and observed outcomes. Studying 300 college students, Paternoster, et al. (1985) offer some evidence that changes in the perceived probability of arrest are negatively correlated with criminal activity and positively correlated with formal sanctions. None of these studies take into account the endogeneity of criminal behavior.

2   The Evolution of Crime and Beliefs

This section outlines a framework for thinking about the interaction of beliefs about the probability of arrest and criminal behavior. The primary goal is to provide intuition about the important issues involved in the empirical study below rather than a rigorous theoretical treatment of the problem.[8] We also discuss a few policy implications that underscore the potential importance of belief updating in determining criminal decisions over the lifecycle.

Suppose individuals begin with prior beliefs about the probability of arrest for different types of crime and then decide whether or not to engage in crime based on those beliefs. Their decision to commit crime and whether they are arrested will affect their future beliefs about the probability of arrest. Beliefs may also respond to information from various other sources. For example, individuals may observe crimes committed by others and whether or not they are arrested, as in Sah (1991). They may move from one neighborhood to another or observe more police on the street. Using all of this information, individuals continually form new beliefs and decide whether or not to engage in crime. This process repeats itself over the lifecycle. Because ex ante identical agents will receive different information about the probability of arrest, their beliefs and criminal behavior will likely differ at any point in time.

First, consider the decision to commit crime when there is uncertainty about the probability of arrest. Following Becker (1968), assume that individuals choose to commit crime if the expected benefits exceed the expected costs. For simplicity, assume the benefits to each individual i from committing a crime at age t, B_{it}, are known beforehand. Individuals also know the punishment, $J_{it} \geq 0$, associated with an arrest, but they do not necessarily know their own probability of arrest. Instead, they have some beliefs about that probability (π_i). Let $F(\pi|H_i^t)$ represent the cumulative distribution function for an individual's perceived arrest probability conditional on information available to him at some date t, H_i^t. Assuming no intertemporal effects of arrest or criminal behavior (except through beliefs), individual i will commit crime in the following period if and only if

$$B_{it} > J_{it} \int\limits_0^1 \pi dF(\pi|H_i^t).$$

For simplicity, this decision rule ignores any incentive to commit crime in order to learn more about the true probability. In this sense, individuals behave myopically each period.[9] Defining the benefit-cost

---

[8]For a more complete theoretical analysis, see Lochner (2004).

[9]Incorporating this type of strategic behavior is straightforward and would create an additional incentive to engage in crime when beliefs are uncertain.

ratio, $R_{it} = B_{it}/J_{it}$, yields the following decision rule for crime:

$$\text{commit crime if and only if} \qquad E(\pi|H_i^t) < R_{it}, \qquad (1)$$

where $E(\pi|H_i^t) = \int_0^1 \pi dF(\pi|H_i^t)$.

Now, consider the evolution of beliefs. Assume that initial beliefs about the probability of arrest are given by $F_0(\pi)$ (where $F_0(0) = 0$ and $F_0(1) = 1$, reflecting the fact that π is itself a probability). Any number of assumptions can be made about how individuals update their beliefs given new information as well as what types of information are relevant for belief updating. Since the criminal decision rule in equation (1) depends on the expectation of the probability of arrest, $E(\pi|H_i^t)$, we consider how this measure of beliefs evolves.

Beliefs about the probability of arrest are likely to depend on an individual's own (past) criminal behavior and arrest outcomes, the criminal and arrest outcomes of others around him, and more general signals that may come from local arrest rates or neighborhood conditions.[10] Let c_{it} be an indicator equal to one if individual i commits a crime between (survey) dates $t-1$ and t and zero otherwise. Similarly, let A_{it} be an indicator equal to one if he is arrested during that time period and zero otherwise.[11] Let $\tilde{c}_{it}$ and $\tilde{A}_{it}$ represent vectors of these indicators for individuals that person i associates with. Finally, we denote any new information about the local environment by Z_{it}. Information accumulates according to $H_i^t = (H_i^{t-1}, c_{i,t}, A_{i,t}, \tilde{c}_{i,t}, \tilde{A}_{i,t}, Z_{i,t})$. A fairly general rule for updating beliefs is given by

$$E(\pi|H_i^t) = g(E(\pi|H_i^{t-1}), c_{i,t}, A_{i,t}, \tilde{c}_{i,t}, \tilde{A}_{i,t}, Z_{i,t}).^{[12]} \qquad (2)$$

Now, consider a number of potential assumptions for this updating equation. First, the current perceived probability of arrest should be increasing in the previous expected probability ($g_1 \geq 0$).[13] Second, the expected probability of arrest should be decreasing in the number of crimes already committed (by oneself or others) holding the number of arrests constant ($g_2 \leq 0$ and $g_4 \leq 0$). Third, the total effect of committing a crime and getting arrested for it should lead to an increase in the

---

[10]By focusing only on information received from others, Sah (1991) neglects the important role that an individual's own criminal and arrest history plays in shaping his own beliefs and, therefore, subsequent criminal decisions. The true probability of arrest is, most probably, quite heterogeneous across individuals. If this type of heterogeneity is substantial, it may imply that information acquired from others plays little role in the development of an individual's beliefs about the probability that he himself will be arrested. Instead, his own history would be the primary determinant.

[11]The timing implicit in this notation makes it easier to follow the empirical analysis below, since all variables with a t subscript are collected in the same survey (i.e. crime and arrests are retrospective since the previous survey while beliefs are recorded at the time of the survey).

[12]A more general rule would allow $E(\pi|H_i^t)$ to depend on the entire distribution of prior beliefs, $F(\pi|H_i^{t-1})$, rather than just $E(\pi|H_i^{t-1})$.

[13]We denote the partial derivative of $g(\cdot)$ with respect to its kth argument by g_k.

expected probability of arrest (i.e. $g_2 + g_3 \geq 0$ and $g_4 + g_5 \geq 0$).[14] Fourth, one might expect that beliefs are increasing in actual local arrest rates. Fifth, the 'broken windows' theory of Wilson and Kelling (1982) suggests that individuals are likely to think the probability of arrest is lower in communities in which buildings are rundown, windows are broken, and lawlessness is rampant. This paper empirically examines each of these potential assumptions.

Before turning to the empirical analysis, it is worth discussing the implications of these assumptions for lifecycle criminal behavior and the evolution of beliefs. To begin, consider an individual who elects to commit a crime. If he avoids arrest, he will unambiguously lower his perceived probability of arrest (assuming no changes in other information). This will raise the likelihood that he commits crime the following period. On the other hand, if he is arrested, he should raise his expected probability of arrest, making him less likely to commit crime in the future. Thus, criminal profiles will be determined, in part, by the randomness associated with an arrest. The 'lucky' individual who manages to avoid an arrest early on is more likely to continue committing crime thereafter than is the 'unlucky' person who gets arrested. Following the same line of argument, individuals with 'lucky' older siblings who engage in crime and get away with it are more likely to engage in crime themselves.

Much more can be said about the evolution of beliefs and crime if we are willing to make stronger assumptions about the structure of information and updating. For example, consider Bayesian decisionmakers who only acquire information about the probability of arrest from their own criminal and arrest histories.[15] They will update their beliefs as follows:

$$E(\pi|H_i^t) = E(\pi|H_i^{t-1}) - \left[\frac{V(\pi|H_i^{t-1})}{1 - E(\pi|H_i^{t-1})} \right] c_{i,t-1} + \left[\frac{V(\pi|H_i^{t-1})}{E(\pi|H_i^{t-1})(1 - E(\pi|H_i^{t-1}))} \right] c_{i,t-1} A_{i,t-1}, \quad (3)$$

where $V(\pi|H_i^{t-1}) = E(\pi^2|H_i^{t-1}) - \left[E(\pi|H_i^{t-1}) \right]^2$ is the variance of beliefs about the probability of arrest given history H_i^{t-1}. Those not committing crime will not change their beliefs, but those choosing to commit a crime will update their beliefs depending on whether or not they are arrested. The expected probability of arrest increases among those who are arrested, while it decreases among those who are not. The magnitude of the change depends on both the variance and mean of the belief distribution. When there is a lot of uncertainty (i.e. $V(\pi|H_i^{t-1})$ is high), the expected probability of arrest changes a lot in response to new information (whether that new information comes from an arrest or the lack of an arrest). This variance is likely to be particularly high early in an individual's life, while it should decline

---

[14]Contrary to these assumptions, the "gambler's fallacy" posits that someone who is not arrested may feel they are due to be arrested next time (and vice versa). Pogarsky and Piquero (2003) explore this hypothesis in regards to drinking and driving. Empirically, we find no evidence of this effect in either the NLSY97 or NYS data.

[15]Alternatively, individuals may receive information from other sources, but it may be largely irrelevant due to the idiosyncratic nature of criminal ability.

as an individual acquires more and more information. This implies that the beliefs of young criminals should respond more to an arrest than should the beliefs of veteran criminals. Additionally, individuals should learn quickly about the probability of arrest for crimes that are committed frequently. At any given age, then, individuals should respond less to new information about the probability of arrest for these crimes.[16]

In this environment, beliefs need not be accurate. Criminals are likely to be optimistic in that they will tend to believe that their probability of arrest is lower than it actually is, while non-criminals will tend to be pessimistic about their chances of evading arrest. This is even true among those who start their criminal careers with unbiased prior beliefs. To understand why, suppose that all individuals begin with unbiased priors. Any change in beliefs, therefore, leads to a bias. Since individuals only commit crime if the expected probability is low enough, those who continue to engage in crime tend to be the lucky ones who have not been arrested for their past crimes. On average, they reduce their perceived probability of arrest leading to a systematic downward bias. At the other extreme, those choosing not to commit crime are likely to have started out with a very high perceived probability of arrest or to have experienced an arrest sometime in the past causing them to revise their beliefs upwards. The latter subgroup of current non-criminals (but former criminals) will bias the average beliefs of all non-criminals upwards. With homogeneity in the true probability of arrest and unbiased prior beliefs, we would expect that, on average, criminals under-estimate the official arrest rate while non-criminals over-estimate the official arrest rate.

When there is heterogeneity in the true probability of arrest across individuals, average beliefs about the probability of arrest will tend to be higher than official arrest rates even if prior beliefs are unbiased for each individual. This is because those with high true probabilities (and, therefore, high prior beliefs about the probability) will not engage in crime. The opposite is true for those with low true and perceived probabilities. Official arrest rates will be lower than the average true probability across all individuals, since they only reflect the probability of arrest for those choosing to commit a crime. The biases in beliefs discussed earlier will arise among non-criminals and criminals, but the overall average belief about the probability of arrest will generally be higher than the official arrest rate due to selection into criminality. The greater the heterogeneity in true probabilities, the greater will be the difference between average beliefs and official arrest rates.

If we continue to assume that individual beliefs only depend on policy-invariant priors and individ-

---

[16]Additionally, when the expected probability, $E(\pi|H_i^{t-1})$, is high, individuals will show little response to an arrest while they will substantially reduce their expected probability if they avoid an arrest. On the other hand, when the expected probability of an arrest is low, individuals that are arrested will substantially revise their probability of arrest upward, while those that are not will revise their expected probability downward by much less.

ual crime and arrest histories so $g_6 = 0$ (e.g. individuals either do not hear about policy changes or do not believe such announcements), then two policy implications contrast sharply with those predicted by standard models that assume the true probability of arrest is known with certainty. First, an increase in the true probability of arrest (e.g. an increase in the number of police or more lax rules on police searches) will have no immediate effect on crime, but it will have lagged effects. This is true of both permanent and temporary changes. Policy affects are lagged because they only affect crime indirectly through beliefs, which take time to evolve. Each additional arrest that occurs as a result of the increased true probability of arrest will cause the affected criminal to revise his perceived probability of arrest upwards. This increases the likelihood that he refrains from committing further crimes in the future. Even with a direct announcement effect on beliefs, the long-run effects of an increase in the probability of arrest would be greater than the short-run effects. On the other hand, when the probability of arrest is known with certainty, all effects on crime would be immediate and would only continue as long as actual arrest rates remain high.[17]

Second, changes in the true probability of arrest should not only affect the level of crime, but they should also affect the age-crime profile as criminals slowly learn about any changes through experience. To the extent that initial criminal decisions only depend on prior beliefs and tastes, there will be no impact of an increase in the true probability of arrest on the initial crime rate of a cohort. But, subsequent crime rates will decline as more and more individuals experience an arrest. Overall, crime should decline more quickly (or rise more slowly) with age. With direct announcement effects on initial beliefs, crime rates would also decline among youth, offsetting some of the learning effect. This learning effect is entirely absent in standard models with fully-informed agents.

Summarizing, this framework suggests that incorporating beliefs about the likelihood of arrest in a criminal choice model can lead to interesting dynamic responses to changes in the probability of arrest that are frequently ignored. It explains why criminals may be optimistic about their chances of evading arrest when non-criminals are pessimistic. It also suggests that the average perceived probability of arrest is likely to be greater than official arrest rates even when beliefs are unbiased. The importance of these effects will depend on the information acquired by individuals as well as the process by which they update their beliefs. In the following sections, we empirically examine these issues.

---

[17]The criminal justice literature commonly refers to two distinct types of deterrence: *general* and *specific*. *General deterrence* refers to the effects of criminal justice policy through general policy announcements or overall arrest probabilities, while *specific deterrence* refers to deterrence achieved through an individual's own interaction with the justice system. The latter is emphasized here.

3 Crime and Perceptions

Crime and Beliefs in the NLSY97

The NLSY97 contains a sample of 9,022 individuals (4,621 males) ages 12-16 in 1997. This study uses four years of panel data covering the years 1997-2000. Information relevant to this study includes data on family background, individual achievement test scores, neighborhood characteristics, criminal behavior, and perceptions about the probability of arrest and various punishments for auto theft.[18]

The extent of criminal activity among young males in the NLSY97 is shown in Table 1. About 5.5% of young males report committing a theft of over $50 in any given year, with blacks reporting the most involvement and whites the least. Slightly more than 1% of the sample reports committing auto theft. Approximately 8% of all young males report an arrest for some offense in any year, and only 1.7% report an arrest for theft.

Unfortunately, the data do not allow us to determine what category or type of theft for which an arrest was made. To the extent that most arrests occur for thefts of something worth more than $50, we can approximate the arrest rate for theft by race/ethnicity. Between 0.28 (hispanics) and 0.36 (blacks) individuals report an arrest (for theft) for every individual who reports having stolen something worth more than $50. A better measure for an arrest rate is given at the bottom of the table, which reports the total number of arrests for theft per reported theft of more than $50. These rates range from 0.06 for hispanics to 0.12 for blacks. According to these figures, about one out of every ten thefts of greater than $50 results in an arrest, and there is little difference in arrest rates between whites and minorities. A number of caveats should be noted. First, some individuals may be arrested even though they have not committed a theft – this would bias arrest rates upward. Second, some arrests may be for thefts of less than $50 in value, again biasing these estimates upward. Third, both arrests and crimes are self-reported, both of which may be under-reported. To the extent that individuals under-report crimes more than arrests, these estimates will be biased upward. Unless arrests are substantially under-reported compared to actual thefts of greater than $50, these arrest rates should over-estimate true arrest probabilities among those choosing to steal.

While these rates are substantially lower than official clearance rates[19] for burglary, larceny-theft, and motor-vehicle theft, they accurately reflect official arrest rates for theft after adjusting for non-

---

[18]Specifically, the survey asks: "What is the percent chance you would be arrested if you stole a car?" It also asks three separate questions about the outcome of arrest: "Suppose you were arrested for stealing a car, what is the percent chance that you would [be released by the police without charges or dismissed at court, pay a fine and be released, serve time in jail]?"

[19]An offense is 'cleared by arrest' when at least one person is: (1) arrested; (2) charged with the commission of the offense; and (3) turned over to the court for prosecution.

reporting (to the police) by victims. Adjusted arrest rates for theft are lowest for the general larceny-theft category (5.4%), slightly higher for burglary (7.6%), and highest for motor vehicle theft (10.0%).[20] Thus, arrest rates for theft among youth surveyed by the NLSY97 closely correspond to official nationwide arrest rates.

Figure 1 shows the kernel density estimated (using a biweight kernel with a bandwidth of 5) distribution of the perceived probability of arrest for auto theft among young males in the NLSY97. Most youth report much higher perceived probabilities of arrest than is reflected in national arrest rates or in the actual arrest rates for thefts committed by this sample. The figure shows strong focal points at probabilities of 0, 0.5, 0.75, 0.9, and 1.

Young males from all racial and ethnic backgrounds tend to report a relatively high probability of arrest as shown in Table 2. While most previous research has shown that official arrest rates do not vary across races (Tonry, 1995), popular discussion might cause one to think that minorities believe they are more likely to face arrest and serious punishment. This does not appear to be the case here.[21] Panel (A) of the table shows that both young black (52%) and hispanic (54%) males tend to have significantly *lower* perceived probabilities of arrest for auto theft than the average young white male (64%).

The fact that perceived probabilities of arrest are substantially higher than true arrest rates does not necessarily imply that individuals over-estimate their own probability of arrest. As noted earlier, individuals that engage in crime may face substantially lower arrest probabilities than those who do not. While this can explain some of the gap between perceptions and actual arrest rates, even young males engaged in crime report fairly high probabilities of arrest.[22] Panel (B) of Table 2 reveals probabilities for young males who reported stealing something worth more than $50 in the previous year; panel (C) shows perceptions for young males who committed auto theft; and panel (D) calculates average perceived probabilities using the number of thefts of over $50 committed in the last year by each individual to weight the observations. Panel (D) accounts for the possibility that individuals who commit the most crime also hold the lowest perceived probabilities of arrest. If each individual's perceived probability is correct, the weighted average of all perceived probabilities for arrest in panel (D) should equal the sample arrest rate.

---

[20]Arrests, offenses known to the police, and clearance rates are taken from the FBI's Uniform Crime Reports, while reporting rates to the police are given by the Bureau of U.S. Department of Justice, Criminal Victimization in the United States.

[21]From a different perspective, police may discriminate against minorities by failing to pursue perpetrators who victimize them. Since most criminals victimize others like them, this would result in lower real and perceived arrest rates among minorities.

[22]In summarizing a number of studies on perceptions in various contexts, Viscusi (1998) reports that individuals tend to overestimate the risk of low probability events, which is consistent with these findings.

Among teenage males who have stolen something worth more than $50, whites believe that their probability of facing arrest is about 10 percentage points higher than hispanics or blacks. Among auto thieves, the gap between whites and the two minorities drops to around 7 percentage points. Weighting beliefs by the number of thefts reduces the gap further by one percentage point. There is little evidence to support the proposition that young blacks and hispanics feel discriminated against in terms of facing higher arrest rates for auto theft.

In general, teenage males that are more involved in crime tend to predict better chances of evading arrest. As discussed in the previous section, these differences in beliefs can be attributed to at least two potential factors: (1) individuals who hold optimistic views about their chances of success (perhaps, because they have successfully avoided arrest in the past) should be more likely to commit crime, and (2) individuals who are better at evading arrest (and truly face lower probabilities of arrest and punishment) can be expected to commit crime at higher rates. It is also the case that individuals not engaged in crime have little incentive to figure out the true probability while those engaged in crime should have more accurate views since such information is crucial for their 'work'; however, there is little reason to expect that this should bias beliefs in one direction or the other. Given the first factor, it is surprising that even those engaged in auto theft report an average expected arrest rate of 40-50%.

An obvious explanation for the discrepancy in beliefs and true arrest rates is that individuals misinterpret the question.[23] Rather than reporting an arrest rate, individuals may respond by reporting the probability that someone who engages in auto theft (perhaps repeatedly) will ever be arrested for that crime. Indeed, this measure for an 'arrest rate' (dividing the total number of individuals arrested for theft by the number of individuals stealing something worth more than $50) is much higher (30% for the entire sample) as seen in Table 1. Alternatively, individuals may report the probability of arrest for stealing a representative (or random) car, while they only choose to steal cars that offer a substantially lower probability of arrest. In this case, reported arrest probabilities would be greater than the official arrest rate. It is impossible to know for sure how people interpret and answer these questions. To the extent that these measures of beliefs change in response to new information and affect behavior in economically interesting ways, it seems likely that they contain important (if noisy) information about true beliefs. Ultimately, this is an empirical question, which we explore in detail below.

Beliefs about the probability of arrest depend not only on subjective perceptions about actual enforcement variables (e.g. number of police) but also on an individual's own perceived ability to evade

---

[23]Examining responses to a variety of questions about the probability of different events occurring in the near future, Walker (2000) finds little evidence that NLSY97 youth are unable to grasp the concept of probability.

detection. In studying variation in beliefs about the likelihood of arrest, it is, therefore, important to consider individual characteristics which might be correlated with criminal abilities as well as those which may affect opinions about the quality of law enforcement. Table 3 uses ordinary least squares (OLS) regression to examine the importance of county-level arrest rates, individual characteristics, family background, and geographic variables in explaining the perceived probability of arrest for auto theft. While the reported results are based on the entire sample of NLSY97 respondents, the results are very similar when restricted to those reporting a theft of something worth more than $50 sometime in the previous year. Column (i) examines the relationship between county arrest rates for motor vehicle theft[24] and the perceived probability of arrest. The estimates suggest a positive correlation with a coefficient of 0.13. Column (ii) adds demographic indicators for age and race. The coefficient on local arrest rates drops by half, suggesting that much of the correlation between beliefs and official arrest rates is due to locational differences in demographics that are correlated with beliefs. Column (iii) adds an indicator for current residence in a Metropolitan Statistical Area (MSA). The effects of county arrest rates decline further, but MSA status is statistically important. Young males living in an MSA believe they are less likely to be arrested, consistent with lower official arrest rates in urban communities. To the extent that most of the true variation in arrest rates across communities depends on metropolitan status and the demographic characteristics of a neighborhood, it is not surprising that the correlation between beliefs and official county arrest rates, which are undoubtedly measured with error, disappears after controlling for these factors.

Given the theory discussed in Section 2, one might expect that older individuals are better informed about the true arrest rate than are younger respondents. However, the results from including interactions between MSA status and age as well as county arrest rates and age in the regressions of Table 3 do not support this conclusion. Coefficient estimates for these interactions are always insignificantly different from zero. On average, beliefs do not more accurately reflect official arrest rates among older individuals.

Column (iv) of Table 3 adds detailed family background measures (specifically, low current family income, whether the respondent lived with both his natural parents in 1997, whether his mother was a teenager at birth) and math achievement test scores.[25] This has little effect on the estimates already

---

[24]County arrest rates are computed from the ratio of arrests per person divided by crimes per person in each county from the following source: U.S. Dept. of Justice, Federal Bureau of Investigation. UNIFORM CRIME REPORTING PROGRAM DATA [UNITED STATES]: COUNTY-LEVEL DETAILED ARREST AND OFFENSE DATA, 1997-2000 [Computer file]. Inter-univerisity Consortium for Political and Social Research, Ann Arbor, MI.

[25]Peabody Individual Achievement Test (PIAT) scores for math are only observed for individuals with less than 10 years of schooling–nearly everyone age 16 in 1997. To maintain the representativeness of the sample, all individuals age 16 in 1997 are dropped from regressions including PIAT scores, making the sample representative of males ages 12-15 in 1997. The large decline in sample size associated with specification (iv) is primarily due to the inclusion of PIAT scores

13

discussed. Young black and hispanic males report a lower probability of arrest than white males even after controlling for age, local arrest rates, residence in a MSA, and other family background measures. However, racial differences are considerably smaller than their unconditional counterparts shown in panel (A) of Table 2. Perhaps surprisingly, family background has little affect on reported beliefs about the probability of arrest. Other than race/ethnicity, only the effects of Peabody Individual Achievement Test (PIAT) scores for math are statistically significant. In contrast to an 'ability to evade' arrest hypothesis, a 10 percentage point higher math PIAT score implies a 1.2 percentage point higher perceived chance of arrest.

The considerable variation in beliefs is not well explained by these rich measures of family background, geographic location, local arrest rates, age, race, and ability – the R^2 statistics for these regressions are no greater than 0.03. Yet, perceptions are fairly stable over time for many respondents as seen in Figure 2, which shows the distribution of changes in the perceived probability of arrest from one year to the next (using kernel density estimation with a biweight kernel and bandwidth of 5). More than 25% of respondents do not change their beliefs about the probability of arrest between any two years. The correlation in perceptions between years is roughly 0.32.

Crime and Beliefs in the NYS

The NYS contains a random sample of 1,725 individuals (918 males) ages 11-17 in 1976. Respondents were surveyed annually from 1976-1980, then again in 1983 and 1986. This paper focuses on the perceptions and criminal behavior of men as reported in the 1983 and 1986 surveys (earlier surveys do not contain information about perceptions of the criminal justice system).[26] Data regarding family background and some neighborhood characteristics are available.

Table 4 reports the extent of selected criminal activities and arrest records from 1984 to 1986. Since most individuals are in their early twenties during these years, criminal participation is much lower than for the younger sample in the NLSY97. Yet, 18% still report stealing something worth less than \$5 over this three-year period, and 9% report physically attacking someone. Substantially fewer individuals engage in more serious forms of theft. Nearly 12% report an arrest over the three-year span, although many of those arrests are for minor crimes. Only 1.9 percent are arrested for a property

---

and family income, both of which are missing for a sizeable fraction of the sample.

[26]Surveys for 1983 and 1986 actually took place early in 1984 and 1987, respectively. Perceptions questions, therefore, refer to beliefs at the beginning of 1984 and 1987. Criminal participation (and most other) questions explicitly ask about the calendar years 1983 and 1986, however. Additionally, the survey taken in early 1987 also asked retrospective questions about criminal participation in 1984 and 1985. In many cases, categorical measures rather than the actual number of crimes committed in a year are reported (especially for 1984 and 1985). In these cases, the number of crimes committed was imputed from the average number of crimes committed among those in that category who reported the actual number of crimes (based on all survey years).

or violent crime.[27]

Sample arrest rates can be calculated from the information on criminal behavior and arrests. When dividing the number of arrests for property crimes by the total number of break-ins and thefts greater than $50 reported in 1983 and 1986, average arrests per property crime are slightly under 5%. A similar arrest rate is obtained for violent crime when dividing the number of arrests for violent crime by the reported number of times individuals used force to obtain something or attacked someone. These arrest rates are less than official arrest rates in the U.S. population adjusted for non-reporting to the police, especially for violent crimes. (For example, 1986 arrest rates for larceny-theft were 5.5%, burglary, 7.4%, and assault, 20.4%.) However, both the number of crimes and number of arrests in this sample are quite small. Furthermore, the denominators are likely to be inflated due to duplication in reporting of crimes (e.g. some break-ins may also be reported as thefts by respondents).

Individuals were asked to report the probability (in increments of 0.1) that they would be arrested if they were to commit various crimes.[28] The distribution of reported probabilities of arrest in the NYS is shown in Figure 3. Table 5 reports average perceived probabilities of arrest in the NYS for four crimes: stealing something worth $5 or less, stealing something worth more than $50, breaking into a building or vehicle, and attacking someone to hurt or kill them. As with teenage boys in the NLSY97, perceived arrest rates are higher than official arrest rates in the U.S. But, the ranking of crimes by perceived arrest probability from most to least likely corresponds to the ranking of actual arrest rates across crime types. Interestingly, black and hispanic men in the NYS report higher perceived arrest probabilities for property crimes than do white men, in sharp contrast to the NLSY97 findings. However, the differences by race are small for all but petty theft.[29]

Table 6 examines whether perceptions vary across criminals and non-criminals. Specifically, the first column reports perceived probabilities for those who did not commit the crime in question, while the second column reports perceived probabilities for those who did. The final column weights perceived probabilities by the number of times an individual reported committing that type of crime. As with the teenage boys in the NLSY97, those committing any particular crime tend to believe their

---

[27]Arrests for property crimes include various forms of theft, evading payment, burglary, breaking and entering, and dealing in stolen goods. Arrests for violent crimes include assault, robbery, and harassment. Other arrests include crimes such as prostitution, vagrancy, panhandling, etc.

[28]Specifically, the survey asks five distinct questions:"Suppose YOU were to [steal something worth $5 or less, steal something worth more than $50, break into a building or vehicle to steal something or just to look around, use force (strongarm methods) to get money or things from other people, attack someone with the idea of seriously hurting or killing him/her]. What are the chances you would be ticketed/arrested?"

[29]Unfortunately, it is impossible to determine whether differences across the NYS and NLSY97 sample are due to differences in time period (mid-1980s vs. late 1990s), differences in the types of crimes studied, or differences in respondents' age (early to mid-teens vs. mid-twenties). Racial differences in beliefs do not appear to differ dramatically by age, suggesting that the latter reason may not be too important.

chance of arrest for that crime is lower than those not engaging in that type of crime. Weighting beliefs by the number of crimes lowers perceived probabilities even more for all crimes except petty theft. Regardless of the sample, perceived probabilities of arrest are high compared to average arrest rates in the U.S.

The effects of age, race, family background, neighborhood conditions, and urban status on perceptions among young men are estimated using OLS and reported in Table 7. (Ordered probits produce similar conclusions.) Even after controlling for other background characteristics, blacks hold a significantly higher perceived probability of arrest than whites for petty theft, but not for other crimes. Men who grew up in intact families and have more educated mothers or fathers think that their likelihood of arrest is lower on average, although the differences are quite small and generally statistically insignificant. Consistent with official arrest patterns, men in rural areas hold higher perceived probabilities of arrest than those in urban communities.[30]

The 'broken windows' theory of Kelling and Wilson (1982) assumes that local neighborhood conditions affect individual perceptions about the likelihood of arrest and/or punishment and that those perceptions, in part, determine criminal behavior. The small and insignificant coefficients on neighborhood crime and disarray fail to support this theory. Instead, the estimates suggest that young men living in neighborhoods in which crime and 'broken windows' are a problem do not view their chances of arrest any differently from those living in cleaner and more orderly environments. We re-examine this issue below.

As in the NLSY97, the substantial heterogeneity in beliefs is not well explained by background and neighborhood characteristics. Perceptions appear to be largely idiosyncratic and difficult to explain; yet, they are also stable for many respondents. Figure 4 shows the distribution of changes in beliefs from 1983 to 1986 for the sample. For each crime, about 20% of the young men do not change their perceived probability of arrest, while about 60% change their perceived probability by twenty percent or less over three years. We would expect many individuals with no involvement in crime to maintain fairly stable beliefs over time. Still, we are particularly interested in those who do change their beliefs and why they change them. Furthermore, we would like to know how changes in beliefs affect subsequent criminal activity. The remaining sections of this paper empirically analyze these issues.

---

[30]State and county of residence are unknown in the NYS, so perceptions cannot be compared with local official arrest rates as in the NLSY97.

4   Information-Based Belief Updating

This section empirically studies factors that may cause individuals to change their beliefs about the probability of arrest. In the NLSY and NYS, we assume the reported measure of the perceived probability of arrest, $p_{i,t}$, relates to $E(\pi|H_i^t)$ as defined in Section 2. The simple Bayesian structure above (see equation 3) suggests estimating the relationship between changes in perceptions and changes in environmental factors $Z_{i,t}$ (e.g. local arrest rates, metropolitan status, neighborhood characteristics, etc.), new arrests $A_{i,t}$ and crimes committed $c_{i,t}$ (both taking place between period $t-1$ and t) by the respondent as well as his siblings, $\tilde{A}_{i,t}$ and $\tilde{c}_{i,t}$:

$$\Delta p_{i,t} = \Delta Z_{i,t}\gamma + \phi A_{i,t} + \lambda c_{i,t} + \phi_s \tilde{A}_{i,t} + \lambda_s \tilde{c}_{i,t} + u_{i,t}. \tag{4}$$

A more general structure of updating can also be estimated as follows:

$$p_{i,t} = Z_{i,t}\gamma + \theta p_{i,t-1} + \phi A_{i,t} + \lambda c_{i,t} + \phi_s \tilde{A}_{i,t} + \lambda_s \tilde{c}_{i,t} + \varepsilon_{i,t}, \tag{5}$$

which relaxes the implicit assumption of the Bayesian model that $\theta = 1$. In this case, some $Z_{i,t}$ variables (e.g. ability, race, family background, etc.) may be time invariant. While these variables are "differenced out" in equation (4), they can play an important role in this more general updating equation. With $|\theta| < 1$ and $Z_{i,t} = Z_i^*$ constant, beliefs would eventually converge to a steady state,

$$p_i^*(Z_i^*) = Z_i^* \frac{\gamma}{1-\theta}, \tag{6}$$

if the individual and his siblings stopped committing crime and were never arrested again (ignoring changes in $\varepsilon_{i,t}$). Consequently, an individual's permanent Z_i characteristics determine his baseline (i.e. no crime/no arrest steady state) level of beliefs. Changes in time-varying $Z_{i,t}$ factors will affect long-run beliefs, perhaps through information gathered from others or from observing changes in local conditions. For example, moving to a new city or neighborhood may cause an individual to gradually shift his beliefs toward thinking the probability of arrest is higher or lower than previously thought, even if he does not engage in crime or face an arrest.

Equation (5) can be re-written as

$$p_{i,t} = (1-\theta)p_i^*(Z_{i,t-1}) + \theta p_{i,t-1} + \Delta Z_{i,t}\gamma + \phi A_{i,t} + \lambda c_{i,t} + \phi_s \tilde{A}_{i,t} + \lambda_s \tilde{c}_{i,t} + \varepsilon_{i,t},$$

which shows how θ determines the rate at which beliefs move toward their baseline level. A value of θ near zero implies that beliefs quickly converge to their steady state level given any new information. This implies that any observed changes affecting beliefs (e.g. an arrest or non-arrest) have short-lived

17

effects on beliefs. This would be the case if individuals continually receive strong signals (unobserved by the econometrician) that their probability of arrest is p_i^*. Or, it may simply imply that individuals have short memories and quickly return to some baseline belief about their own probability of arrest.

Let T represent the final period of observation in the data. With at least three periods of data (i.e. $T \geq 3$), we can allow for unobserved individual fixed effects: $\varepsilon_{i,t} = \mu_i + \nu_{i,t}$ assuming that

$$E(\nu_{i,t}|p_i^{t-1}, Z_i^T, A_i^t, c_i^t, \tilde{A}_i^t, \tilde{c}_i^t) = 0 \qquad \forall t = 2, ..., T, \tag{7}$$

where $x_i^t = (x_{i1}, x_{i2}, ..., x_{it})$ represents the entire history through period t for any variable x. This assumes that only Z variables are 'strictly exogenous', while all other regressors are 'pre-determined'. As Arellano and Honore (2001) show, the assumption of equation (7) is fairly weak in that it does not rule out feedback effects of lagged dependent variables or disturbances on current and future values of the 'pre-determined' variables. That is, by conditioning on c_i^t and A_i^t (rather than c_i^T and A_i^T as with traditional fixed or random effects strategies) in equation (7), we explicitly allow for the fact that $c_{i,t+1}$ and $A_{i,t+1}$ (as well as all future values of crime and arrests) may depend on the current disturbance $\nu_{i,t}$.[31] This is important, because we expect that subsequent crime depends on current beliefs about the probability of arrest – an issue we examine more closely in the following sections.

For

$$\Delta \nu_{i,t} = \Delta p_{i,t} - [\theta \Delta p_{i,t-1} + \Delta Z_{i,t}\gamma + \phi \Delta A_{i,t} + \lambda \Delta c_{i,t} + \phi_s \Delta \tilde{A}_{i,t} + \Delta \lambda_s \tilde{c}_{i,t}],$$

we estimate the general model with fixed effects using GMM and the following moment conditions:

$$
\begin{aligned}
E[p_{i,t-2}\Delta \nu_{i,t}] &= 0, \\
E[\Delta Z_{i,t}\Delta \nu_{i,t}] &= 0, \\
E[A_{i,t-1}\Delta \nu_{i,t}] &= 0, \qquad E[\tilde{A}_{i,t-1}\Delta \nu_{i,t}] = 0, \\
E[c_{i,t-1}\Delta \nu_{i,t}] &= 0, \qquad E[\tilde{c}_{i,t-1}\Delta \nu_{i,t}] = 0.
\end{aligned}
$$

These moment restrictions are applied for t equal to 1999 and 2000 in the NLSY97.[32] This method cannot be used with the NYS data, since only two periods of perceptions data are reported.

Table 8 reports estimates related to belief updating in the NLSY97 for the following: (A) OLS regression for the difference equation (4); (B) OLS regression for the quasi-difference equation (5); and (C) GMM for the quasi-difference equation (5) accounting for individual fixed effects. Each panel

---

[31]Equation (7) also allows for the possibility that $\tilde{c}_{i,t+1}$, $\tilde{A}_{i,t+1}$, and all future sibling criminal behavior and arrest outcomes are affected by the disturbance term.

[32]It is possible to use additional lags as instruments. However, conditional on using one-period lags, little is gained from adding more (the qualitative results are unchanged when more lags are added), and concerns about weak instruments become heightened.

reports two specifications. The first includes indicators for whether the individual or his male siblings committed crime or were arrested for theft between survey dates. The second includes indicators for the actual number of times individuals and their siblings committed crimes and were arrested.[33] Measures based on sibling crimes and arrests refer to male siblings who are also in the main NLSY97 sample. As a result, their ages are always within a few years of the respondent.[34] In general, all specifications show strong evidence of belief updating in response to the respondent's own criminal history. Individuals who reported stealing something worth more than $50 or selling drugs were likely to report a lower perceived probability of arrest (conditional on prior beliefs and the arrest outcome) in the next survey year. The effect of at least one of these crimes is statistically significant at the 10% level (most at the 5% level) in every specification. Those arrested for a theft increased their perceived probability of arrest (significantly so in most specifications). As shown at the bottom of Table 8, joint tests for whether the coefficients on all individual crime and arrest variables equal zero are rejected for all but the final specification.[35]

The effects of sibling crime and arrests are less precisely estimated given that only 27% of the respondents have at least one sibling that is also in the NLSY97 sample. Still, a number of coefficient estimates on measures of sibling crime are statistically significant and negative, as expected. The estimated effects of a sibling's arrest are generally not significantly different from zero and are often of the wrong sign. Joint tests for whether the sibling crime and arrest coefficients equal zero are rejected (at the 10% level) in all but the last two specifications. Alternatively, joint tests for whether the coefficients on sibling crimes and arrests equal the corresponding coefficients on respondent crime and arrests cannot be rejected for any but the final specification.

The most noticeable difference between the OLS and GMM estimates of equation (5) is the change in the estimated coefficient on the lagged measure of beliefs. After controlling for unobserved individual fixed effects, the autocorrelation of the perceived probability of arrest (θ) drops from 0.3 to below 0.04, which implies that there is little persistence in the effects of new information on reported beliefs. Instead, unobserved differences in baseline beliefs appear to explain why some individuals hold a high

---

[33]Ideally, we would use measures for the crime of auto theft and arrests for auto theft in our updating specifications, but auto thefts are rarely observed in the NLSY97 data and arrests for auto theft cannot be identified. Assuming beliefs about the probability of arrest are positively correlated across crimes – in the NYS, correlations in beliefs about the probability of arrest across crimes range from a low of 0.33 between attack and minor thefts to a high of 0.69 between minor and major thefts – we should expect beliefs about the probability of arrest for auto theft to change in response to other crimes and arrests.

[34]Though not reported, specifications controlling for the number of siblings present in the household show nearly identical results – there is little effect of household size on beliefs. Also, estimates are qualitatively similar when using a restricted sample of individuals who reported a theft of greater than $50 in at least one of the previous two years.

[35]All tests for the first differences and quasi-first difference specifications are F-tests, while tests for the GMM specifications are Wald tests.

perceived probability of arrest year after year, while others believe the probability of arrest is much lower. As we see in Section 6, measurement error in reported beliefs may contribute substantially to the low estimate of θ obtained here.[36]

A more limited analysis is performed using young men in the NYS. Because the NYS only records beliefs for two periods, we cannot estimate the quasi-difference model with fixed effects using the GMM procedure outlined above. In Table 9, we report estimates of equation (5) using 1983 and 1986 measures of beliefs, accounting for crimes and arrests that take place between the two surveys (estimates of the first-difference specification are quite similar).[37] Again, we employ two specifications for each type of crime studied. The first includes an indicator variable for whether the individual committed the crime under study (e.g. in column 1, the indicator is one if the individual reported stealing something worth less than \$5 and zero otherwise) or was arrested for a violent or property crime during the 1984-86 period. The second includes measures of the number of crimes committed and arrests over that period.

As in the NLSY97, these men report lower perceived probabilities of arrest for all crime categories at the end of 1986 if they engaged in that type of crime in 1984-86 (three of the four estimates are statistically significant). Coefficients on arrest are always positive and quantitatively large, but they are only significantly different from zero for break-ins. Joint tests of whether the coefficients on crime and arrests are zero are rejected in nearly all of the columns. While the model above suggests that the net effect of a crime and arrest should be negative, the specifications with only an indicator for committing a crime and getting arrested generally reveal a coefficient on the indicator for crime that is larger in absolute value than the coefficient on the arrest indicator. This is because those who are arrested commit more than one crime, on average. Looking at the specifications controlling for the number of crimes and arrests, we always observe a larger coefficient on an arrest than on a crime. Thus, the net effect of a crime and arrest on the perceived probability of arrest is always positive.

It is interesting to note that using measures for any arrest rather than arrests for more serious property and violent crimes (as in Table 9) generally produces smaller and insignificant effects on beliefs (except for small thefts). This suggests that police attempts to crack down on vagrancy, public intoxication, and other petty crimes are not likely to influence beliefs about the probability of arrest for more serious crimes in any significant way.

---

[36]To address concerns about measurement error in beliefs, one could substitute the moment condition $E(p_{i,t-3}\Delta\nu_{i,t}) = 0$ for $E(p_{i,t-2}\Delta\nu_{i,t}) = 0$. Unfortunately, this reduces the sample size by half (since we can no longer use the moments for $t = 1999$) and generates extremely noisy estimates.

[37]Specifications also control for age, race/ethnicity, whether the individual's parents earned less than \$10,000 in 1976, and whether the individual lived with both natural parents in 1976.

Table 9 also reports coefficient estimates on central city status and rural residential status. The effects of these measures are insignificant once we control for crime and arrest histories and previous measures of beliefs (in contrast to those in Table 7). Measures of neighborhood lawlessness and disarray also have no significant effect on beliefs. Again, we find no evidence to support the 'broken windows' theory of Wilson and Kelling (1982).

While we do not have sibling measures of beliefs or crime in the NYS, one might wonder whether information about the arrests of other random criminals has any affect on beliefs as Sah's (1991) theory would suggest. To that end, we examine whether individuals who are victimized by a crime alter their beliefs in response. The underlying assumption here is that victims are likely to learn whether or not the perpetrator is ever arrested. In a world in which all individuals face identical probabilities of arrest, information as a victim should be as useful as information as a perpetrator. Unfortunately, the data do not record whether someone was arrested for the crime, but it is reasonable to assume that no arrest was made in most cases given the low official arrest rates reported earlier. Then, we should expect, on average, that individuals will adjust the probability of arrest downward after a victimization.[38] The estimated coefficients on victimization are small and statistically insignificant for all crimes in Table 9. While not an ideal test of information from the arrest outcomes of others, these estimates suggest that individuals put little weight on the information provided by the crime and arrest outcomes of random criminals. Arrest probabilities may be too individual-specific to make such information useful in determining one's own arrest probability.

Altogether, the NLSY97 and NYS estimates strongly suggest patterns consistent with belief updating among respondents that is based on their own history of interaction with the criminal justice system. When young men participate in crime, they tend to lower their perceived probability of arrest if they evade arrest. If arrested, they raise their perceived probability. One could potentially explain the first finding by arguing that individuals chose to commit crime between survey dates, because they had already (for some exogenous reason) lowered their perceived probabilities (but were unable to report those new perceptions until surveyed the second time). Or, those engaged in crime could have gained experience at crime, lowering their true (and perceived) arrest probability. However, such scenarios cannot explain why those arrested between sample dates maintain higher perceived probabilities of arrest at the time of the second interview. Only an information-based model of belief updating

---

[38]Of course, if those who observe an arrest adjust their beliefs upward much more than those who do not observe an arrest adjust their beliefs down, this need not be the case. Given that official arrest rates range from 5-20% for the crimes under study, those observing an arrest would have to adjust their beliefs upwards by 5 to 20 times as much as those not observing an arrest adjust theirs downward for the effects to cancel. This is unlikely, given that the estimated negative coefficients on (own and sibling) crime measures remain significantly negative when leaving out arrest outcomes in updating regressions (i.e. Tables 8 and 9).

can readily explain both findings. The fact that sibling criminal histories affect beliefs in a similar way to an individual's own history strengthens this conclusion; although, the results for sibling arrest histories do not. The importance of an individual's own criminal and arrest history in determining beliefs strongly supports the very simple model outlined in Section 2 and developed further in Lochner (2004). Sah's (1991) theory also finds some support in that the criminal history of an individual's siblings affects beliefs about the probability of arrest, but information about arrest outcomes from a sibling or other random persons does not seem to be important. Taken together, these results suggest a limited role for learning from others.

5   The Influence of Perceptions on Criminal Behavior

Given the considerable variation in perceptions about the probability of arrest, it is natural to question whether individuals act differently based on stated beliefs. And, do they behave differently in periods when they report a high perceived probability of arrest than when they report a low probability? Rational choice theory suggests that (holding all else constant), individuals facing a higher probability of arrest and/or punishment should commit less crime. We examine this relationship in both the NLSY97 and NYS.

In this section, we examine the relationship between beliefs and crime using a linear probability model:

$$c_{it} = W_{i,t}A + Bp_{i,t-1} + v_{i,t}, \tag{8}$$

where $W_{i,t}$ represents observed individual characteristics that may affect the costs of or returns to crime and $v_{i,t}$ are iid shocks to current criminal returns/costs. As before, $c_{i,t}$ is an indicator for whether an individual committed a crime between survey dates $t-1$ and t, and $p_{i,t}$ reflects beliefs about the probability of arrest as of survey date t. Since we explore the effects of elicited perceptions in the previous survey year on retrospective crime reported in the current survey, we are left with three years of belief-crime data in the NLSY97 and a single cross-section in the NYS.

In the NLSY97, we can control for age, race, ethnicity, whether or not the youth lived with both his natural parents, whether or not the youth's mother was a teenager at birth, math PIAT scores, and official county-level arrest rates, in addition to the perceived probability of arrest for auto theft. Since we only observe the perceived probability of arrest for the crime of auto theft, we restrict our attention to its estimated effects on major thefts (thefts of something worth more than $50) and on auto theft. Assuming $v_{i,t}$ shocks are orthogonal to $W_{i,t}$ and beliefs about the probability of arrest, equation (8) can be estimated by OLS. These results are presented in columns 1 and 2 of Table 10 and

22

suggest that the perceived probability of arrest for auto theft significantly reduces both auto theft and theft more generally. To put the estimates into perspective, they imply that a ten percentage point increase in the perceived probability of arrest would reduce major thefts by nearly 4% and auto theft by 7%.[39] It is interesting that official county-level arrest rates are estimated to have no impact on auto theft and a perverse effect on major thefts.

A similar approach can be taken with the NYS, examining the effects of beliefs in 1983 on crime committed over the 1984-86 period. Here, we have crime-specific beliefs, which we include in the regressions along with controls for age, race/ethnicity, whether the respondent lived with both natural parents in 1976, whether parental income was below $10,000 in 1976, and rural and central city residential status. All estimated coefficients on the crime-specific perceived probability of arrest are negative, supporting the case for deterrence. The estimates suggest that a 10 percentage point increase in the perceived probability of arrest would reduce all four crimes studied (minor thefts of something worth less than $5, major thefts of something worth more than $50, breaking into a building or vehicle, and attacks) by 7-12%. Estimated effects on minor thefts and attack are statistically significant at the 5% level.

It is possible that perceptions of arrest probabilities are correlated with more general unobserved preferences for risk and crime. Then, these estimated relationships would capture both the deterrent effect of a higher perceived probability of arrest and the influence of unobserved tastes for crime that are correlated with beliefs about the likelihood of arrest. Two sets of results suggest that this may not be an important concern. First, estimates using NLSY97 data (and identical specifications to those above) suggest that the effects of beliefs about the probability of arrest for auto theft are small and statistically insignificant for minor delinquent activities like smoking and drinking.[40] This is reassuring, since one would not expect the probability of arrest for auto theft to be very highly correlated with the probability that these young males will be punished for smoking or drinking. Second, the estimated effects of beliefs on crime in the NYS remain even after controlling for parental and peer approval levels for crime as well as the individual's own moral attitudes towards crime.[41]

Still, it is worthwhile exploring whether unobserved heterogeneity affects these estimates. With the NLSY97 data, it is possible to incorporate permanent unobserved differences, assuming $v_{i,t} =$

---

[39]Estimates using a logit or probit specification yield very similar results. Specifications which also include the perceived conditional probability of going to jail if arrested yield similar estimates for the impact of arrest probabilities. Specifications which control for the perceived unconditional probability of going to jail (i.e. the interaction of the probability of arrest with the probability of going to jail conditional on arrest) produce qualitatively similar results as well.

[40]The estimates suggest that a ten percentage point increase in the perceived probability of arrest would only reduce drinking by 0.2% and smoking by 0.8%.

[41]Specifically, these specifications control for whether the respondent's parents or peers would disapprove of them stealing something and whether they themselves believe stealing is wrong.

$\xi_i + \omega_{i,t}$ with $\xi_i \perp\!\!\!\perp \omega_{i,t}$.[42] In general, we would expect ξ_i to be correlated with $p_{i,t-1}$, since ξ_i is a determinant of past participation in crime and since past criminal behavior affects current beliefs. By first-differencing equation (8), we can eliminate the unobserved fixed effect. However, $\Delta p_{i,t-1}$ will tend to be correlated with $\Delta\omega_{i,t-1}$ since $\omega_{i,t-1}$ affects $c_{i,t-1}$, which is a determinant of $p_{i,t-1}$. Assuming that $\omega_{i,t}$ is independent of W_i^T, p_i^{t-1}, and all lagged values of crime and arrests for the individual and his siblings, we can use the following instruments for $\Delta p_{i,t-1}$: $p_{i,t-2}$, the number of thefts of something worth more than \$50 by the individual and his siblings reported in survey year $t-2$, and the number of arrests for theft reported by the individual and his siblings in survey year $t-2$. Consistent with the estimates presented in the previous section, these instruments are quite strong in predicting changes in beliefs from period $t-2$ to $t-1$. Results from two stage least squares regression of the first-differenced crime equation are reported in columns 3 and 4 of Table 10.[43] The estimated effect of the perceived probability of arrest for auto theft on committing any major theft and on engaging in auto theft are quite similar to the OLS estimates. The effect on auto theft is statistically significant at the 5% level.

Treating these fixed effect instrumental variable estimates as the deterrent effect of perceived arrest probabilities, it is possible to study the extent to which differences in beliefs are responsible for differences in criminal participation by race or ability. The estimated 7.8 percentage point difference in perceived arrest probabilities between whites and blacks (Table 3, column iv) implies a 6.8% higher participation rate in auto theft by blacks. Hispanics are predicted to have a 7.6% higher participation rate in auto theft than whites due to differences in perceived arrest probabilities. The predicted difference in auto theft participation rates between individuals at the 75th and 25th percentiles in PIAT math scores is 5.2%. These simple comparisons suggest that important variation in criminal participation rates across individuals may be due to differences in information and beliefs.

6 A Structural Estimation of Belief Updating, Crime, and Arrests

In this section, we use the NLSY97 data to estimate a structural model of belief updating, criminal participation, and arrests (like that sketched in Section 2) accounting for permanent unobserved individual heterogeneity. While this approach adds more structure to the belief updating and criminal participation problems described in the previous two sections (most notably, it imposes distributional assumptions on the disturbance terms), it offers a number of important contributions over that analysis. First, it guarantees that estimates of the belief updating and criminal behavior decision are

---

[42]Because we do not observe crime subsequent to both surveys with perception measures in the NYS, we are unable to address unobserved heterogeneity in these data.

[43]We estimate a stacked regression for changes in crime from 1998-99 and 1999-2000 for all individuals. Standard errors are corrected for clustering at the individual level.

internally consistent and, more importantly, facilitates an analysis of counterfactual experiments that fully incorporate the dynamics of belief updating, crime, and arrests. Second, it allows us to estimate how unobserved factors that affect criminal participation (ξ_i) are related to unobserved components determining baseline beliefs about the probability of arrest (μ_i). Third, the introduction of an arrest equation allows us to analyze whether unobserved and observed factors affecting the perceived probability of arrest also determine the actual probability of arrest faced by individuals. Fourth, it enables a more formal treatment of the criminal participation decision that avoids the well-known concerns associated with the linear probability model. Finally, it offers a natural way to estimate the extent of measurement error in beliefs and how that measurement error affects the estimates presented above. More generally, the structural approach provides a form of robustness check on the estimates presented in Sections 4 and 5.

We draw on a simplified version of the belief updating equation (5) with unobserved fixed effects (dropping measures of sibling crime and arrests):

$$p_{i,t} = Z_{i,t}\gamma + \theta p_{i,t-1} + \phi A_{i,t} + \lambda c_{i,t} + \mu_i + \nu_{i,t}, \tag{9}$$

where

$$\nu_{i,t} \sim N(0, \sigma^2) \ iid \ \forall i, t, \qquad \nu_{i,t} \perp\!\!\!\perp \left(c_i^t, A_i^t, p_i^{t-1}, Z_i^T, W_i^T, \mu_i, \xi_i \right).$$

The assumptions on $\nu_{i,t}$ are quite similar to those used in Section 4 (see equation 7) with the added assumption of normality.

We assume that an individual uses the following decision rule for criminal participation:

$$c_{i,t} = \begin{cases} 1 & \text{if } c_{i,t}^* \equiv W_{i,t}A + Bp_{i,t-1} + \xi_i - \omega_{i,t} > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\omega_{i,t}$ is assumed to be iid standard normal and

$$\omega_{i,t} \perp\!\!\!\perp \left(c_i^{t-1}, A_i^{t-1}, p_i^{t-1}, Z_i^T, W_i^T, \mu_i, \xi_i \right).$$

This is the probit analogue to the linear probability model estimated in Section 5, assuming that previous crimes, arrests, and beliefs are predetermined. It is important that these assumptions do not rule out the potential for current $\omega_{i,t}$ shocks to impact future crimes, arrests, and beliefs, since theory suggests that they should.

Finally, we assume that the probability of arrest is given by

$$Pr(A_{i,t} = 1 | c_{i,t}, Z_{i,t}, \mu_i) = \Phi\left(\pi_0 + \pi_1\mu_i + Z_{i,t}\pi_2 + \pi_3(1 - c_{i,t}) \right),$$

where $\Phi(\cdot)$ is the standard normal cumulative density function. The same individual and environmental characteristics that affect beliefs are assumed to affect true arrest rates, including the unobserved characteristic, μ_i. This allows us to estimate whether individuals with a higher baseline perceived probability of arrest for auto theft actually face a higher true probability of arrest for theft. Furthermore, the arrest probability explicitly depends on current criminal participation; although, it need not be zero for those who do not engage in crime.

Under these assumptions, it is straightforward to estimate the parameters of the belief updating equation, crime equation, and arrest equation with panel data on beliefs, crime, and arrests as well as an assumption about initial beliefs. Regarding the latter, one could simply specify an *ad hoc* distribution for initial beliefs conditional on $Z_{i,0}$. Instead, we assume that at age 10 (prior to observation in our data), beliefs are drawn from the stationary no crime/no arrest distribution:

$$p_{i,0} \sim N\left(p^*(Z_{i,0}, \mu_i), (\sigma^*)^2\right),$$

where $p^*(Z_{i,0}, \mu_i) = \frac{Z_{i,0}\gamma + \mu_i}{1-\theta}$ and $\sigma^* = \frac{\sigma}{\sqrt{1-\theta^2}}$. This is consistent with beliefs evolving over time according to equation (9) and an assumption that individuals do not engage in theft or face an arrest for theft prior to age 10. As discussed in the Appendix, this assumption requires that we integrate our likelihood function over all possible crime and arrest histories leading up to the first period of observation for each individual.

Finally, we must specify the distribution for unobserved characteristics. We assume that there is a finite number, J, types for $\Omega = (\mu, \xi)$ and that Ω is independent of (Z^T, W^T) in the population. Given the number of types, this model can be estimated using maximum likelihood. See the Appendix for a discussion of the likelihood function.

Table 11 presents estimation results for the NLSY97 when $J = 4$, $p_{i,t}$ reflects the perceived probability of arrest for auto theft, $c_{i,t}$ is an indicator for any thefts worth more than \$50, and $A_{i,t}$ is an indicator for any arrests for theft. The results presented here assume that there are two unobserved types who never commit crime (i.e. $\xi_i = -\infty$) and two types that have the potential to engage in crime. The latter values of ξ_i are estimated along with the probability of each type. We assume there are two values of μ_i, but do not impose any correlation between μ_i and ξ_i.[44]

The left hand side of Table 11 reports estimates (and standard errors) for the model assuming beliefs are measured without error – consistent with results previously reported in Tables 8 and 10. Estimates of the belief updating equation are similar to those in Table 8. Focusing on the estimated

---

[44]Attempts to introduce additional types did not prove fruitful. In general, either the probability of new types went to zero or new values for μ and ξ converged to those estimated for the current types.

autocorrelation in beliefs (θ), we observe that our estimation procedure does a good job of capturing the low level of persistence in beliefs with only two values for μ. The estimated values for μ suggest that there are substantial differences across persons in their baseline beliefs about the probability of arrest that cannot be accounted for by a rich set of personal, family, and neighborhood characteristics. As expected, participation in crime reduces the perceived probability of arrest while an arrest raises it; although, the latter effect is not statistically significant. Turning to the estimates of the crime equation, we find that both living in an intact family and higher PIAT math scores reduce the propensity to engage in theft. However, we find no effect of the perceived probability of arrest on theft, in contrast to our previous findings in Section 5 and deterrence theory. We return to this issue below. Estimated parameters of the arrest function imply negligible differences in the actual probability of arrest across individuals from different backgrounds. Based on the positive estimate for π_1, the unobserved factor affecting baseline beliefs about the probability of arrest (μ) is positively related to the true arrest probability, but the estimate is not statistically significant. Only PIAT math scores are statistically significant, suggesting that those with better math skills are less likely to be arrested. Interestingly, this is in sharp contrast to the relationship between PIAT scores and the perceived probability of arrest.

While these estimates suggest that changes in beliefs over time do not induce changes in criminal participation, unobserved tastes for crime (ξ) are negatively correlated with baseline beliefs about the perceived probability of arrest (μ) among the two criminal types. To the extent that reported beliefs are measured with error, baseline beliefs may better represent true beliefs driving criminal behavior. Much of the variation in beliefs over time, especially for those not engaged in crime, may simply represent random measurement error. In this case, it would not be surprising if year-to-year changes in reported beliefs did not generate changes in criminal behavior for most respondents. (Measurement error would also explain why reported beliefs are only weakly correlated over time after controlling for permanent unobserved heterogeneity.) Ideally, we would like to link changes in behavior to true changes in beliefs, like those changes that come from changes in one's environment or engagement in crime and the arrest outcomes associated with criminal activity.

To address this issue, we introduce measurement error in $p_{i,t}$ that is additive, normally distributed with mean zero and variance σ_η^2, and independent of all other variables and disturbance terms. When we explicitly allow for measurement error in the probability of arrest for auto theft, the estimates suggest that the variance of the structural shocks to beliefs (σ^2) is negligible. The right hand side of Table 11, therefore, reports estimates assuming $\sigma^2 = 0$ while estimating the variance of measurement

error, σ_η^2. (See the Appendix for details on estimated likelihood equations.) In this case, true changes in beliefs over time come from observable events like a change in residential location, commission of a crime, or an arrest, and not unobserved idiosyncratic shocks, $\nu_{i,t}$. These observable events will drive identification of the autocorrelation in beliefs, θ, and the effect of beliefs on crime, B – two of the more important parameters from a theoretical perspective.

Accounting for measurement error leads to a number of changes in the estimated parameter values, three of which are worth highlighting. First, the estimated autocorrelation in beliefs (θ) increases dramatically from 0.07 to 0.40. This is not surprising, but it changes the dynamic implications of any policy designed to affect individual beliefs. We discuss this further below. Second, the perceived probability of arrest now has a significant negative effect on criminal participation. Again, this is not surprising, since changes in measured beliefs associated with changes in residential status, a theft, or an arrest are likely to be negatively correlated with subsequent criminal activity in a way that changes in reported beliefs due to measurement error are not. Third, the correlation between ξ and μ among non-criminals switches from negative (in the case ignoring measurement error) to positive (in the measurement error specification). However, the unobserved factor affecting baseline beliefs (μ) remains positively correlated with the true arrest probability.

Using the estimates from the measurement error specification in Table 11, we explore the quantitative importance of beliefs in determining criminal activity. In particular, we simulate how average beliefs and criminal behavior evolve over adolescence and early adulthood under a number of different scenarios. In addition to the baseline set of parameters, we explore the effects of (a) raising all μ values by $5 \times (1 - \theta)$, which effectively raises the perceived probability of arrest by about 5 percentage points each year; (b) raising initial beliefs, $p_0(Z_{i,0}, \mu_i)$, by 5 for everyone; and (c) increasing the true probability of arrest by about one-third (specifically, increasing π_0 by 0.2). The effects of these changes on beliefs and criminal participation are shown in Figures 5a and 5b.

Increasing μ such that the perceived probability of arrest increases by about 5 percentage points in every year reduces major theft rates by about 50%. The interaction of belief updating and criminal behavior is evident in the growing effects on crime with age. A higher μ raises the perceived probability in each period. This causes individuals to commit less crime, which affects what they learn about their probability of arrest. Committing less crime each year, they do not revise their beliefs downward as much, on average, as those with a lower value of μ. These two forces work together, so that the perceived probability of arrest declines less and criminal activity increases less with age when μ is increased. The dynamic role of belief updating is even more evident when examining the effects of

28

an increase in initial beliefs. With some persistence in beliefs (i.e. $\theta > 0$), a one-time increase in the perceived probability of arrest has lasting effects that fade over time, as seen in Figure 5a. With θ estimated to be around 0.4, it takes about five years before beliefs return to where they would have been in the absence of a five percentage point shock at age 10. This implies that any effects on crime also take about five years to fade out (Figure 5b). Overall, these results suggest that deterrence based on the perceived probability of arrest is quite strong.[45]

Figures 5a and 5b also show how a change in the true arrest probability affects criminal behavior through individual experiences with an arrest. Recall from Section 2, that an increase in the true arrest probability will cause more individuals to be arrested, which will cause them to raise their perceived probability of arrest and reduce subsequent criminal activity. Given the low baseline arrest rate (around 15% of the persons engaging in theft experience an arrest for theft) and the modest estimate of the effect of arrests on beliefs, even a dramatic increase in the true arrest probability has quite small effects on beliefs and the crime rate through learning and individual experience. As noted in Section 2, the effects are delayed and appear to grow over time, but they never become quantitatively important. These results suggest a fairly limited role for specific deterrence achieved through an individual's own arrest experiences, at least for crimes like theft that have such a low arrest rate.

Differences in criminal behavior across races, ethnicity, or academic ability can arise, because individuals differ in their baseline beliefs about the probability of arrest, their criminal tastes (or costs), or the true probability of arrest. Given the results just discussed, it is not surprising that the latter plays little role in explaining these differences. However, differences due to heterogeneous beliefs and tastes for crime appear to be large and offsetting when comparing whites, blacks, and hispanics. That is, simulations which compare the rates of theft if everyone in the NLSY97 is assumed to be white imply similar crime rates if everyone is assumed to be black. But, if we take the sample of whites and give them the belief parameter values of blacks or hispanics, we predict that these individuals would commit about twice as much crime, with the effect growing slightly with age.[46] The fact that blacks and hispanics have a lower taste for crime (or higher costs) essentially offsets this, such that minorities and whites commit theft at very similar rates. A similar decomposition for math PIAT scores suggests that the effects of perceptions dominate the effects of tastes for crime. According to the estimates

---

[45]It is important to note that the standard error on the estimate for B is sizeable, despite the fact that B is statistically significant at the 5% level. Reducing B by two standard deviations (i.e. setting B at its lowest point in its 95% confidence interval) reduces the effects of the perceived probability of arrest on crime by nearly 70%. Even this conservative estimate leaves an important role for belief-based deterrence.

[46]This experiment is very similar to that of changing μ by $5 \times (1 - \theta) = 3.0$, since the estimated difference in the belief index for whites and blacks is -3.5 (-3.7 for whites and hispanics).

in Table 11 (accounting for measurement error), more able individuals commit less crime on average, because they believe they are more likely to be arrested – not because they have less taste for crime or higher costs of engaging in crime.

7   Conclusions

Empirically, we uncover substantial heterogeneity in beliefs among young males in the National Longitudinal Survey of Youth 1997 Cohort and the National Youth Survey. On average, individual beliefs about the probability of arrest for various crimes are substantially higher than official arrest rates, and those beliefs are fairly stable across time for individuals. Perceived arrest rates are lower, on average, among those actively engaged in crime, which is consistent with standard deterrence theory as well as an information-based model of belief updating. There is little evidence that minority men believe they are more likely to be arrested than do white men, which is consistent with studies suggesting that there is little, if any, discrimination in official arrest rates across race (Tonry, 1995). Less than 5% of the heterogeneity in beliefs can be explained by differences in family background, neighborhood, or cognitive abilities.

Beliefs are correlated with county-level official arrest rates and metropolitan or urban residential status. But, contrary to the 'broken windows' theory developed by Wilson and Kelling (1982), perceptions are not correlated with other neighborhood conditions like general lawlessness, abandoned buildings, or the presence of winos on the streets. Furthermore, perceptions are not significantly affected by one's own criminal victimization, which might provide additional information about the likelihood of arrest. Instead, individual beliefs about one's own probability of arrest are largely idiosyncratic, stable, and unrelated to the local environment. It is difficult to know whether variation in beliefs across individuals reflects actual variation in the true probability of arrest or simple differences in beliefs, but our structural estimates suggest negligible differences in true arrest rates based on most observable and unobservable characteristics.

While beliefs seem to be largely orthogonal to most outside influences, they do respond to an individual's own experiences with crime and police. Individuals who engage in crime while avoiding arrest tend to reduce their perceived probability of arrest; those who are arrested raise their perceived probability. Interestingly, beliefs respond similarly to changes in the criminal history of their siblings, but they do not appear to adjust in response to a sibling's arrest. Thus, individuals may share information and learn from other family members, but the evidence on this is mixed.

Finally, there is robust evidence in favor of deterrence theory based on an individual's perceived

probability of arrest (and not actual local arrest rates). The strongest evidence in favor of deterrence is based on a structural model that incorporates unobserved heterogeneity in criminal ability and baseline beliefs as well as measurement error in beliefs. These estimates suggest that the effects of differences in beliefs about the probability of arrest are at least as important as differences in tastes for (or costs of) crime in explaining differences in criminal participation by race, ethnicity, or ability. Furthermore, even temporary changes in beliefs can impact criminal behavior for years into the future.

Altogether, the empirical findings support the economic model of crime and belief updating outlined in Section 2 of this paper. Beliefs are heterogeneous and idiosyncratic, but they respond to an individual's own arrest and non-arrest outcomes in predictable ways. While most of the literature on criminal deterrence assumes that individuals know true arrest rates and that an increase in those arrest rates will immediately deter crime, this paper suggests that this may not be the case. Individuals appear to learn about the probability of arrest as they gain more experience with the criminal justice system. As a result, responses to changes in enforcement are likely to differ across individuals with different crime and arrest histories, and the full impacts of any enforcement policy may not be realized for many years. Furthermore, policies that directly intervene to alter individual perceptions about the probability of arrest are likely to be quite effective at deterring crime.

References

Anderson, L. and C. Holt, "Information Cascades in the Laboratory," *American Economic Review*, 1997, *87* (5), 847–62.

Arellano, M. and B. Honore, "Panel Data Models: Some Recent Developments," in J. J. Heckman and E. Leamer, eds., *Handbook of Econometrics*, Vol. 5, New York: North-Holland, 2001, chapter 53.

Ayres, I. and S. Levitt, "Measuring the Positive Externalities from Unobservable Victim Precaution: An Empirical Analysis of Lojack," *Quarterly Journal of Economics*, Feb. 1998, *113* (1), 43–77.

Bachman, R., R. Paternoster, and S. Ward, "The Rationality of Sexual Offending: Testing a Deterrence/Rational Choice Conception of Sexual Assault," *Law and Society Review*, 1992, *26* (2), 343–72.

Banerjee, A., "A Simple Model Of Herd Behavior," *Quarterly Journal of Economics*, 1992, *107*, 797–817.

Becker, G., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*,

March/April 1968, *73*, 169–217.

Bikhchandani, S., D. Hirshleifer, and I. Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 1992, *100*, 992–1026.

Blumstein, A., D. Nagin, and J. Cohen, *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, Washington, DC: National Academy of Sciences, 1978.

Cameron, S., "The Economics of Crime Deterrence: A Survey of Theory and Evidence," *Kyklos*, 1988, *41*, 301–23.

Corman, H., T. Joyce, and N. Lovitch, "Crime, Deterrence and the Business Cycle in New York City: A VAR Approach," *Review of Economics and Statistics*, Nov. 1987, *69* (4), 695–700.

Ehrlich, I., "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy*, May/June 1973, *81* (3), 521–65.

— , "On the Usefulness of Controlling Individuals: An Economic Analysis of Rehabilitation, Incapacitation, and Deterrence," *American Economic Review*, June 1981, *71* (3), 307–22.

Elliott, D. and S. Ageton, "Reconciling Race and Class Differences in Self-Reported and Official Estimates of Delinquency," *American Sociological Review*, Feb. 1980, *45*, 95–110.

Grasmick, H. and R. Bursik Jr., "Conscience, Significant Others, and Rational Choice: Extending the Deterrence Model," *Law and Society Review*, 1990, *24* (3), 837–61.

Grogger, J., "Certainty vs. Severity of Punishment," *Economic Inquiry*, 1991, *29*, 297–301.

Heavner, D. L. and L. Lochner, "Social Networks and the Aggregation of Individual Decisions," 2002. NBER Working Paper No. 8979.

Horney, J. and Marshall I., "Risk Perceptions Among Serious Offenders: The Role of Crime and Punishment," *Criminology*, 1992, *30* (4), 575–94.

Klepper, S. and D. Nagin, "Tax Compliance and Perceptions of the Risks of Detection and Criminal Prosecution," *Law and Society Review*, 1989, *23* (2), 209–40.

Levitt, S., "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, June 1997, *87* (3), 270–90.

— , "Juvenile Crime and Punishment," *Journal of Political Economy*, Dec. 1998, *106* (6), 1156–85.

— , "Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?," *Economic Inquiry*, July 1998, *36*, 353–72.

Lochner, L., "Belief Updating and Criminal Careers," 2004. Working Paper.

Maguire, K. and A. Pastore, eds, *Sourcebook of Criminal Justice Statistics - 1998*, Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics, U.S. Government Printing Office,

1998.

Manski, C., "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?," in C. Clotfelter and M. Rothschild, eds., *Studies in Supply and Demand of Higher Education*, Chicago: University of Chicago Press, 1992, chapter 2.

Minor, W. and J. Harry, "Deterrent and Experiential Effects in Perceptual Deterrence Research: A Replication and Extension," *Journal of Research in Crime and Delinquency*, 1982, *19*, 190–203.

Myers, S., "Estimating the Economic Model of Crime: Punishment vs. Deterrent Effects," *Quarterly Journal of Economics*, 1983, *98*, 157–66.

Nagin, D., "Enduring Individual Differences and Rational Choice Theories of Crime," *Law and Society Review*, 1993, *27* (3), 467–96.

— , "Criminal Deterrence Research at the Outset of the Twenty-First Century," *Crime and Justice*, 1998, *23*, 1–42.

Paternoster, R. and S. Simpson, "Sanction Threats and Appeals to Morality: Testing a Rational Choice Model of Corporate Crime," *Law and Society Review*, 1996, *30* (3), 549–83.

— , **L. Saltzman, G. Waldo, and T. Chiricos**, "Estimating Perceptual Stability and Deterrent Effects: The Role of Perceived Legal Punishment in the Inhibition of Criminal Involvement," *The Journal of Criminal Law and Criminology*, 1983, *74* (1), 270–97.

— , — , — , **and** — , "Assessments of Risk and Behavioral Experience: An Exploratory Study of Change," *Criminology*, 1985, *23* (3), 417–36.

Piliavin, I., R. Gartner, C. Thornton, and R. Matsueda, "Crime, Deterrence, and Rational Choice," *American Sociological Review*, 1986, *51* (1), 101–19.

Pogarsky, G. and A. Piquero, "Can Punishment Encourage Offending? Investigating the 'Resetting' Effect," *Journal of Research in Crime and Delinquency*, 2003, *40* (1), 95–120.

Richards, P. and C. Tittle, "Gender and Perceived Chances of Arrest," *Social Forces*, 1981, *59* (4), 1182–99.

Sah, R., "Social Osmosis and Patterns of Crime," *Journal of Political Economy*, 1991, *99* (6), 1271–1295.

Saltzman, L., R. Paternoster, G. Waldo, and T. Chiricos, "Deterrent and Experiential Effects:The Problem of Causal Order in Perceptual Deterrence Research," *Journal of Research in Crime and Delinquency*, 1982, *19*, 172–89.

Schneider, A. and L. Ervin, "Specific Deterrence, Rational Choice, and Decision Heuristics: Applications in Juvenile Justice," *Social Science Quarterly*, Sept. 1990, *71* (3), 585–601.

Tauchen, H., A. D. Witte, and H. Griesinger, "Criminal Deterrence: Revisiting the Issue with a Birth Cohort," *Review of Economics and Statistics*, Aug. 1994, *76* (3), 399–412.

Tittle, C., "Assessments of Risk and Behavioral Experience: An Exploratory Study of Change," *Social Forces*, 1977, *55* (3), 579–96.

Trumbull, W., "Estimations of the Economic Model of Crime Using Aggregate and Individual Level Data," *Southern Economic Journal*, 1989, *94*, 423–39.

Viscusi, W. K., "The Risks and Rewards of Criminal Activity: A Comprehensive Test of Criminal Deterrence," *Journal of Labor Economics*, 1986, *4* (3, pt. 1), 317–40.

— , *Rational Risk Policy*, Oxford: Clarendon Press, 1998.

Waldfogel, J., "Criminal Sentences as Endogenous Taxes: Are they 'Just' or 'Efficient'?," *Journal of Law and Economics*, April 1993, *36(1) pt.1*, 62–81.

Walker, J., "Adolescents' Expectations on Birth Outcomes: A Comparison of the 1979 and 1997 NLS Cohorts," 2000. Working Paper.

Wilson, J. Q. and G. Kelling, "The Police and Neighborhood Safety: Broken Windows," *The Atlantic*, March 1982, pp. 29–38.

Witte, A. D., "Estimating the Economic Model of Crime with Individual Data," *Quarterly Journal of Economics*, Feb. 1980, *94* (1), 57–84.

Appendix: The Structural Likelihood Functions

This appendix develops the likelihood for the structural model of Section 6. Define $x_{i,t} = (p_{i,t}, c_{i,t}, A_{i,t})$ and $\Omega_i = (\mu_i, \xi_i)$. Denote by $x_i^t = (x_{i,1}, x_{i,2}, ..., x_{i,t})$ the history through date t for x. We observe data for $t = 1, 2, ..., T$. We also define $\varphi(\cdot)$ and $\Phi(\cdot)$ as the pdf and cdf, respectively, of the standard normal distribution. Let $\varphi(\cdot; \sigma)$ be the pdf for a normal random variable with a mean of zero and standard deviation of σ.

Define the following conditional probabilities:

$$
\begin{aligned}
g(p_{i,t}|c_{i,t}, A_{i,t}, p_{i,t-1}, Z_{i,t}, \mu_i) &= \varphi\left(p_{i,t} - Z_{i,t}\gamma - \theta p_{i,t-1} - \phi A_{i,t} - \lambda c_{i,t} - \mu; \sigma\right) \\
h(A_{i,t}|c_{i,t}, Z_{i,t}, \mu_i) &= \Phi\left(\pi_0 + \pi_1\mu_i + Z_{i,t}\pi_2 + \pi_3(1 - c_{i,t})\right)^{A_{i,t}} \\
&\quad \times \left[1 - \Phi\left(\pi_0 + \pi_1\mu_i + Z_{i,t}\pi_2 + \pi_3(1 - c_{i,t})\right)\right]^{1-A_{i,t}} \\
\kappa(c_{i,t}|p_{i,t-1}, W_{i,t}, \xi_i) &= \Phi\left(W_{i,t}A + Bp_{i,t-1} + \xi_i\right)^{c_{i,t}} \cdot \left[1 - \Phi\left(W_{i,t}A + Bp_{i,t-1} + \xi_i\right)\right]^{1-c_{i,t}}.
\end{aligned}
$$

Dropping i subscripts, let

$$
\begin{aligned}
f_t(x_t|x^{t-1}, Z^T, W^T, \Omega) &= g(p_t|c_t, A_t, p_{t-1}, Z_t, \mu) \cdot h(A_t|c_t, Z_t, \mu) \cdot \kappa(c_t|p_{t-1}, W_t, \xi) \\
&= f_t(x_t|p_{t-1}, Z_t, W_t, \Omega),
\end{aligned}
$$

which reflects the fact that the probability of state $x_t = (p_t, c_t, A_t)$ conditional on history $x^{t-1} = (x_1, x_2, ..., x_{t-1})$ only depends on p_{t-1} and current observables and unobservables.

Now, consider the likelihood of observing the sequence of beliefs, crime, and arrests $x_1, ..., x_T$ conditional on initial beliefs, observables, and unobservables:[47]

$$
\begin{aligned}
f(x_1, ..., x_T|p_0, Z^T, W^T, \Omega) &= f_T(x_T|x^{T-1}, Z^T, W^T, \Omega) \cdots f_2(x_2|x^1, Z^T, W^T, \Omega) \cdot f_1(x_1|p_0, Z^T, W^T, \Omega) \\
&= \prod_{t=1}^{T} f_t(x_t|p_{t-1}, Z_t, W_t, \Omega)
\end{aligned}
$$

Since p_0 and Ω are unobserved, we need to 'integrate them out' of the likelihood. Consider J values for unobserved types, each denoted by $\Omega_j = (\mu_j, \xi_j)$. Then, after integrating out over p_0 and unobserved types, the individual likelihood becomes

$$
\begin{aligned}
f(x_1, ..., x_T|Z^T, W^T) &= \sum_{j=1}^{J} \left[\int_{p_0} f(x_1, ..., x_T, p_0, \Omega_j|Z^T, W^T)\, dp_0\right] \\
&= \sum_{j=1}^{J} \left[\int_{p_0} \left(\prod_{t=1}^{T} f_t(x_t|p_{t-1}, Z_t, W_t, \Omega_j)\right) f_0(p_0|Z^T, W^T, \Omega_j) Pr(\Omega_j|Z^T, W^T)\, dp_0\right].
\end{aligned}
$$

---

[47] Conditioning on initial beliefs, p_0, is equivalent to conditioning on the initial state x_0.

At this point, one must specify the distribution of initial beliefs conditional on unobservables and observables as well as the distribution of unobservables conditional on observables.

If Ω is independent of Z^T and W^T, then $Pr(\Omega_j|Z^T, W^T) = \rho_j$. Furthermore, if individuals are young enough when they are first observed, such that everyone is in a no crime/no arrest state prior to observation, then it seems reasonable to assume that beliefs are drawn from the stationary no crime/no arrest distribution at time 0.[48]

We, therefore, make the following assumptions:

Assumption 1: $\Omega \perp\!\!\!\perp (Z^T, W^T)$.

Assumption 2: $p_0 \sim N\left(p^*(Z, \mu), (\sigma^*)^2\right)$.

These assumptions imply that the individual likelihood is

$$f(x_1, ..., x_T|Z^T, W^T) =$$
$$\sum_{j=1}^{J} \left[\left(\prod_{t=2}^{T} f_t(x_t|p_{t-1}, Z_t, W_t, \Omega_j) \right) \left(\int_{p_0} f_1(x_1|p_0, Z_1, W_1, \Omega_j) \varphi\left(p_0 - p^*(Z_0, \mu); \sigma^*\right) dp_0 \right) \rho_j \right].$$

When data on Z_0 is unavailable, it is natural to assume $Z_0 = Z_1$ for most variables.

Observing individuals at later ages

Now, consider the case in which individuals are not observed until later ages $k, k+1, ..., T$ where $k > 1$ (and individuals may have committed crime prior to the sampling period). In this case, the likelihood becomes

$$\begin{aligned}
f(x_k, ..., x_T|Z^T, W^T) &= \sum_{j=1}^{J} \left[\int_{p_{k-1}} f(x_k, ..., x_T, p_{k-1}, \Omega_j|Z^T, W^T) \, dp_{k-1} \right] \\
&= \sum_{j=1}^{J} \left[\int_{p_{k-1}} \left(\prod_{t=k}^{T} f_t(x_t|p_{t-1}, Z_t, W_t, \Omega_j) \right) \hat{f}_{k-1}(p_{k-1}|Z^T, W^T, \Omega_j) Pr(\Omega_j|Z^T, W^T) \, dp_{k-1} \right].
\end{aligned}$$

One could make simple but arbitrary assumptions about the density for beliefs at time $k-1$, $\hat{f}_{k-1}(p_{k-1}|\cdot)$, and how it depends on observables and unobservables. However, it is more natural to make assumptions about beliefs at time 0 (before individuals begin their criminal careers) using the structure of the model to determine the implied distribution of beliefs at date $k-1$. This is because period $k-1$ beliefs are likely to be correlated with observed and unobserved characteristics even if period 0 beliefs are not. Under assumptions 1 and 2, $Pr(\Omega_j|Z^T, W^T) = \rho_j$ and p_0 are drawn from the stationary no crime/no arrest distribution. The conditional density function, $\hat{f}_{k-1}(p_{k-1}|Z^T, W^T, \Omega_j)$,

---

[48]Here, we implicitly assume that Z has remained constant for some time as well.

is non-trivial to calculate when we do not observe persons from the beginning of their potential criminal lives, since entire sequences of beliefs, choices, and arrest outcomes from periods 1 to $k-1$ must be 'integrated over' (in addition to beliefs at time 0).

$$
\begin{aligned}
\hat{f}_{k-1}(p_{k-1}|Z^T, W^T, \Omega_j) &= \int_{p_{k-2}} \cdots \int_{p_0} f(p_{k-1}, p_{k-2}, ..., p_0|Z^T, W^T, \Omega_j) \, dp_0 \cdots dp_{k-2} \\
&= \int_{p_{k-2}} \cdots \int_{p_0} \left(\prod_{t=1}^{k-1} \tilde{f}_t(p_t|p_{t-1}, Z_t, W_t, \Omega_j) \right) \varphi\left(p_0 - p^*(Z_0, \mu); \sigma^* \right) \, dp_0 \cdots dp_{k-2}
\end{aligned}
$$

where

$$
\tilde{f}_t(p_t|p_{t-1}, Z_t, W_t, \Omega_j) = \sum_{c_t=0}^{1} \sum_{A_t=0}^{1} g(p_t|c_t, A_t, p_{t-1}, Z_t, \mu_j) h(A_t|c_t, Z_t, \mu_j) \kappa(c_t|p_{t-1}, W_t, \xi_j)
$$

integrates over all possible crime and arrest choices (assuming there are only one or zero possible crimes and arrests each period).

Combining terms, we obtain the likelihood (without measurement error) to be estimated:

$$
\begin{aligned}
f(x_k, ..., x_T|Z^T, W^T) &= \sum_{j=1}^{J} \left[\rho_j \left(\prod_{t=k+1}^{T} f_t(x_t|p_{t-1}, Z_t, W_t, \Omega_j) \right) \int_{p_{k-1}} \cdots \int_{p_0} f_k(x_k|p_{k-1}, Z_k, W_k, \Omega_j) \right. \\
&\quad \times \left. \left(\prod_{\tau=1}^{k-1} \tilde{f}_\tau(p_\tau|p_{\tau-1}, Z_\tau, W_\tau, \Omega_j) \right) \varphi\left(p_0 - p^*(Z_0, \mu); \sigma^* \right) \, dp_0 \cdots dp_{k-1} \right]
\end{aligned}
$$

For those Z_τ ($\forall \tau < k$) that are unknown, we assume $Z_\tau = Z_k$.

Numerical Approximation

Consider the multiple integral

$$
\begin{aligned}
& \int_{p_{k-1}} \cdots \int_{p_0} f_k(x_k|p_{k-1}, Z_k, W_k, \Omega_j) \left(\prod_{t=1}^{k-1} \tilde{f}_t(p_t|p_{t-1}, Z_t, W_t, \Omega_j) \right) \varphi\left(p_0 - p^*(Z_0, \mu); \sigma^* \right) \, dp_0 \cdots dp_{k-1} \\
&= \int_{p_{k-1}} \cdots \int_{p_0} f_k(x_k|p_{k-1}, Z_k, W_k, \Omega_j) \prod_{t=1}^{k-1} \left(\sum_{c_t=0}^{1} \sum_{A_t=0}^{1} h(A_t|c_t, Z_t, \mu) \kappa(c_t|p_{t-1}, W_t, \xi) \right. \\
&\quad \times \left. \varphi\left(p_t - Z_t\gamma - \theta p_{t-1} - \phi A_t - \lambda c_t - \mu; \sigma \right) \right) \varphi\left(p_0 - p^*(Z_0, \mu_j); \sigma^* \right) \, dp_0 \cdots dp_{k-1}
\end{aligned}
$$

This can be approximated by drawing L sequences of i.i.d. random shocks $\left\{ \tilde{\nu}_{i,t}^{\ell} \right\}_{t=0}^{k-1}$ for all N persons where each shock, $\tilde{\nu}_{i,t}^{\ell}$, is drawn from the standard normal distribution. For an individual, the above integral is approximated by

$$
\begin{aligned}
\frac{1}{L} \sum_{\ell=1}^{L} & \left[\sum_{c^{k-1}} \sum_{A^{k-1}} f_k \left(x_k | p_{k-1}^{\ell} \left(\left\{ \tilde{\nu}_j^{\ell} \right\}_{j=0}^{k-1}, c^{k-1}, A^{k-1} \right), \Omega_j, Z_k, W_k \right) \right. \\
& \times \left. \prod_{t=1}^{k-1} h(A_t|c_t, Z_t, \mu_j) \kappa \left(c_t | p_{t-1}^{\ell} \left(\left\{ \tilde{\nu}_j^{\ell} \right\}_{j=0}^{t-1}, c^{t-1}, A^{t-1} \right), W_t, \xi_j \right) \right].
\end{aligned}
$$

Here, $p_t^\ell \left(\left\{ \tilde{\nu}_j^\ell \right\}_{j=0}^t, c^t, A^t \right)$ represents the perceived probability of arrest given the random shock history sequence ℓ and a particular crime (c^t) and arrest (A^t) history all through period t.[49] Note that the interior sums are over all potential crime and arrest histories through period $k-1$ (there are 4^{k-1} such histories). The perceived probability for $t \in \{1, 2, ..., k-1\}$ can be determined recursively from:

$$p_t^\ell \left(\left\{ \tilde{\nu}_j^\ell \right\}_{j=0}^t, c^t, A^t \right) = \begin{cases} Z_t \gamma + \theta p_{t-1}^\ell \left(\left\{ \tilde{\nu}_j^\ell \right\}_{j=0}^{t-1}, c^{t-1}, A^{t-1} \right) + \phi A_t + \lambda c_t + \mu_j + \sigma \tilde{\nu}_t^\ell & \text{if } t \geq 1 \\ p^*(Z_0, \mu_j) + \sigma^* \tilde{\nu}_0^\ell & \text{if } t = 0. \end{cases}$$

Intuitively, this approach draws L sequences of random belief shocks for periods 0 to $k-1$. For each sequence of shocks, $\left\{ \tilde{\nu}_t^\ell \right\}_{t=0}^{k-1}$, we compute the implied evolution of beliefs for all possible crime and arrest paths through period $k-1$. The probability of each of those paths can be analytically determined by rolling forward through time using $f_t(x_t | p_{t-1}, Z_t, W_t, \Omega_j)$. At the end, we observe x_k in the data and compute the probability of observing that state for each of the possible crime and arrests paths (and for each draw of the shock sequence). The overall probability of observing x_k for each draw of the shock sequence can then be determined using the probability of each crime and arrest path. Finally, we average this probability over all L random shock sequences to get an approximation of the multivariate integral. This method can be easily be modified to account for missing observations in middle survey years.

Measurement Error

Assume that measurement error, $\eta_{i,t}$, is distributed $N(0, \sigma_\eta^2)$ and is independent of all variables and disturbance terms. Then, observed beliefs are $p_{i,t}^o = p_{i,t} + \eta_{i,t}$. When there is no structural error, the likelihood can be written (and computed directly without the need for numerical approximation) as:

$$\begin{aligned} f(x_k^o, ..., x_T^o | Z^T, W^T) &= \sum_{j=1}^J \left[\rho_j \sum_{c^{k-1}} \sum_{A^{k-1}} \left\{ \left(\prod_{\tau=k}^T \varphi[p_\tau^o - p_\tau(c^\tau, A^\tau, \mu_j, Z_\tau); \sigma_\eta] \right) \right. \right. \\ &\quad \left. \left. \times \left(\prod_{t=1}^T h(A_t | c_t, Z_t, \mu_j) \kappa(c_t | p_t(c^t, A^t, \mu_j, Z_t), W_t, \xi_j) \right) \right\} \right] \end{aligned}$$

where

$$p_t(c^t, A^t, \mu_j, Z_t) = \begin{cases} Z_t \gamma + \theta p_{t-1}(c^{t-1}, A^{t-1}, \mu_j, Z_{t-1}) + \phi A_t + \lambda c_t + \mu_j & \text{if } t \geq 1 \\ p^*(Z_0, \mu_j) & \text{if } t = 0. \end{cases}$$

The likelihood with both measurement error and a structural error reflects a combination of features from this likelihood and that with the structural error. This is available upon request.

---

[49]The perceived probability p_t^ℓ also depends on Z^t and μ_j.

Figure 1: Distribution of Perceived Probability of Arrest for Auto Theft (NLSY97)
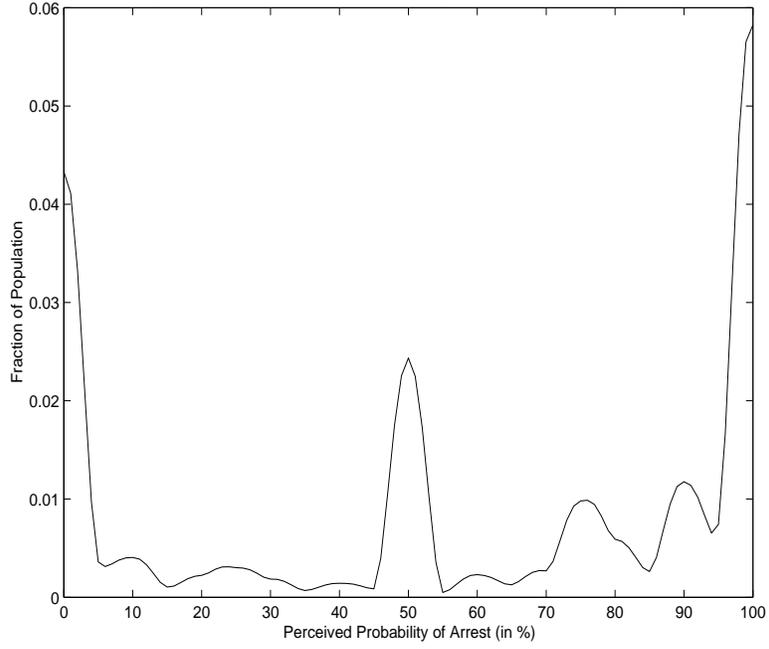
Figure 2: Distribution of Changes in Perceived Probability of Arrest for Auto Theft (NLSY97)
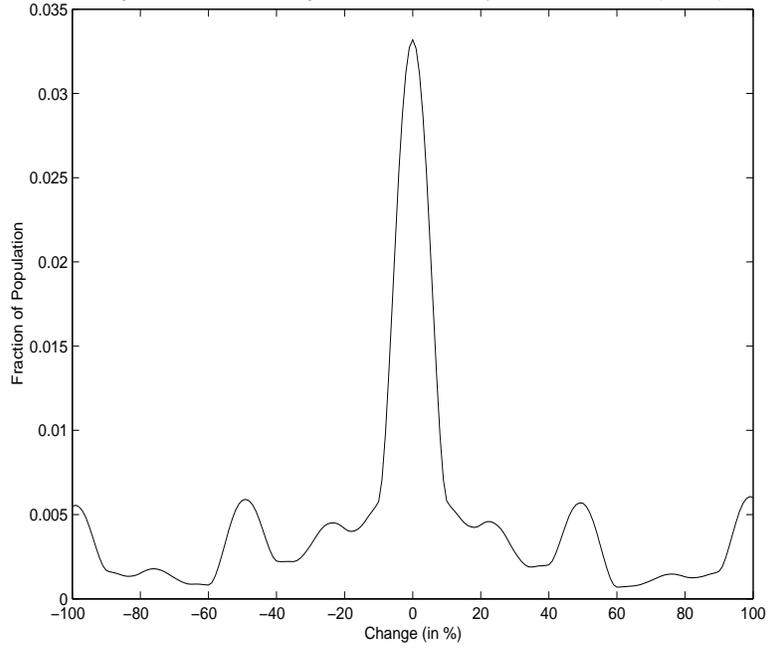
Figure 3: Distribution of Perceived Probability of Arrest in NYS
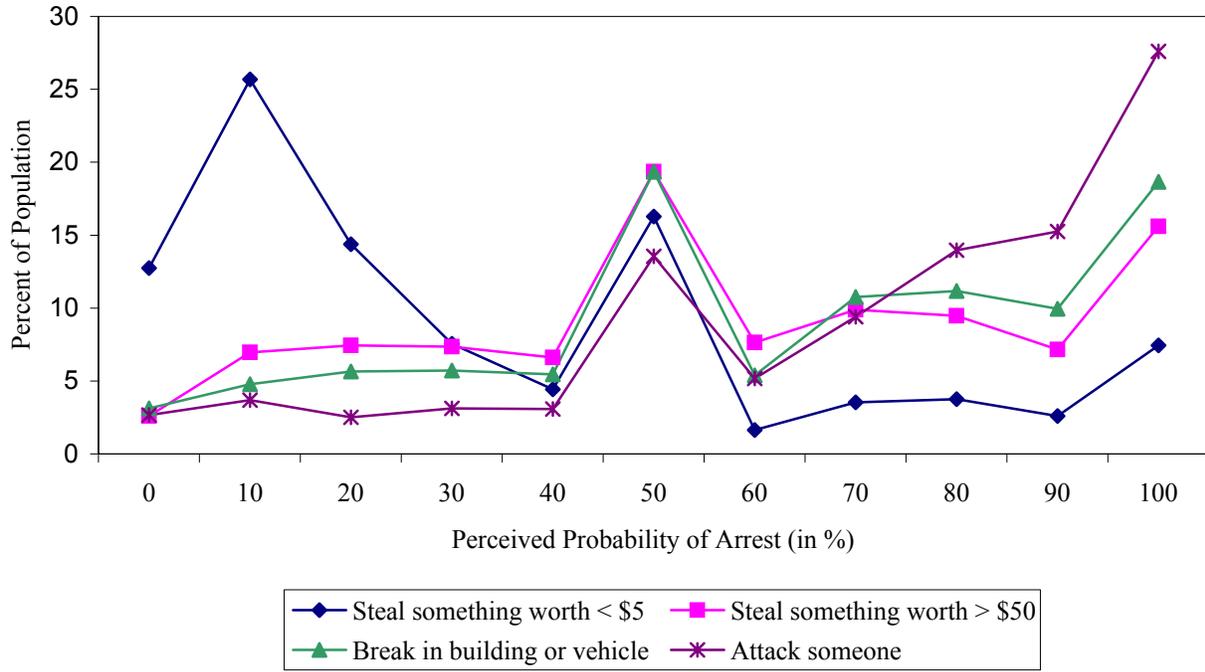
Figure 4: Distribution of Changes in Perceived Probability of Arrest in NYS
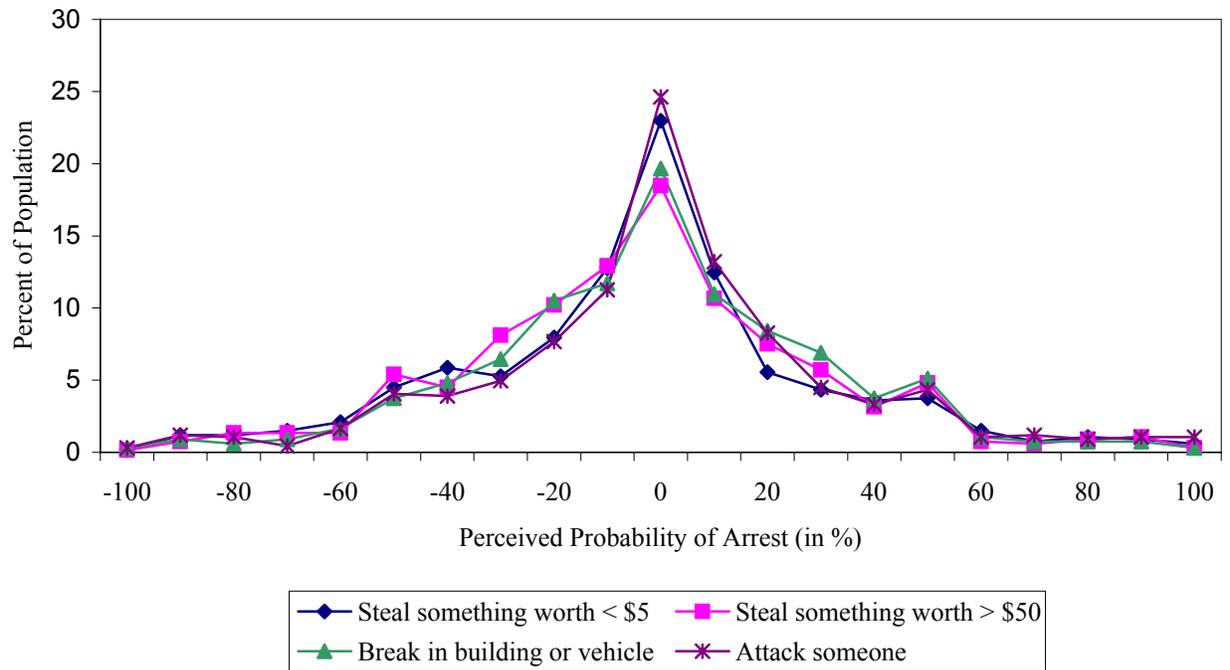
Figure 5: Simulated Effects of Changes in Beliefs and True Arrest Rates

a) Average Perceived Probability of Arrest
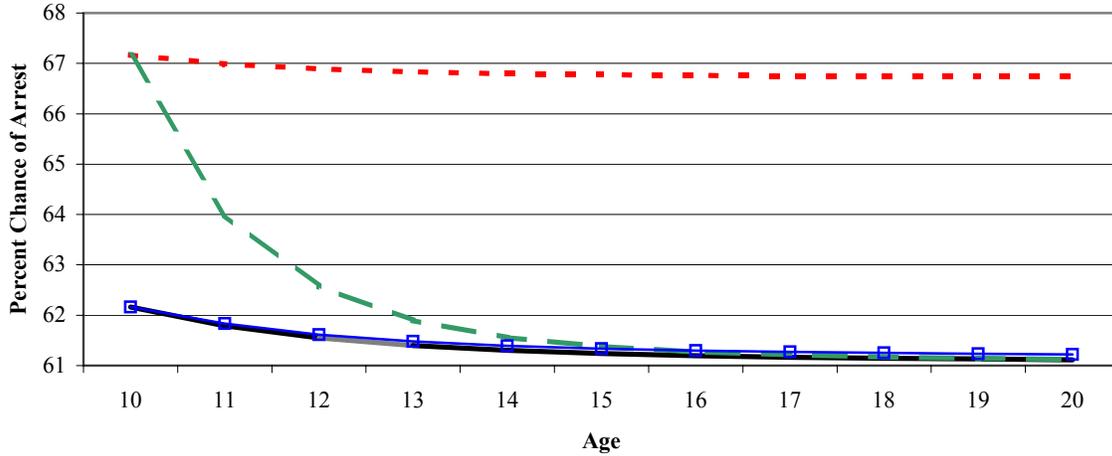
b) Crime Rate
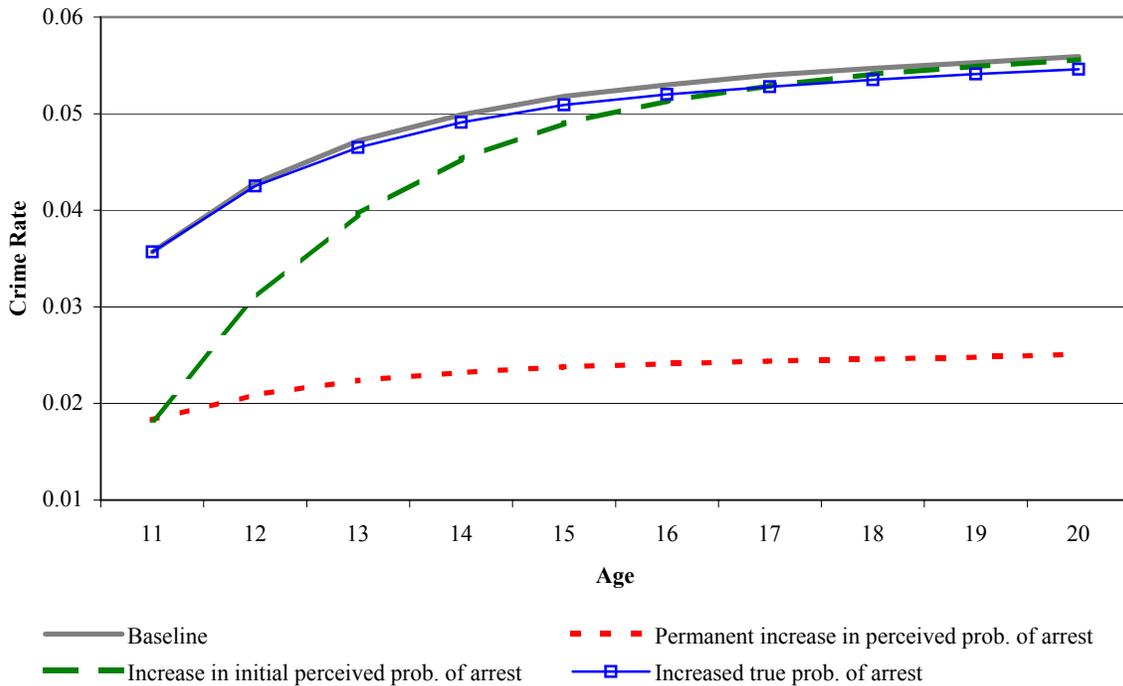
Notes: Permanent increase in the perceived probability of arrest reflects an increase in μ of $5*(1-\theta)$.   Increase in initial perceived probability reflects an increase in p_0 by 5. Increased true probability of arrest reflects an increase in π_0 by 0.2.

Table 1: Annual Self-Reported Crime and Arrests Among Males in the NLSY97

	All	Blacks	Hispanics	Whites
Number of respondents	4,559	1,169	977	2,413
Percent who stole something worth > $50	5.45	6.37	6.31	5.09
Percent who stole a vehicle	1.20	1.31	1.77	1.06
Avg. number of thefts > $50	0.36	0.44	0.57	0.30
Avg. number of thefts > $50 (of those who stole)	6.69	7.18	9.16	5.99
Percent arrested for any offense	8.37	11.22	9.07	7.61
Percent arrested for theft	1.72	2.28	1.74	1.60
Avg. number of arrests for theft	0.03	0.05	0.03	0.03
Persons arrested for theft / persons who stole > $50	0.32	0.36	0.28	0.31
Persons arrested for theft / persons who stole a vehicle	1.43	1.74	0.98	1.51
Arrests for theft / number of thefts > $50	0.09	0.12	0.06	0.09

Notes:
   All measures computed using panel sample weights.

Table 2: Average Perceived Probabilities (in %) of Arrest for Auto Theft (Males in NLSY97)

	All	Blacks	Hispanics	Whites
A) All Individuals	60.53	51.79	53.67	63.74
	(0.48)	(1.00)	(1.04)	(0.59)
B) Individuals who reported stealing something worth more than $50	50.46	43.49	43.61	53.88
	(1.67)	(3.57)	(2.81)	(2.22)
C) Individuals who reported stealing a car	44.78	40.50	39.72	47.42
	(2.97)	(5.82)	(4.65)	(4.16)
D) Weighted by number of thefts worth more than $50	39.31	35.76	35.77	41.71
	(3.02)	(6.50)	(5.96)	(4.09)

Notes:

Panel weights used in calculating all statistics. Standard errors, corrected for clustering across years for each individual, are in parentheses.

Table 3: OLS Estimates of Perceived Probability (in %) of Arrest for Auto Theft (Males in NLSY97)

Variable	(i)	(ii)	(iii)	(iv)
county arrest rate for motor vehicle theft	0.130	0.076	0.034	0.054
	(0.038)	(0.038)	(0.039)	(0.049)
age		-0.602	-0.634	-0.293
		(0.338)	(0.337)	(0.566)
black		-11.671	-11.625	-7.821
		(1.200)	(1.200)	(1.829)
hispanic		-9.600	-9.100	-8.713
		(1.239)	(1.250)	(1.799)
living in MSA			-4.699	-3.635
			(1.266)	(1.625)
family income less than $10,000				2.430
				(2.267)
living with both natural parents in1997				0.278
				(1.373)
PIAT score (percentile)				0.120
				(0.021)
mother a teenager at birth				-1.620
				(2.137)
R-square	0.002	0.019	0.020	0.030
Number of observations	13,800	13,800	13,800	7,141

Notes:

All specifications are weighted by panel weights and include a constant. Specifications (ii)-(iv) also control for year dummies. Standard errors, corrected for clustering across years for each individual, are in parentheses.

Table 4: Total Self-Reported Crimes and Arrests from 1984-1986 (Males in NYS)

Percent black	16.40
Percent hispanic	4.25
Percent who stole something worth < $5	17.70
Avg. number of thefts < $5	2.21
Percent who stole something worth > $50	3.72
Avg. number of thefts > $50	0.23
Percent who broke into a building or vehicle	2.29
Avg. number of breakins	0.12
Percent attacking someone to hurt or kill them	8.88
Avg. number of attacks	0.25
Percent arrested	11.90
Percent arrested for a property or violent offense	1.86
Average number of arrests	0.17
Average number of arrests for property or violent offense	0.02

Notes:

Arrests for property offenses include various forms of theft, evading payment, burglary, breaking and entering, and dealing in stolen goods. Arrests for violent offenses include assault, robbery, and harassment.

Table 5: Average Perceived Probabilities (in %) of Arrest (Males in NYS, 1983 & 1986)

Crime	All	Blacks	Hispanics	Whites
(i) Steal something worth $5 or less	33.84	43.55	38.37	31.86
	(0.90)	(2.54)	(4.60)	(0.97)
(ii) Steal something worth more than $50	57.81	63.10	58.57	56.78
	(0.87)	(2.25)	(4.49)	(0.97)
(iii) Break into a building or vehicle	62.49	67.22	66.33	61.54
	(0.88)	(2.26)	(4.71)	(0.98)
(iv) Attack someone to hurt or kill them	72.00	72.12	70.61	72.08
	(0.82)	(2.18)	(5.58)	(0.90)

Notes:
Standard errors, corrected for clustering across years for each individual, are in parentheses.

Table 6: Mean Perceived Probabilities (in %) of Arrest (Males in NYS, 1983 & 1986)

Crime	Did not commit this type of crime	Commited this type of crime	Weighted by Number of Crimes Committed
(i) Steal something worth $5 or less	35.64	19.19	20.43
(standard error)	(0.97)	(1.72)	(4.97)
[sample size]	[1,307]	[161]	[161]
(ii) Steal something worth more than $50	57.94	53.00	46.55
(standard error)	(0.88)	(5.08)	(8.86)
[sample size]	[1,428]	[40]	[40]
(iii) Break into a building or vehicle	62.77	51.67	44.67
(standard error)	(0.89)	(6.12)	(16.12)
[sample size]	[1,432]	[36]	[36]
(iv) Attack someone to hurt or kill them	73.43	54.78	52.76
(standard error)	(0.81)	(3.34)	(4.05)
[sample size]	[1,355]	[113]	[113]

Notes:

Standard errors, corrected for clustering across years for each individual, are in parentheses. Sample sizes in brackets.

Table 7: OLS Estimates of Perceived Probability (in %) of Arrest Among Males in NYS

Variable	(i) Steal something worth < $5	(ii) Steal something worth > $50	(iii) Break into building or vehicle	(iv) Attack Someone
Age	-0.596	-1.111	-0.430	0.393
	(0.343)	(0.340)	(0.344)	(0.350)
Black	9.866	4.974	4.958	-0.721
	(3.559)	(3.257)	(3.260)	(3.189)
Hispanic	5.054	1.344	4.964	-0.042
	(5.596)	(5.362)	(5.267)	(5.668)
Rural	4.625	6.866	5.129	3.121
	(2.163)	(2.083)	(2.208)	(2.129)
Central city	-1.813	-1.108	0.584	0.952
	(2.161)	(2.108)	(2.097)	(1.937)
Living with both parents in 1976	-1.545	-0.807	-5.109	-1.053
	(2.387)	(2.353)	(2.333)	(2.216)
Family income < $10,000 in1976	2.982	0.383	-1.278	-2.039
	(2.606)	(2.490)	(2.592)	(2.442)
Mother graduate from HS	-4.323	-1.189	-2.032	-0.819
	(2.354)	(2.237)	(2.225)	(2.057)
Father graduate from HS	-1.338	-2.554	-3.822	-0.862
	(2.364)	(2.258)	(2.426)	(2.247)
Neighborhood crime a problem	-1.477	0.845	0.354	-2.472
	(1.867)	(1.799)	(1.777)	(1.745)
Neighborhood disarray a problem	0.453	0.212	-1.737	1.444
	(2.223)	(2.162)	(2.181)	(2.107)
R-square	0.039	0.031	0.025	0.006

Notes:
All specifications also include an intercept term. Standard errors, corrected for clustering across years for each individual, are in parentheses. Sample size is 1,272.

Table 8: Belief Updating Among Males in the NLSY97
Dependent Variable: Perceived probability of arrest (in %)

Variable	(A) OLS (First differences)		(B) OLS (Quasi-first differences)		(C) GMM (Quasi-first diff. with fixed effects)	
	(i)	(ii)	(i)	(ii)	(i)	(ii)
County arrest rate (in percentage terms)	-0.036 (0.047)	-0.034 (0.047)	0.029 (0.040)	0.030 (0.040)	-0.058 (0.064)	-0.058 (0.064)
Perceived probability of arrest in previous year (in percentage terms)			0.298 (0.012)	0.299 (0.012)	0.038 (0.025)	0.039 (0.025)
Stole something worth > $50 in previous year	-4.060 (2.421)		-8.678 (2.289)		-9.303 (3.584)	
Sold drugs in previous year	-4.515 (1.779)		-5.853 (1.756)		-6.387 (3.402)	
Arrested for theft in previous year	8.712 (4.106)		9.256 (3.876)		10.021 (6.338)	
Num. times stole something worth > $50 in previous year		-0.314 (0.130)		-0.361 (0.118)		-0.352 (0.184)
Num. times sold drugs in previous year		0.015 (0.044)		-0.110 (0.043)		0.016 (0.094)
Num. times arrested for theft in previous year		2.692 (1.629)		3.555 (1.553)		0.562 (3.024)
Sibling stole something worth > $50 in previous year	6.813 (4.317)		2.346 (4.214)		-3.060 (5.418)	
Sibling sold drugs in previous year	-7.957 (3.087)		-10.594 (3.129)		-3.914 (7.305)	
Sibling arrested for theft in previous year	-3.828 (7.226)		1.696 (7.994)		-13.798 (14.318)	
Num. times siblings stole something worth > $50 in previous year		-0.334 (0.214)		-0.281 (0.260)		0.315 (0.258)
Num. times siblings sold drugs in previous year		-0.086 (0.071)		-0.174 (0.073)		-0.085 (0.249)
Num. times sibling was arrested for theft in previous year		-1.105 (2.043)		-0.846 (1.803)		-5.024 (2.840)
Tests (P-value):						
No effect of respondent information	0.003	0.076	0.000	0.000	0.008	0.251
No effect of sibling information	0.051	0.037	0.008	0.028	0.731	0.226
Equal respondent and sibling information	0.118	0.199	0.125	0.261	0.346	0.058

Notes:

First difference specifications regress changes in beliefs on changes in MSA status and the variables shown in the table. OLS quasi-first difference specifications regress current beliefs on the variables shown in the table as well as controls for race, age, MSA status, year dummies, PIAT percentile, whether the respondent lived with both natural parents at age 14, and whether the respondent's mother was a teenager when he was born. GMM (quasi-first difference with fixed effects) specifications control for age and MSA status in addition to the variables in the table. Tests of no effect of respondent (or sibling) information jointly test whether all coefficients on own (or sibling) crimes and arrests are zero. Test of equal respondent and sibling information tests whether all coefficients on crimes and arrests are equal for siblings and respondents. Tests in panels (A) and (B) are F-tests, while those in panel (C) are Wald tests. Sample weights are used. Standard errors for coefficient estimates are in parentheses.

Table 9: Belief Updating Among Males in the NYS
Dependent Variable: Perceived Probability of Arrest (in %) in 1986

Variable	Steal something worth < $5 (i)	Steal something worth < $5 (ii)	Steal something worth > $50 (i)	Steal something worth > $50 (ii)	Break in (i)	Break in (ii)	Attack Someone (i)	Attack Someone (ii)
Perceived probability of arrest in 1983 (in percentage terms)	0.318 (0.037)	0.319 (0.047)	0.329 (0.036)	0.325 (0.036)	0.371 (0.037)	0.370 (0.036)	0.249 (0.036)	0.268 (0.036)
Committed respective crime since 1984	-12.387 (3.283)		-8.669 (5.678)		-25.639 (7.024)		-19.164 (3.591)	
Number of times committed respective crime since 1984		-0.318 (0.139)		-0.870 (0.392)		-4.283 (1.708)		-4.642 (0.863)
Arrested for violent or property crime since 1984	6.568 (8.167)		7.103 (8.136)		18.097 (7.980)		12.012 (7.672)	
Number of times arrested for violent or property crime since 1984		4.300 (6.415)		10.008 (6.729)		15.297 (6.681)		9.349 (6.245)
Central city status	1.528 (2.536)	-1.772 (3.074)	0.946 (2.450)	0.751 (2.452)	2.740 (2.439)	2.522 (2.449)	1.094 (2.404)	1.196 (2.405)
Rural status	0.945 (3.056)	0.192 (4.031)	3.480 (2.944)	3.173 (2.946)	3.446 (2.931)	3.488 (2.943)	-0.150 (2.903)	-0.059 (2.904)
Neighborhood crime a problem	-1.085 (2.431)	-0.650 (3.004)	1.647 (2.360)	0.915 (2.353)	2.789 (2.347)	1.720 (2.349)	-1.977 (2.324)	-2.475 (2.317)
Neighborhood disarray a problem	1.345 (2.777)	6.053 (3.543)	-1.647 (2.691)	-1.111 (2.702)	-2.793 (2.676)	-2.297 (2.701)	-0.734 (2.642)	0.195 (2.647)
Victim of a crime since 1984	1.878 (2.322)	-4.332 (2.855)	3.698 (2.249)	3.605 (2.243)	2.618 (2.228)	2.761 (2.240)	1.788 (2.205)	1.589 (2.208)
Tests (P-value):								
No effect of respondent information	0.001	0.068	0.288	0.066	0.001	0.014	<.0001	<.0001
No effect of neighborhood crime or disarray	0.852	0.217	0.725	0.889	0.396	0.628	0.573	0.536

Notes:

All specifications also control for age, race/ethnicity (black and hispanic), whether the individual's parents earned less than $10,000 in 1976, and whether the individual lived with both natural parents in 1976. Test for no effect of respondent information is an F-test whether the coefficients on arrests and crimes committed since 1984 are both zero. Test for no effect of neighborhood crime or disarray is an F-test whether the coefficients on changes in neighborhood crime and disarray indicators are both zero.

Table 10: Effects of Beliefs on Criminal Activity Multiplied by 100 (Linear Probability Model, NLSY97)

	OLS		Fixed Effects 2SLS	
Variable	Steal something worth > $50	Auto Theft	Steal something worth > $50	Auto Theft
percent chance of arrest for auto theft	-0.021	-0.008	-0.016	-0.010
	(0.008)	(0.004)	(0.012)	(0.005)
age	0.153	0.083	1.959	0.274
	(0.224)	(0.112)	(1.303)	(0.585)
black	-0.387	-0.468		
	(0.981)	(0.484)		
hispanic	0.361	0.235		
	(0.977)	(0.506)		
living in MSA	1.668	0.081	-5.950	-0.058
	(0.842)	(0.408)	(3.846)	(0.119)
living with both natural parents in 1997	-2.895	-1.130		
	(0.768)	(0.352)		
mother a teenager at birth	-1.019	-0.602		
	(1.044)	(0.468)		
PIAT score (percentile)	-0.027	-0.015		
	(0.011)	(0.006)		
local arrest rate	0.058	0.004	-0.025	-0.009
	(0.028)	(0.011)	(0.040)	(0.016)

Notes: Fixed Effects 2SLS estimates use year t-2 reported number of thefts > $50 for individuals and siblings, year t-2 reported number of arrests for theft for individuals and siblings, and the year t-2 perceived probability of arrest for auto theft as instruments for the year t-2 to t-1 change in the perceived probability of arrest. Sample weights are used. Standard errors corrected for clustering at individual level.

Table 11: Parameter Estimates for Structural Model of Belief Updating, Crime, and Arrests

Parameters	No Measurement Error		Measurement Error	
	Estimates	Standard Errors	Estimates	Standard Errors
Perceptions Equation:				
black	-5.410	1.252	-3.489	1.333
hispanic	-5.998	1.354	-3.747	1.414
living with both natural parents in 1997	0.793	1.067	0.327	0.661
mother a teenager at birth	-0.601	1.527	-0.505	0.935
living in MSA	-3.670	1.248	-2.448	1.056
PIAT score (percentile)	0.096	0.016	0.060	0.022
perceived probability of arrest at previous survey (θ)	0.075	0.016	0.400	0.190
stole something worth > $50 since previous survey (λ)	-4.519	2.004	-13.965	4.863
arrested since previous survey (φ)	4.802	3.250	5.783	17.269
μ_1	29.509	2.005	19.586	6.253
μ_2	67.698	2.101	44.788	14.170
standard deviation of transitory belief shock (σ)	33.995	0.376		
standard deviation of measurement error (σ_η)			33.685	0.355
Crime Equation:				
black	-0.053	0.096	-0.444	0.160
hispanic	-0.019	0.094	-0.443	0.177
living with both natural parents in 1997	-0.325	0.072	-0.204	0.096
mother a teenager at birth	-0.137	0.121	-0.160	0.134
living in MSA	0.093	0.091	-0.179	0.132
PIAT score (percentile)	-0.004	0.001	0.003	0.003
age	0.020	0.021	0.003	0.020
perceived probability of arrest at previous survey (B)	0.000	0.001	-0.065	0.022
ξ_1	-0.822	0.360	0.654	0.836
ξ_2	-1.149	0.361	3.721	1.631
Arrest Equation:				
intercept (π_0)	-0.662	0.196	-0.866	0.210
coefficient on μ_i (π_1)	0.001	0.003	0.007	0.005
black	0.048	0.086	0.054	0.085
hispanic	-0.037	0.095	-0.029	0.096
living with both natural parents in 1997	-0.131	0.079	-0.134	0.078
mother a teenager at birth	-0.052	0.115	-0.045	0.116
living in MSA	-0.109	0.097	-0.111	0.096
PIAT score (percentile)	-0.005	0.001	-0.005	0.001
no theft worth > $50 since previous survey (π_3)	-1.410	0.084	-1.405	0.085
Unobserved Type Probabilities:				
probability of non-criminal with $\mu_i = \mu_1$ (ρ_1)	0.205	0.019	0.058	0.103
probability of non-criminal with $\mu_i = \mu_2$ (ρ_2)	0.448	0.035	0.377	0.068
probability of $\xi_i = \xi_1$, $\mu_i = \mu_1$ (ρ_3)	0.109	0.016	0.257	0.104

Note: Sample size is N=2,518.