# RISE | RICE INITIATIVE *for the* STUDY *of* ECONOMICS

RISE Working Paper 14-020

# "Structural Estimation of a Becker-Ehrlich Equilibrium Model of Crime: Allocating Police Across Cities to Reduce Crime"

by

Chao Fu, Kenneth I. Wolpin

RICE Department of Economics
Baker Hall, MS22
6100 Main Street, Houston, Texas 77005
https://economics.rice.edu

# Structural Estimation of a Becker-Ehrlich Equilibrium Model of Crime: Allocating Police Across Cities to Reduce Crime*

Chao Fu† Kenneth I. Wolpin‡

August 18, 2014

**Abstract**

We develop a model of crime in which the number of police, the crime rate, the arrest rate, the employment rate and the wage rate are joint outcomes of a subgame perfect Nash equilibrium. The local government chooses the size of its police force and citizens choose among work, home and crime alternatives. We estimate the model using MSA-level data. We use the estimated model to examine the effects on crime of targeted federal transfers to local governments to increase police. We find that knowledge about unobserved MSA-specific attributes is critical for the optimal allocation of police across MSA's.

*We thank Fumihiko Suga for excellent research assistance. We thank John Kennan, Thomas MaCurdy, Xiaoxia Shi and Chris Taber for insightful discussions and workshop participants at SITE conference, the Ohio State University and the University of Virginia for helpful comments. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

†Department of Economics, University of Wisconsin-Madison. 1180 Observatory Dr. Madison, WI 53706. Email: cfu@ssc.wisc.edu

‡Rice University and University of Pennsylvania

# 1 Introduction

The modern literature on the economics of crime, originating with Becker (1968), recognized that the crime rate is the equilibrium outcome of the joint determination of the supply of crime, resulting from the uncoordinated decisions of citizens about the supply of their labor to legal and/or illegal income-generating activities, and the demand for crime, determined by a government policy maker who decides on the level of resources to commit to preventing crime.[1] Much of the empirical literature, beginning with Ehrlich (1973), has used aggregate data based on either cross-sectional or time-series variation.[2] Coupled with the supply/demand theoretical framework, the use of aggregate data led naturally to the adoption of a simultaneous equations econometric structure.[3] The estimating equations in that system consisted of a supply of offenses function representing the decision rule of potential criminals, an apprehension production function and the policy maker's decision rule governing the level of resources devoted to apprehension (and punishment).

The econometric structure was meant to approximate the solution to the equilibrium model. As such, its parameters are combinations of those of the underlying behavioral structure, that is, of the preference function of potential offenders (the citizenry), the apprehension production function, the objective function of the policy maker and the distribution of the unobservables that enter those functions. An alternative approach, pursued in this paper, is to estimate the behavioral structure by specifying and solving a parametric model of agent decision-making. This approach relaxes two potentially important aspects of the approximation in the traditional approach. First, it explicitly aggregates the individual decisions about whether to engage in illegal activities over the citizen population, allowing for nonlinearities and location-specific effects as

[1] For a recent updated version of this framework, see Ehrlich (2010).

[2] For a recent review, see Tauchen (2010).

[3] This contrasts with the empirical literature on labor supply that took a decidely partial equilibrium approach given the availability of micro-level data (for example, Heckman (1971)). However, recent studies have used longitudinal data to estimate the supply of crime function, for example, Lochner (2004), Lochner and Moretti (2004) and Imai and Krishna (2004). For an early specification of dynamic models of criminal choice, see Flinn (1986).

would naturally arise, and second, it accounts for the existence of multiple equilibria that is inherent in the Beckerian structure.

It is useful to illustrate these issues in a simplified version of the model that we estimate. Assume a cross-section of cities, where each city ($j$) has a government and a continuum of citizens ($i$), each endowed with a preference over criminal behavior and potential legal sector earnings drawn from the distribution $F_j(\eta_{ji}, w_{ji}^L)$. Illegal sector earnings are $w_j^I$, assumed for simplicity to be the same for all potential criminals. Citizens optimally choose whether to engage in criminal activity or work in the legal sector taking into account the probability they will be apprehended and the severity of the punishment. The probability that a criminal is apprehended, $\pi_j$, depends on the level of enforcement expenditures chosen by the government, $s_j$, the aggregate level of crime, $\mu_j$ (the fraction of citizens who choose to engage in crime) and the city's efficiency of enforcement, $\epsilon_j$.[4] The apprehension technology is such that given $s_j$, the rate at which criminals can be apprehended falls with increases in the crime rate as enforcement resources are spread more thinly. The government maximizes an objective function that takes into account the cost of enforcement and the level of crime.[5]

Let $d_{ji} = 1$ if citizen $i$ residing in city $j$ chooses to engage in crime ($= 0$ otherwise), which depends on the apprehension rate in the city, the citizen's legal sector wage and preference for criminal behavior, and the (city-level) illegal sector wage, that is, $d_{ji} = d(\pi_j, w_{ji}^L, w_j^I, \eta_{ji})$. The overall crime rate in city $j$, $\mu_j$, is given by $\mu_j = \int d(\pi_j, w_{ji}^L, w_j^I, \eta_{ji}) dF_j(\eta_{ji}, w_{ji}^L)$. If the propensity to engage in illegal activity is monotonically declining in $\pi_j$ at the citizen level, then the crime rate, $\mu_j$ is declining in $\pi_j$ as well.[6] The apprehension production function

[4] The number of crimes per criminal is set to one for all cities.

[5] The severity of punishment is taken to be the same across cities.

[6] Note that the number of additional citizens who choose to engage in crime when the apprehension rate falls will in general not be homogeneous across cities as long as the joint distribution of preferences and legal earnings opportunities ($F_j$) differs across cities. Thus, even if one wants to maintain (log) linearity (and additive separability) in the relationship between the crime rates in cities and their apprehension rates, as in the traditional approach, a more appropriate specification would allow for random coefficients. This observation is particularly important in the context of instrumental variables (see Angrist and Imbens (1994), Heckman and Vytlacil(1998)) as is a common procedure used in estimating the deterrent

is given by

$$\pi_j = \Pi(s_{j,}\mu_j, \epsilon_j), \tag{1}$$

where $\pi_j$ is monotonically increasing in the level of enforcement resources, $s_j$, decreasing in the level of crime $\mu_j$ and increasing in enforcement efficiency, $\epsilon_j$.

The equilibrium crime rate for given $s_j$ is the solution to

$$\mu_j = \int d(\Pi(s_{j,}\mu_j, \epsilon_j), w_{ji}^L, w_j^I, \eta_{ji})dF_j(\eta_{ji}, w_{ji}^L). \tag{2}$$

If, in (1), the apprehension rate is one when the crime rate is zero and if preferences are such that no citizen will commit a crime when apprehension is certain, then for positive $s_j$ a zero crime rate is always an equilibrium. In addition, as shown in Conley and Wang (2006), with multi-dimensional heterogeneity across individuals, for positive $s_j$ there may be multiple interior equilibria with a positive crime rate. Given multiple solutions for $\mu_j$ in (2), there will be a corresponding apprehension rate for each $\mu_j$ (given $s_j$). The fundamental reason for the occurrence of multiple equilibria is the existence of the spillover effect of increasing crime on the apprehension probability. As seen in (2), the crime rate would be unique if $\mu_j$ did not enter $\pi_j$ in (1). Because it does, there is in essence a coordination problem that may lead to multiple equilibria.

The existence of multiple equilibria makes the interpretation of the regressions in the traditional approach potentially problematic, because each city $j$, facing different fundamentals, selects an equilibrium from a different set. Such sets may differ across cities in the number of equilibria in the set and by the values of the equilibrium objects, $\mu_j$ and $\pi_j$, in each equilibrium within the set. In order that the regression coefficient on, for example, the apprehension rate reflect what would happen to the crime rate within a location if the apprehension rate were to exogenously change, the goal of estimation, it is necessary that the equilibrium selected across locations that differ in their apprehension rates is in some sense the same. Examples of equilibrium selection rules that satisfy a sameness criterion include choosing the equilibrium with the lowest

effect of apprehension with aggregate data.

crime rate or choosing the Pareto dominant equilibrium (if one always exists). If different cities select different equilibria, then the regression coefficient would reflect that change in equilibria.

The previous discussion took $s_j$ as given. However, the policy maker can obviously manipulate the apprehension probability through the choice of $s_j$. It seems reasonable then to assume that the policy maker acts strategically, choosing the level of law enforcement resources recognizing its effect on the apprehension rate and, thus, on the incentives for citizens to engage in crime. The government, as a strategic player, accounts for the spillover in determining the optimal level of enforcement. However, because the equilibrium crime rate that results from a choice of $s_j$ is not necessarily unique, the government must know (or have some belief about) the equilibrium selection rule. Given a government objective function, for example as in Becker (1968) minimizing the expected societal loss from crime, the optimal choice of $s_j$ would be a function of the model fundamentals, that is,

$$s_j = s(F_j(\eta_{ji}, w_{ji}^L), w_j^I, \upsilon_j, \theta_\mu, \theta_\pi, \theta_R), \tag{3}$$

where $\upsilon_j$ represents location-specific unobservables in the government objective function, $\theta_\mu$ and $\theta_\pi$ are the parameters in the crime decision function $(d\,(\cdot))$ and apprehension production function and $\theta_R$ are the parameters of the equilibrium selection rule.

The model that we estimate specifies the optimization problem of potential criminals and of the policy maker in a city. Each city has a continuum of individuals of different types in terms of their preferences for staying at home and for committing crimes, and in terms of their legal sector human capital. Types are correlated with observable characteristics. Besides the distributions of their citizens' characteristics, cities also differ in unobservable characteristics that include their production technology, apprehension efficiency and marginal costs of the police force. Each citizen chooses whether to be a criminal, work in the legal sector or stay at home. City output is produced with the human capital of citizens who choose to work in the legal sector. Criminals meet victims randomly and receive a fraction of the victim's legal sector income. For a given city, the arrest rate varies with the crime rate and the size of the police force. Given any size of the police force, it is feasible to solve for all possible equilibria because the model structure yields an ordering over citizens in their propensity to be criminals that is invariant across equilibria.

Acting as a Stackelberg leader, the policy maker chooses the number of police to maximize the expected value of an objective function that includes the number of police (negatively), the crime rate (negatively) and the apprehension rate (positively), where the expectation is taken over the distribution of equilibria that we estimate. The solution to the model yields equilibrium values of the number of police, the employment rate, the crime rate, the apprehension rate and the competitively determined rental price of human capital.

The estimation of the model is by simulation. At any given set of parameter values and set of location-specific unobservable characteristics, the model can be solved for the equilibrium objects for each location and for all of the equilibria. Doing multiple simulations for the same parameter values by randomly drawing from the distribution of unobservables and averaging over the simulations within each location provides statistics for the equilibrium outcomes that can be matched with data. The data are from two sources: the Uniform Crime Reports of the FBI and the Current Population Survey. We focus on a single year, 2008, and on the 238 MSA's that can be matched between the two data sets. The CPS provides demographics, employment and wage data, while the UCR provides data on crimes, arrests and the number of police. We apply the model to property crimes.

With our estimated model, we conduct two sets of counterfactual policy experiments that are not feasible within the conventional simultaneous equations estimation framework using cross-sectional data. These experiments are motivated by federal programs that provide subsidies to local governments for increasing the number of police. For example, the Community Oriented Police Services (COPS) program, initiated by the Clinton administration in 1994, aimed at a nationwide increase in the number of police of 20 percent.[7] In all of

[7]See Zhao, Scheider and Thurman (2002) and Evans and Owens (2007) for an analysis of

the experiments, we assume that the federal government can perfectly monitor the use of the resources to any locality, ensuring that the intended increase in the size of the police force is realized. In the first set of experiments, we explore two scenarios. In the first, unlike the COPS program, the planner (federal government) uniformly increases the size of the police force by 20 percent for each MSA; the program leads to an 8.2 percent reduction in the national crime rate.

In the second scenario, as in the COPS program, instead of a uniform transfer, the planner subsidizes newly hired police, where the number of new hires is chosen by each local government. First, we determine the subsidy rate such that the planner's total spending is the same as under the uniform transfer. At such a subsidy rate (48 percent), there is a 42 percent increase in the number of police and the national crime rate falls by 21 percent, or 2.6 times the decrease in the uniform policy. Second, we determine the subsidy rate such that the increase in police nationwide, hence the total cost (federal and local), is the same as in the uniform transfer case, i.e., a 20 percent increase. The subsidy rate in that case is 27 percent, which leads to a 10.6 percent decrease in crime, as compared to 8.2 percent in the uniform transfer case. Allowing local governments to optimally choose the number of police, instead of a uniform transfer, leads to a greater reduction in crime. Finally, we determine the subsidy rate such that the total spending mimics the intended cost of the COPS program. Such a subsidy rate (34 percent) leads to a 26 percent increase in the number of police, as compared to the 20-percent goal set by COPS, and reduces the crime rate by 13 percent, which would be the effect of a program like COPS if the federal government were able to perfectly monitor the use of its grants.

Our second set of experiments explores the effects of various targeting schemes to allocate federally-sponsored additional police across locations, given a fixed resource constraint. We illustrate the idea by focusing on pair-wise allocation problems, where the total additional resource to be allocated between a pair of MSA's is equivalent to 20 percent of their current total police force. To achieve the optimal targeting, we first use the model estimates, for each MSA, to determine the values of the MSA-level unobserved characteristics: the MSA's arrest

that program.

efficiency, marginal cost of police, value of leisure and productivity. We explore five different allocation rules. The first assumes the planner has information on both the observable and unobservable characteristics of the MSA's and chooses the allocation to minimize the overall crime rate of the pair. The second assumes the planner solves the same optimization problem but with information only on observables. The other three rules are based on current crime rates, on current arrest rates and on current GDP per capita. We study the cases for five pairs of MSA's. In the case of Philadelphia versus Phoenix, for example, the allocation with complete information leads to a 16 percent decrease in the crime rate, while the crime reduction ranges from 5.6 percent (allocation based on crime rates) to 10.1 percent (allocation based on observables) across the other four allocation rules. Although the allocation with complete information always dominates the other rules, the relative effectiveness of the other four allocation rules is found to be case-dependent.

Our paper contributes to the literature that studies crime from an equilibrium perspective. Via different channels, various theoretical studies have demonstrated the existence of multiplicity of equilibria in models of crime. For example, Conley and Wang (2006) study an equilibrium model where agents, with heterogenous working abilities and tastes for crime, choose either to commit crimes or invest in education and become workers. The arrest rate depends on the crime rate and the number of police. They establish that when individuals differ in more than one dimension, multiple interior equilibria with different positive crime rates may exist. Taking the arrest rate as an exogenous parameter, Burdett, Lagos and Wright (2003, 2004) introduce crime into an otherwise classical random search equilibrium framework, where firms post wages and meet workers with an exogenous probability.[8] Besides the random job offers they receive, workers (unemployed or employed) may also receive a criminal opportunity at random. Workers choose whether or not to accept the job offer in the case they receive one and whether or not to commit a crime in the case they receive a criminal opportunity. Multiple interior equilibria with a positive

[8] Examples of other papers that build crime in a search model framework include Engelhardt, Rocheteau and Rupert (2008) and Huang, Liang and Wang (2004).

crime rate may arise due to matching externalities.[9] Our paper is the first to empirically implement a model of crime with multiple equilibria.

Imrohoroglu, Merlo and Rupert (2004) develop and calibrate a dynamic (supply-side) equilibrium model to study the trend in crime between 1980 and 1996. Taking the arrest rate as exogenous, they study individuals' dynamic choices of whether or not to be a criminal after a stochastic period-specific employment status is realized. An equilibrium requires, among others, that the aggregate crime rate is consistent with individuals' choices. Imrohoroglu, Merlo and Rupert (2000) embed a static equilibrium model of crime in a political economy framework. Individuals choose to specialize in either legitimate or criminal activities. The police force is funded by tax revenues from labor income, where the tax rate is determined via a majority-voting rule. The number of police is the sole determinant of the arrest rate, regardless of the crime rate.

The rest of the paper is as follows. The next section describes the model. Section 3 describes the data and Section 4 our estimation strategy and results. Counterfactual experiments are presented in Section 5. The last section concludes the paper. Some details and additional tables are in the appendix.

# 2 Model

There are $J$ cities $j = 1, ..., J$, each with a government and a continuum of individual citizens. Cities are considered as closed economies.[10] Each government acts as a Stackelberg leader by choosing the size of its police force. Observing their government's decision, individuals in each city choose one of the three mutually exclusive and exhaustive discrete options: work in the legal sector, work in the criminal sector or remain at home.

Each citizen is endowed with a human capital level ($l$), a taste for crime

[9] Besides matching externalities and multi-dimensional heterogeneity among agents, multiple equilibria in models of crime may also arise from other sources, for example, geographical externalities (e.g., Freeman, Grogger and Sonstelie (1996) and Verdier and Zenou (2004)) and direct interpersonal spillovers (e.g., Sah (1991) and Glaeser, Sacerdote and Scheinkman (1996)).

[10] This assumption rules out mobility between cities, an extension we leave for future work.

($\eta$) and a value of "leisure" when at home ($\kappa$). The triplet $(l, \eta, \kappa)$ defines an individual's type, which is unobservable to the researcher, but correlated with observable characteristics ($x$).[11] The distribution of $x$, $G_j(x)$, is city-specific as is the distribution of types. Each component of an individual's type is assumed to be discrete with $l \in \{l_1, ...l_{N_l}\}$, $\eta \in \left\{\eta_1, ..., \eta_{N_\eta}\right\}$ and $\kappa \in \{\kappa_1, ..., \kappa_{N_\kappa}\}$. Therefore, there are $N = N_l \times N_\eta \times N_\kappa$ types of individuals. We let $n \in \{1, ..., N\}$ be the index of a type defined as $(l_n, \eta_n, \kappa_n) \in \{l_1, ...l_{N_l}\} \times \left\{\eta_1, ..., \eta_{N_\eta}\right\} \times \{\kappa_1, ..., \kappa_{N_\kappa}\}$. Denote the proportion of individuals of type $n$ in city $j$ as $p_{jn}$. We denote the discrete choices for a type $n$ individual as $d_{n1} = 1$ if working in the legal sector ($= 0$ otherwise), $d_{n2} = 1$ if at home ($= 0$ otherwise) and $d_{n3} = 1$ if working in the criminal sector ($= 0$ otherwise).

## 2.1 The Legal Sector

Legal sector output in city $j$, $Y_j$, is produced using the aggregate stock of human capital of those citizens in city $j$ who choose to work in that sector, $L_j$. The production technology is given by

$$Y_j = \tau_j L_j^\theta,$$

where $\theta \in (0, 1)$ is the elasticity of output with respect to aggregate human capital and $\tau_j$ is a city-specific Hicks-neutral technology factor drawn from the distribution $\tau_j \sim \ln N(-0.5\sigma_\tau^2, \sigma_\tau^2)$. Assuming a competitive labor market in each city, the rental rate for a unit of human capital is given by its marginal product,

$$r_j = \tau_j \theta L_j^{\theta-1}, \tag{4}$$

and earnings for an individual of type $n$ residing in city $j$, $y_{jn}$, is the product of the rental price in city $j$ and the individual's level of human capital, that is, $y_{jn} = r_j l_n$.

[11] The characterisitcs included in $x$ are age, gender, race, education and the number of young children, all treated as discrete variables.

## 2.2 The Criminal Sector

Each law-abiding citizen in city $j$, whether working or at home, faces an equal probability $\mu_j$ of being the victim of a crime. If victimized, an individual loses a fraction $\alpha$ of his income to the criminal.[12] The income of workers is described above. The income of non-workers, denoted as $b_j$, is assumed to be equal to the human capital rental price in the city times the lowest level of human capital, $l_1$.[13] A citizen who chooses to work in the criminal sector faces probability $\pi_j$ of being arrested. The probability of an arrest (the arrest production function) depends positively on the size of the police force, $s_j$, and negatively on the crime (victimization) rate, $\mu_j$. There is a city-specific unobserved component of technology, $\epsilon_j$, which captures the unobservable factors that affect the efficiency of criminals in avoiding arrests, or equivalently police inefficiency.[14] A city with a lower value of $\epsilon_j$, that is, with higher police efficiency, has a higher arrest rate for a given number of police. The arrest technology function is given by

$$\pi_j = \Pi(s_j, \mu_j, \epsilon_j) = \exp(-\frac{\gamma(\mu_j \epsilon_j)^\rho}{s_j}), \tag{5}$$

where $\rho > 0$, $\epsilon_j \sim \ln N(-0.5\sigma_\epsilon^2, \sigma_\epsilon^2)$, and $\gamma > 0$ is a normalizing constant.[15] The functional form ensures that $\Pi(\cdot) \in [0, 1]$. Notice that the arrest rate declines with the crime rate and that the rate of decline depends on whether $\rho \gtreqless 1$. Note also that the parameterization of the degree of police (in)efficiency (which has

[12] We restrict attention only to property crimes.

[13] The model does not distinguish between non-participation and unemployment. The income of law-abiding non-workers is thus intended to capture both non-labor income and unemployment insurance. It is set to a low value, although one that varies with a city's productivity via $r_j$, because those sources of income comprise, on average, a small proportion of total household income across all households within the population.

[14] Given that the only information available about crime is the aggregate crime rate, it is necessary to rule out heterogeneity of criminal abilities across individuals within a city; by including $\epsilon_j$, we allow for such heterogeneity at the city level.

[15] In the implementation , $\gamma$ is further parameterized to be a deterministic function of the education distribution of the city. Specifically,

$$\gamma_j = b_0 + b_1 \ln(1 + P_j(edu = 1)),$$

where $b_0 > 0$, $b_1 \geq 0$ and $P_j(edu = 1)$ is the fraction of citizens in city $j$ without a high school degree.

mean one) as multiplicative with the crime rate implies that a city with twice the crime rate and twice the police efficiency as another city will have the same arrest rate (given the same number of police).

## 2.3 The Individual's Decision Problem

The decision of an individual in city $j$ depends on his own characteristics as well as city-level variables: the crime rate $\left(\mu_j\right)$, the human capital rental rate $(r_j)$, the aggregate labor input $(L_j)$, the arrest rate $(\pi_j)$ and the expected utility of a successful criminal $(A_j)$, where the expectation is taken over the victim's income.[16] Define the vector $\Omega_j\left(s_j\right) \equiv \left[\mu_j, r_j, L_j, \pi_j, A_j\right]$, which will vary with the number of police $(s_j)$.[17] Flow utility is assumed to be logarithmic in disposable income (consumption) and additive in the taste for crime and in the value of staying home. We also assume that criminals cannot target their victims and that they cannot steal from other criminals. Letting $d_n = [d_{n1}, d_{n2}, d_{n3}]$ be the individual's choice vector, the alternative-specific values for a type $n$ individual residing in city $j$ is given by

$$V_{nj}\left(d_n|\Omega_j\left(s_j\right)\right) = \begin{cases} \mu_j \ln\left((1-\alpha)y_{jn}\right) + (1-\mu_j)\ln(y_{jn}) \text{ if } d_{n1} = 1, \\ \mu_j \ln\left((1-\alpha)b_j\right) + (1-\mu_j)\ln(b_j) + \kappa_n \text{ if } d_{n2} = 1, \\ \mu_j \ln(b_j) + (1-\mu_j)\left[\pi_j \ln(\underline{c}) + (1-\pi_j)A_j\right] + \eta_n \text{ if } d_{n3} = 1. \end{cases} \tag{6}$$

The first (second) row in (6) shows the value if the individual chooses to work (stay home). With probability $\mu_j$, the individual is victimized and consumes $(1-\alpha)y_{jn}$ if employed or $(1-\alpha)b_j$ if at home. With probability $(1-\mu_j)$, the individual is not victimized and consumes his income. If he chooses to be at home, he also enjoys the value of staying home, $\kappa_n$. The third row of in (6) shows the value if the individual chooses to be a criminal. With probability $\mu_j$, a criminal fails to find a victim (criminals cannot be victims); in that case,

[16] $A_j$ is given below when a market equilibrium is defined.

[17] To save on notation, we have suppressed the dependence of each component in the vector on $s_j$. In addition, we include both $L_j$ and $r_j$ in the individual's state space for completeness in describing the vector of outcomes although only one is necessary for the individual's choice problem.

we assume that the criminal has the same income as a law-abiding non-worker, $b_j$. With probability $(1-\mu_j)$, a criminal meets a victim. In this case, with probability $\pi_j$, he is arrested and punished, consuming $\underline{c}$. With probability $(1-\pi_j)$, he is not arrested, and has expected utility $A_j$. Engaging in crime also directly increases utility by the value $\eta_n$, which can be negative.

### 2.3.1 Optimal Decisions

It can be shown that an individual in city $j$ with $(l_n, \eta_n, \kappa_n)$ will engage in crime if only if[18]

$$\begin{aligned} & (1-\mu_j)\left[\pi_j \ln(\underline{c}) + (1-\pi_j)A_j\right] + \mu_j \ln(\frac{b_j}{1-\alpha}) - \ln(r_j) \\ > \; & \max\left\{\ln(l_n), \ln(l_1)+\kappa_n\right\} - \eta_n. \end{aligned} \tag{7}$$

If the individual does not choose to be a criminal, the individual will choose to work if only if

$$\ln(l_n) \geq \ln(l_1) + \kappa_n. \tag{8}$$

We denote the optimal decision of an individual by $d_n(\Omega_j\left(s_j\right))$.

From condition (7), it can be seen that an individual's propensity to engage in crime can be summarized by

$$T_n \equiv \max\left\{\ln(l_n), \ln(l_1)+\kappa_n\right\} - \eta_n. \tag{9}$$

Index these propensities such that $T_n \leq T_{n+1}$, in which case the lower is $n$, the higher is one's criminal propensity. Thus, if $T_n$ type chooses to be a criminal, all $T_{n'}$ types will do so for $n' < n$.

## 2.4 Market Equilibrium

**Definition 1** *Given the size of the police force $s_j$, a market equilibrium in city $j$ consists of a vector $\widetilde{\Omega}_j\left(s_j\right) = \left[\mu_j, r_j, L_j, \pi_j, A_j\right]$, together with a set of optimal*

[18] We assume that an individual who is indifferent between legal and illegal activities will choose to be decent. This assumption is not without loss of generality given that types are discrete; however, our counterfactual experiment results are robust to this assumption.

*individual decision rules* $\{d_n(\cdot)\}$ *for* $n = 1, ..., N$ *such that*
(*a*) *for all* $n$, $d_n\left(\widetilde{\Omega}_j\left(s_j\right)\right)$ *is an optimal decision for type* $n$, *i.e., conditions* (7) *and* (8) *hold;*
(*b*) $\widetilde{\Omega}_j\left(s_j\right)$ *is consistent with individual choices where*

$$\begin{aligned}
\textit{crime rate} &: \quad \mu_j = \sum_{n=1}^{N} p_{jn} d_{n3}\left(\widetilde{\Omega}_j\left(s_j\right)\right), \qquad (10) \\
\textit{rental rate} &: \quad r_j = \tau_j \theta L_j^{\theta-1}, \\
\textit{aggregate labor} &: \quad L_j = \sum_{n=1}^{N} p_{jn} l_n d_{n1}\left(\widetilde{\Omega}_j\left(s_j\right)\right), \\
\textit{arrest rate} &: \quad \pi_j = \Pi(s_j, \mu_j, \epsilon_j), \\
\textit{crime utility}|\textit{success} &: \\
A_j &= \sum_{n=1}^{N} p_{jn} \frac{1 - d_{n3}\left(\widetilde{\Omega}_j\left(s_j\right)\right)}{1-\mu_j} \ln\left(\alpha\left[\begin{array}{c} y_{jn} d_{n1}\left(\widetilde{\Omega}_j\left(s_j\right)\right) \\ +b_j d_{n2}\left(\widetilde{\Omega}_j\left(s_j\right)\right)\end{array}\right] + b_j\right).
\end{aligned}$$

Multiple $\Omega_j\left(s_j\right)$'s can be supported as market equilibria. However, as seen from (9) ,the ranking of $T_n$ types is independent of equilibrium objects. Thus, all of the equilibria can be ordered by their equilibrium crime rates. The total number of equilibria, $N^*$, is at most equal to the total number of $T_n$ types and is bounded above by $N_\eta \times (N_\kappa + N_l - 1)$.[19] Letting $h_{jn}$ denote the measure of type $T_n$ in city $j$ and $H_{jn} = \sum_{n' \leq n} h_{jn'}$ the cumulative distribution of criminal propensities, then $\Omega_j^n\left(s_j\right)$ with $\mu_j^n = H_{jn}$ can be supported as an equilibrium if

$$T_n < (1-\mu_j^n)\left[\pi_j^n \ln(\underline{c}) + (1-\pi_j^n)A_j^n\right] + \mu_j^n \ln(\frac{b_j^n}{1-\alpha}) - \ln(r_j^n) \leq T_{n+1}, \quad (11)$$

where the superscript $n$ indexes the $n^{th}$ potential equilibrium. In equilibrium $\widetilde{\Omega}_j^n\left(s_j\right)$, an individual will choose to be a criminal if and only if the individual's criminal propensity is ranked among the top $n$ groups. The fact that there are at most $N^*$ equilibria and that they can be ordered, as given by (11), allows

[19] Note that $N^*$ is parameter-dependent. See online Appendix B1 for details about $N^*$ and $h_{jn}$.

us to compute all of the market equilibria given $s_j$.[20]

## 2.5 Government Problem

A government cares about the crime rate in the city and can affect the level of criminal activity by choosing the size of the police force. In addition, a government may also care directly about the arrest rate for political reasons; a government without the ability to catch criminals could be considered as ineffectual in combatting crime. Therefore, the government's loss function will be positively related to the crime rate and, given its cost, to the size of the police force and negatively related to the arrest rate. The government is assumed to minimize its expected loss, where the expectation is taken over all possible market equilibria. Formally, the government in city $j$ solves the following problem

$$\min_{s_j} \left\{ \sum_{n=1}^{N^*} q_j^n(s_j) \left( \omega_1 \exp(\mu_j^n) - \omega_2 \ln(\pi_j^n) + \nu_j s_j \right) \right\}, \tag{12}$$

where $\omega_1$ and $\omega_2$ are the weights governments put on the crime and arrest rates relative to the cost of the police force. $\nu_j \sim \ln N(-0.5\sigma_\nu^2, \sigma_\nu^2)$ is the city-specific marginal (opportunity) cost of the police force.[21] $q_j^n(s_j)$ is the probability that $\Omega_j^n(s_j)$ is realized as a market equilibrium, to be specified in Section 4. Although given any $s_j$ there may be multiple market equilibria, the government optimal choice is generically unique.

## 2.6 Subgame Perfect Nash Equilibrium

**Definition 2** *A subgame perfect Nash equilibrium in city j is* $\left\{ s_j^*, d(\cdot), \left\{ \widetilde{\Omega}_j^n(\cdot) \right\} \right\}$ *such that*
*(a) Given any* $s_j$, $\left( d(\cdot), \widetilde{\Omega}_j^n(s_j) \right)$ *is a market equilibrium;*
*(b)* $s_j^*$ *solves the government's problem.*

[20] If individuals' crimininal propensities cannot be ranked independent of equilibrium objects, there will be $2^N$ potential market equilibria; and it will be infeasible to compute all possible equilibria at every candidate parameter configuration during the estimation.

[21] For example, cities may differ in their needs for resources in areas other than crime prevention.

# 3 Data

We make use of data from two sources. One source is the Current Population Survey (CPS), which provides micro-level data on demographics ($x$), wages and employment. We focus on the population aged 16 to 64.[22] We aggregate individuals within a metropolitan statistical area (MSA), which is the counterpart of a "city" in our model. The distribution of $x$ within an MSA is taken as the "city-specific" distribution of $x$ in our model. We define an individual as employed if he/she works for more than 13 weeks during the year, and define the employment rate as the fraction of those employed among the age 16-64 population. For those who are employed, we also use information on their annual earnings.

The other data source is the Uniform Crime Reports (UCR), which contains agency-level reports on crimes, arrests, number of police, and population size. We focus on property crimes, which include robbery, burglary, larceny-theft and motor vehicle theft.[23] Within each MSA, we aggregate agencies with non-missing reports and treat them as representative of the entire MSA. For each MSA, we define the crime rate as the total number of actual crimes divided by the total 16-64 population covered by non-missing agencies, the arrest rate as the total number of arrests divided by the total number of actual crimes and the size of police force as the total number of police divided by the total 16-64 population covered by non-missing agencies.

For both the CPS and the UCR, we focus on the year 2008. We are able to match 245 MSA's between the two data sets. We exclude 7 MSA's with zero arrests as extreme outliers, which results in a final sample consisting of 238 MSA's and 86,248 individuals living in these MSA's. Table 1 summarizes the MSA-specific marginal distributions of individual characteristics. The first (second) row gives the mean (standard deviation) of the within-MSA marginal distribution over the 238 MSA's. As seen, MSA's are more diverse in their

[22] The age restriction is consistent with the literature, for example, Imrohoroglu et. al. (2000, 2004).

[23] Note that our definition of property crimes is consistent with the literature but different from the UCR definition, which doesn't include robbery.

educational and racial compositions than in their age or gender composition. For example, the coefficient of variation across MSA's in the fraction of people without a high school degree is 0.43, but only 0.23 for the fraction of residents under the age of 25.

Table 1 Summary Statistics: $x$ Distribution

| | Education | | Age | | Race | Gender |
|---|---|---|---|---|---|---|
| % | Less than HS | BA and Above | Under 25 | Above 50 | Black/Hispanic | Male |
| Mean[a] | 16.08 | 24.32 | 21.43 | 25.75 | 24.07 | 49.32 |
| Std. Dev.[a] | 6.99 | 9.34 | 5.03 | 6.78 | 18.70 | 3.80 |

Number of Obs: 238 MSA's in the U.S.

[a]Cross-MSA mean and std deviation of various marginal distributions.

Table 2 reports statistics from the data on the equilibrium outcomes of the model. Specifically, it shows the cross-MSA mean, standard deviation and coefficient of variation (CV) of the crime rate (per 1,000), the arrest rate (percent), the number of police per 1,000 people, the employment rate (percent) and mean earnings. As seen, the mean of the crime rate across the MSA's is 56.5 with a standard deviation of 17.0. As reflected by the CV, the variation in the crime rate is similar to that of the arrest rate (.36) and of the number of police (.40). The labor market outcomes, on the other hand, exhibit less variation across MSA's; the CV in the employment rate is 0.09 and that in earnings 0.19.

Table 2 Summary Statistics: Outcomes

| | Crime (per 1,000) | Arrest (%) | Police (per 1,000) | Employment (%) | Mean Earnings[b] ($) |
|---|---|---|---|---|---|
| Mean[a] | 56.52 | 19.03 | 4.78 | 73.34 | 38,142 |
| Std Dev.[a] | 16.96 | 6.85 | 1.91 | 6.87 | 7,340 |
| CV[a] | 0.30 | 0.36 | 0.40 | 0.09 | 0.19 |

Number of Obs: 238 MSA's in the U.S.

[a]Cross-MSA mean, std deviation and coefficient of variation of variable in each column.

[b]Within-MSA mean earnings.

# 4 Empirical Implementation and Estimation

## 4.1 Additional Empirical Specifications

### 4.1.1 Distribution of Individual Types

To implement the model, we need to specify the joint distribution of the unobservables $(l_n, \eta_n, \kappa_n)$, that is, $p_{jn}$. Conditional on $x$, the distributions of $l$, $\eta$ and $\kappa$ are assumed to be independent, so $p_{jn}$ is the product of the marginal distributions of its components. The means of the distributions of $l$ and $\eta$ are (log)linear in $x$ and are assumed to be city-independent.[24] The distribution of human capital is given by[25]

$$p_{l_m}(x) = \begin{cases} \Phi(\frac{\ln(l_m)-x'\lambda^l}{\sigma_l}) - \Phi(\frac{\ln(l_{m-1})-x'\lambda^l}{\sigma_l}) & \text{for } 1 < m < N_l, \\ \Phi(\frac{\ln(l_m)-x'\lambda^l}{\sigma_l}) & \text{for } m = 1, \\ 1 - \Phi(\frac{\ln(l_{m-1})-x'\lambda^l}{\sigma_l}) & \text{for } m = N_l. \end{cases} \tag{13}$$

The mass points $l's$ are assumed to be quantiles from $\ln N(\overline{x}'\lambda^l, \sigma_l^2)$, where $\overline{x}$ is the mean of $x$.

The distribution of the preference for engaging in crime is given by

$$p_{\eta_m}(x) = \begin{cases} \Phi(\frac{\eta_m-x'\lambda^\eta}{\sigma_\eta}) - \Phi(\frac{\eta_{m-1}-x'\lambda^\eta}{\sigma_\eta}) & \text{for } 1 < m < N_\eta, \\ \Phi(\frac{\eta_m-x'\lambda^\eta}{\sigma_\eta}) & \text{for } m = 1, \\ 1 - \Phi(\frac{\eta_{m-1}-x'\lambda^\eta}{\sigma_\eta}) & \text{for } m = N_\eta. \end{cases} \tag{14}$$

It can be shown that there exists an $\eta^*$ such that people with $\eta \leq \eta^*$ will never commit a crime independent of equilibrium outcomes.[26] We therefore set the lowest $\eta$, $\eta_1 = \eta^*$. We also assume that the largest $\eta$ $(\eta_{N_\eta})$ is such that individuals with $\eta_{N_\eta}$ will always commit crimes.[27] The other mass points $\eta$'s

[24] We set $N_l = 20$, $N_\eta = 10$ and $N_\kappa = 10$ such that there are 2000 types of individuals.

[25] Heterogeneity in wages across cities arises from variation in $x's$ and from heterogeneity in human capital rental prices. Recall that we allow productivity ($\tau$) to differ across cities. Given that we use only cross-sectional variation in wages, it would be difficult to identify city-specific unobservables in both human capital levels and rental prices.

[26] As is shown in online Appendix B1.1, $\eta^* = \ln(l_1) + \kappa_1 + \ln(1-\alpha) - \ln(\alpha l_{N_l} + l_1)$.

[27] A zero-crime equilibrium is thus never possible.

are assumed to be quantiles from the distribution $N(\overline{x}'\lambda^\eta, \sigma_\eta^2)$ that are above $\eta^*$.

The mean of the distribution of the value of the home option ($\kappa$) is log-linear in $x$, and is assumed to be city-specific. In city $j$, the distribution of $\kappa$ is given by

$$p_{\kappa_m}(x,j) = \begin{cases} \Phi(\frac{\ln(\kappa_m)-\overline{\kappa}_j - x'\lambda^\kappa}{\sigma_\kappa}) - \Phi(\frac{\ln(\kappa_{m-1})-\overline{\kappa}_j - x'\lambda^\kappa}{\sigma_\kappa}) & \text{for } 1 < m < N_\kappa, \\ \Phi(\frac{\ln(\kappa_m)-\overline{\kappa}_j - x'\lambda^\kappa}{\sigma_\kappa}) & \text{for } m = 1, \\ 1 - \Phi(\frac{\ln(\kappa_{m-1})-\overline{\kappa}_j - x'\lambda^\kappa}{\sigma_\kappa}) & \text{for } m = N_\kappa, \end{cases} \tag{15}$$

where the mass points $\kappa$'s are quantiles from the distribution $\ln N(\overline{x}'\lambda^\kappa, \sigma_\kappa^2)$. The city-specific component $\overline{\kappa}_j$ is assumed to be an i.i.d. random variable across cities, drawn from the distribution $N(0, \sigma_{\overline{\kappa}}^2)$.

#### 4.1.2 Probability Distribution of Market Equilibria

Given the size of police force ($s_j$), $q_j^n(s_j) = 0$ if $\Omega_j^n(s_j)$ is not supportable as a market equilibrium. As shown, all the supportable $\Omega_j^n(s_j)$'s can be ranked by their crime rates. Based on this fact, we assume that the probability a particular market equilibrium is realized depends on the ranking of its crime rate in the set of equilibria. It is not clear a priori whether a low-crime or a high-crime equilibrium is more likely, and we use the following structure that allows for various possible scenarios.[28]

Let $N_j^*(s_j)$ be the number of market equilibria in city $j$ under $s_j$, and $\left\{\widetilde{\Omega}_j^{n'}(s_j)\right\}_{n'=1}^{N_j^*(s_j)}$ the set of these equilibria ranked from low crime rate to high crime rate. The probability of the $n^{th}$ element in $\left\{\widetilde{\Omega}_j^{n'}(s_j)\right\}_{n'=1}^{N_j^*(s_j)}$, i.e., the

[28]There are a number of papers that have also estimated equilibrium selection rules jointly with the other model parameters. Examples of this approach include Ackerberg and Gowrisankaran (2006), Bajari, Hong and Ryan (2010), Bjorn and Vuong (1984), Card and Giuliano (forthcoming), and Jia (2008).

probability that the $n^{th}$-ranked equilibrium $\widetilde{\Omega}_j^n(s_j)$ is realized is given by

$$q_j^n(s_j) = \frac{\exp\left[1+\zeta_1\left(n_{j1}^*-n\right)I\left(n<n_{j1}^*\right)+\zeta_2\left(n-n_{j2}^*\right)I\left(n>n_{j2}^*\right)\right]}{\sum_{n'=1}^{N_j^*(s_j)}\exp\left[1+\zeta_1\left(n_{j1}^*-n'\right)I\left(n'<n_{j1}^*\right)+\zeta_2\left(n'-n_{j2}^*\right)I\left(n'>n_{j2}^*\right)\right]},$$

where $I(\cdot)$ is the indicator function. If $N_j^*(s_j)$ is odd, $n_{j1}^*(s_j) = n_{j2}^*(s_j) = \frac{N_j^*(s_j)+1}{2}$, that is, the median of $\left\{1,...,N_j^*(s_j)\right\}$. If $N_j^*(s_j)$ is even, $n_{j1}^*(s_j) = \frac{N_j^*(s_j)}{2}+1$, the first number to the right of the median, and $n_{j2}^*(s_j) = \frac{N_j^*(s_j)}{2} - 1$, the first number to the left of the median. The two parameters, $\zeta_1$ and $\zeta_2$, capture the relationship between the probability of an equilibrium and its ranking. For example, if $\zeta_1 = \zeta_2 = 0$, all equilibria are equally likely; if $\zeta_1 > 0$, $\zeta_2 > 0$, the distribution of equilibria is U-shaped; if $\zeta_1 < 0$, $\zeta_2 < 0$, the distribution of equilibria is inverse-U-shaped; if $\zeta_1 > 0$, $\zeta_2 < 0$, lower crime-rate equilibria are more likely, and if $\zeta_1 < 0$, $\zeta_2 > 0$, higher crime-rate equilibria are more likely.[29]

## 4.2 Estimation

We estimate the model using simulated generalized method of moments (SGMM). For each parameter configuration, we solve for the equilibria of the model, compute the model-predicted moments for each equilibrium and integrate over them. The parameter estimates minimize the distance between the model-predicted moments $(M(\Theta))$ and the data moments $\left(M^d\right)$:

$$\widehat{\Theta} = \arg\min_{\Theta}\left\{\left(M(\Theta)-M^d\right)' W\left(M(\Theta)-M^d\right)\right\},$$

where $\Theta$ is the vector of structural parameters, and $W$ is a positive-definite weighting matrix.[30] $\Theta$ includes parameters governing the distributions over individuals in their human capital and tastes for crime and for home, the

[29] When one and only one $\zeta$ is zero, then all equilibria below or above the median are equally likely and the probabilities for others are monotonically increasing or decreasing with their rank depending on the sign of the other $\zeta$.

[30] In particular, $W$ is a diagonal matrix, the $(k,k)^{th}$ component of which is the inverse of the variance of the $k^{th}$ moment, estimated from the data.

distributions of city-level unobservables, the arrest technology, the production technology, the return to crime, the consumption level if arrested, government preferences and the probability distribution of market equilibria.[31]

The estimation routine involves an outer loop searching over the parameter space, and an inner loop determining the set of equilibria. For each of the 238 MSA's, we simulate $R$ replicas that copy its distribution of observable characteristics $(x)$. Each of the simulated $238R$ replicas serves as the counterpart of a city in our model. The outer loop of the estimation procedure uses a simplex routine. The following describes the inner loop.

Step 1: Given a set of parameters $\Theta$, calculate the mass points of $l$ and $\kappa$ as the quantiles from $\ln N(\overline{x}'\lambda^l, \sigma_l^2)$ and $\ln N(\overline{x}'\lambda^\kappa, \sigma_\kappa^2)$, respectively. Calculate the first mass point $\eta_1 = \eta^*$ as defined in online Appendix B1.1, and other mass points $\eta$ as quantiles from $N(\overline{x}'\lambda^\eta, \sigma_\eta^2)$ that are above $\eta^*$. Calculate the criminal propensity $T_n$ according to equation (9), and the number of $T_n$ types $N^*$ as in online Appendix B1.2. Index $T_n$ such that $T_n \leq T_{n+1}$.

Step 2: For each city $j$, draw the city-level unobservable characteristics $(\nu_j, \tau_j, \epsilon_j, \overline{\kappa}_j)$. For each vector of observable characteristics $x$, which is assumed to be discrete, calculate the probability vectors $p_l(x)$, $p_\eta(x)$ and $p_\kappa(x, j)$, according to equations (13) to (15). Given the distribution of observable citizen characteristics $G_j(x)$, calculate the measure of each $(l, \eta, \kappa)$-type in city $j$. Derive the measure $(h_{jn})$ of each $T_n$-type in city $j$ according to online Appendix B1.3.

Step 3: Pick an $s_j$ from the grid for the size of police force and solve for all market equilibria.[32] For each $n \in \{1, ..., N^*\}$, suppose all $T_{n'}$ are criminals for

[31] Given the nonlinear nature of the model, we are not able to provide a constructive identification argument. However, to provide some evidence on identification, we have conducted Monte Carlo exercises in which we first simulated data with known parameter values, treated as the "truth" and then, using moments from the simulated data, started the estimation of the model from a wide range of initial guesses of parameter values. In all cases, we were able to recover parameter values that are close to the "truth."

[32] The grid points for $s_j$ are parameter-city-specific. They are bounded between $[0, \infty)$, and they always include the points $\{s_j^n\}_{n=1}^{N^*}$ such that

$$
\begin{aligned}
T_{n+1} &= (1-\mu_j^n)\left[\pi_j^n \ln(\underline{c}) + (1-\pi_j^n)A_j^n\right] + \mu_j^n \ln\left(\frac{b_j^n}{1-\alpha}\right) - \ln(r_j^n), \\
s.t.\ \pi_j^n &= \Pi\left(s_j^n, \mu_j^n, \gamma_j, \epsilon_j\right).
\end{aligned}
$$

$n' \leq n$, that is, the crime rate $\mu_j^n = \sum_{n' \leq n} h_{jn'}$ and for any $n' > n$, the $T_{n'}$-type chooses between work and home according to condition $(8)$, which implies the aggregate human capital employed $L_j^n$ as in $(10)$. Derive the rest of the components of $\Omega_j^n(s_j)$, the arrest rate as in equation $(5)$, $A_j^n$ as in equation $(10)$, and $r_j^n$ as in equation $(4)$. Calculate the value of the middle term in $(11)$. $\Omega_j^n(s_j)$ is an equilibrium if only if inequality (11) is satisfied.

Step 4: Calculate the government cost under $s_j$ by integrating over potential equilibria, as in (12).

Repeat Steps 3-4 until the optimal size of police force $s_j^*$ and the associated set of market equilibria $\left\{\widetilde{\Omega}_j^n\left(s_j^*\right)\right\}$ are found, and do this for every city $j \in \{1, ..., 238R\}$.

Step 5: Calculate the model predicted moments as

$$\frac{1}{238R} \sum_j \sum_{n=1}^{N^*} q_j^n\left(s_j^*\right) M_j^n\left(s_j^*; \Theta\right),$$

where $M_j^n\left(s_j^*; \Theta\right)$ is the vector of model predicted statistics in city $j$ if $\Omega_j^n\left(s_j^*\right)$ is a market equilibrium.

#### 4.2.1 Target Moments

We target 171 moments that include, among others,

1. A set of unconditional moments with each MSA as an observation:

1) First moments across MSA's of the number of police, crime rate, arrest rate, employment rate, and the within-MSA average and standard deviation of earnings;

2) Cross moments of the variables in 1), except that between mean earnings and the standard deviation of earnings;

3) Second moments of the first 5 outcome variables.

4) The fractions of cities with crime rates below the $10^{th}$, $20^{th}$,...,$90^{th}$ percentiles

That is, $s_j^n$ is the level of police force that makes type $T_{n+1}$ at the margin between crime and decency (and choose decency by assumption), and thus supports the equilibrium with $\mu_j^n$ exactly. Note that $s_j^n$ may also support other equilibria, to be determined in Step 3.

of crime rates in the data.

2. First moments of MSA outcomes by MSA characteristics.[33]

The average size of the police force, the average crime rate and the average arrest rate conditional on the within-MSA marginal distributions of age, gender, education and race, all treated as discrete variables. For example, we target the average outcomes among MSA's where the fraction of college graduates is ranked below the 50th percentile among all MSA's.

3. First moments of individual outcomes by individual characteristics.

1) The employment rate and average earnings among individuals by age, by education, by gender, by race, by the number of kids and by the number of kids among females.

2) It is well documented that the crime rate is significantly higher among youths. However, our data do not contain age-specific crime rates. To extract as much information as possible about the criminal versus non-criminal choices, we also use data from the CPS on school enrollment status. Specifically, we target, within this age group (less than age 25), their employment rate and school attendance rate by current education attainment.[34]

[33] To keep the number of observations constant across moments that use MSA's as units of observations, we target joint moments of characteristics and outcome instead of conditional moments. For example, letting $I(\cdot)$ be the indicator function, we target $E(y_{jm}, I(G_j(x) \leq z))$ instead of $E(y_{jm}|G_j(x) \leq z)$, where $y_{jm}$ is the $m^{th}$ outcome, $G_j(x)$ is the distribution of within-MSA demographics, and $z$ is some quantile. Similarly, in the next (the third) set of moments, which use individuals as units of observations, we target joint moments of individual outcome and individual characteristics, instead of conditional moments to keep the number of observations constant across this set of moments.

[34] To incorporate school enrollment, we assume that, for people in the youngest age group, the home option includes attending school. With a probability that is specific to one's current education level, a youth who chooses the home option is enrolled in school. These probabilities are treated as parameters to be estimated jointly with other parameters in the model. For people of older ages, we restrict the home option to not include attending school, given that school enrollment rates are low for them..

### 4.3 Parameter Estimates

Table 3 shows selected parameter estimates; the others are shown in the appendix. Standard errors (in parentheses) are calculated via bootstrap.[35] As shown in the upper panel of Table 3, the elasticity of output with respect to aggregate human capital ($\theta$) in the legal sector is 0.84. The value of $\alpha$ implies that a criminal steals ten percent of the victim's income, or approximately \$3,250.[36] However, if apprehended, consumption is only \$900.[37]

Table 3 Selected Parameter Estimates

| | | |
|---|---|---|
| Production $\theta$ | 0.84 | (0.01) |
| Return to Crime $\alpha$ | 0.10 | (0.003) |
| Consumption if Arrested $\underline{c}$ | 0.09 | (0.003) |
| Dispersion of MSA-Level Unobservables | | |
| Productivity $\sigma_\tau$ | 0.12 | (0.09) |
| Arrest Efficiency $\sigma_\epsilon$ | 5.89 | (0.03) |
| Value of Leisure $\sigma_{\overline{\kappa}}$ | 0.13 | (0.07) |
| Marginal Cost of Police $\sigma_\nu$ | 0.25 | (0.10) |
| Probability Distribution of Market Equilibria | | |
| Below Median Slope $\zeta_1$ | -0.20 | (0.18) |
| Above Median Slope $\zeta_2$ | -0.60 | (0.24) |

The middle panel of Table 3 shows the standard deviations of each of the four location-specific unobservables. To understand their magnitudes, we have explored four counterfactual scenarios, where we improve, one at a time, the productivity, the arrest efficiency, the value of leisure and the marginal cost of police for every location by one standard deviation. We then calculate the percentage changes in outcomes at the national level as compared to the baseline.

[35] There are 160 bootstrap iterations. In each iteration, we sample 238 MSA's from the original sample of 238 MSA's with replacement. All individuals in the matched CPS data are included who reside in MSA's in the bootstrap sample.

[36] If they could choose only between work and home, i.e., criminal activity were not an option, only 31% of the current criminals would choose work over home, and their average income would be \$2,550. Obviously, criminals come from the lower tail of the income distribution.

[37] All parameters in monetary units are in units of \$10,000.

Without changing the number of police, the national crime rate is reduced by 6.3, 17.7 and 8.3 percent when productivity, arrest efficiency and the value of leisure are individually increased by one standard deviation above their current levels in all MSA's (see Appendix Table A1). Allowing for the government to optimally adjust the size of their police force leads to reductions in the number of police and thus smaller decreases in the crime rate of 2.8, 11.9 and 3.6 percent. Finally, if the marginal cost of police is reduced by one standard deviation, the national police force increases by 14.8 percent leading to a 7.8 percent reduction in crime.[38]

The lower panel of Table 3 shows the slope parameters that govern the distribution of market equilibria. Both parameters are estimated to be negative, although $\zeta_1$ is not precisely estimated. In a joint Wald test, the p-value for the joint test that both of $\zeta$'s are zero is 0.023. The point estimates suggest that for a given city's size of police force, the distribution of market equilibria is inverse-U-shaped. That is, the equilibrium whose crime rate is ranked in the middle is most likely to be realized. Moreover, as $|\zeta_1| < |\zeta_2|$, the distribution is asymmetric: given the same ranking distance from the median, the high crime rate equilibrium is less likely than the low crime rate equilibrium.[39]

Appendix Tables A2-A4 show the parameters governing the distributions of values of the home option, levels of human capital, and tastes for crime conditional on $x$. Education is positively correlated with one's human capital and value of leisure. With the presence of young children, the value of leisure decreases (increases) for males (females), which is consistent with the fact that the employment rate for males (females) with children is higher (lower) when there are children in the household. The preference for engaging in crime is lower for those who are more educated, older and/or have young children.[40]

[38] These effects, as well as those that follow, are obtained by integrating over market equilibria.

[39] For example, if there are 3 market equilibria ranked by their crime rates from low to high, then the probability distribution will be $\{0.35, 0.42, 0.23\}$. In the case of 4 market equilibria, the distribution will be $\{0.29, 0.35, 0.23, 0.13\}$.

[40] Appendix Table A5 shows the probabilities of attending school among youth who choose the home option. Table A6 shows estimates for government preferences and other parameters in the arrest technology.

## 4.4 Model Fit

Table 4 shows the predicted means of various endogenous outcomes across MSA's. As seen, they closely match the data. The fit to the standard deviations of outcomes across MSA's, as shown in Table 5, is generally close but not as good.

Table 4 Means

| | Crime (%) | Arrest (%) | Police (%) | Employment (%) | Mean Wage ($) | Std Dev Wage ($) |
|---|---|---|---|---|---|---|
| Data | 5.65 | 19.03 | 0.48 | 73.33 | 38,142 | 40,067 |
| Model | 5.63 | 19.01 | 0.47 | 73.82 | 39,551 | 39,598 |

Table 5 Standard Deviations

| | Crime (%) | Arrest (%) | Police (%) | Employment (%) | Mean Wage ($) |
|---|---|---|---|---|---|
| Data | 1.70 | 6.85 | 0.19 | 6.69 | 7,340 |
| Model | 1.53 | 7.79 | 0.12 | 4.34 | 5,996 |

Model predictions of correlations between pairs of outcomes can be calculated from the predicted first, second and cross moments. Table 6 contrasts these correlations with the data. Overall, the model correctly predicts the signs of these correlations. For example, the correlation between the number of police and the crime rate is positive both in the data and for the model prediction. The positive correlation arises from two offsetting effects. On the one hand, having additional police leads to a reduction in crime through an increase in the apprehension rate. On the other, governments in MSA's that can be expected to be more prone to crime because of the distribution of citizen characteristics and/or MSA-level characteristics will optimally choose a larger police force. When the model and the data diverge in sign, between police size and the wage and between the arrest rate and the wage, the correlations in both the data and the model are weak.

Table 6 Correlations

| | Data | Model |
|---|---|---|
| (Police, Crime) | 0.35 | 0.17 |
| (Police, Arrest) | -0.08 | -0.04 |
| (Police, Employment) | -0.18 | -0.10 |
| (Police, Average Wage) | 0.18 | -0.09 |
| (Employment, Crime) | -0.30 | -0.38 |
| (Employment, Arrest) | 0.10 | 0.22 |
| (Employment, Average Wage) | 0.06 | 0.18 |
| (Crime, Arrest) | -0.20 | -0.45 |
| (Crime, Average Wage) | -0.16 | -0.31 |
| (Arrest, Average Wage) | -0.11 | 0.01 |

The model fits well the relationships between the crime rate and the within-MSA distribution of demographics. For example, the correlation between the crime rate and the fraction of blacks/Hispanics (not shown in the tables) is 0.41 in the data and 0.40 in the model.[41]

# 5 Counterfactual Experiments

The counterfactual experiments are motivated by federal programs that have provided subsidies to local governments for increasing the number of police. For example, the COPS program, initiated by the Clinton administration in 1994, planned to add 100,000 (20 percent) more police nationwide by fiscal year 2000, as part of the federal government's effort to reduce crime.[42]

[41] In the online appendix, Figure 1 shows the correlation between the crime rate and the distribution of education levels; Table B1 shows the fit of employment rates and earnings by individual characteristics.

[42] Since 1994, COPS has provided $11.3 billion in assistance to state and local law enforcement agencies to help in hiring additional police officers.

## 5.1 Increasing the Police Force Nationwide

The first counterfactual experiment considers scenarios where the planner (federal government) aims at an increase of 20 percent in the total number of police through a subsidy to local governments and where the use of the subsidy is monitored. First, we consider a scenario where the program is "uniform" in the sense that the size of the police force is increased by 20% of its current size in each MSA. Second, we consider three scenarios where the planner subsidizes a certain percentage of the newly hired police. However, the number of new hires in each MSA is chosen by each local government, that is, the take-up of the subsidy is voluntary. The three scenarios differ in the subsidy rate.

Table 7 shows the outcomes of these four scenarios at the national level. The first row shows the percentage of MSA's that experience an increase in their police force. Rows 2-4 show the change nationwide in the total police force, the change in the crime rate and the implied elasticity of the crime rate with respect to the number of police. Row 5 shows the cost of the policy, measured by the number of police per 1,000 people funded by the program. The last row shows the efficiency of each policy, as measured by the ratio of the reduction in crime to the cost ("efficiency ratio").

Under a uniform policy (column 1), the police force increases by 20 percent and the crime rate falls by 8.23 percent. The planner funds an additional 0.94 police officers per 1,000 people and for each unit of federal resources, as measured by the number of additional police per 1,000 people, there is a 8.8 percent reduction in crime, that is, the efficiency ratio is 8.8.

The next set of scenarios require that the subsidy rate lead to outcomes that satisfy three different constraints. Because the hiring decisions of local governments vary with the subsidy policy, we solve for the desired subsidy rate by simulation in each case. We simulate equilibrium outcomes for a given subsidy rate and repeat the process until we find the subsidy rate that satisfies our constraint. In the first scenario (column 2), we find the subsidy rate such that the total cost to the planner is the same as in column 1, that is, the employment of 0.94 additional police per 1,000 people. That rate turns out to be 47.7 percent. At that subsidy rate, over 92 percent of local governments accept the subsidy and increase their police force. The number of police increases by 42 percent nationwide, considerably greater than the target increase of 20 percent, and the national crime rate decreases by 21.2 percent. The efficiency ratio is 22.7, more than twice that of the uniform policy.

Table 7 National Outcomes

| | Uniform | Subsidy 1 | Subsidy 2 | Subsidy 3 |
|---|---|---|---|---|
| % MSA with Increase in Police | 100 | 92.27 | 65.84 | 77.77 |
| % Change in Total Police | 20 | 41.89 | 20 | 25.57 |
| % Change in Crime Rate | -8.23 | -21.19 | -10.60 | -13.44 |
| Elasticity Crime to Police | -0.41 | -0.51 | -0.53 | -0.53 |
| Cost (Federally-funded police per 1000 people) | 0.94 | 0.94 | 0.25 | 0.40 |
| $\frac{\text{\% Reduction in Crime Rate}}{\text{Cost}}$ | 8.80 | 22.65 | 41.92 | 33.40 |

In the next scenario, we find the subsidy rate such that the nationwide increase in total police is the same as the planner intended, that is, 20 percent. That subsidy rate turns out to be 27 percent, the cost to the planner is only 0.25 police officers per 1,000 people and each unit of federal resources leads to an almost 42 percent decrease in crime. Notice that the total costs, federal and local, are the same in this scenario as in the scenario with the uniform 20 percent nationwide increase in the police force. However, in this case there is a more significant reduction in crime (10.6 vs. 8.2 percent). With monitoring, as is assumed in both cases, allowing local governments to choose their own police force, instead of a uniformly imposed police increase, leads to larger reduction in crime nationwide and an efficiency ratio that is almost 5 times as large.

In the last scenario, we consider the effects of a COPS-like budget on crime reduction. We first calculate the intended cost of the COPS program. With the 75 percent COPS subsidy rate and a goal of 100,000 new police, the cost is equivalent to 0.4 police funded per 1,000 people in the year 2000. The subsidy rate that satisfies this budget constraint is 34 percent. About 78 percent of local governments take the offer, leading to a 26 percent increase in police nationwide, higher than the 20 percent increase set as a goal by COPS. The crime rate is reduced by 13 percent and the efficiency ratio in this case is over 33. This ratio is what one could expect from a COPS-like program if the federal government were able to monitor local governments perfectly so that the grants are used exclusively to add new police on net. That is, this scenario provides an upper bound on the effects of such programs.[43]

The effects of subsidies can differ substantially across locations. For example, Table 8 shows the standard deviations and coefficients of variation of various outcomes across MSA's under the two scenarios represented in columns 1 and 3 of Table 7, where the total increases in police are both set at 20 percent. In the first column, although the percentage change in police is uniform across MSA's, the reduction in crime is dispersed with a coefficient of variation (CV) in the crime rate of about 0.5. In column 2, the CVs of both the increase in police and the change in crime are not far from one.

Table 8 Distribution of Outcomes Across MSA's

| | Uniform | | Subsidy 2 | |
|---|---|---|---|---|
| | Std Dev | CV | Std Dev | CV |
| % Change in Police | 0 | 0 | 20.17 | 0.98 |
| % Change in Crime Rate | 3.86 | 0.46 | 8.79 | 0.87 |

## 5.2 Targeted Allocation of the Police Force

Given that the effects of additional police on crime differ significantly across MSA's, determining the efficient allocation of extra resources across MSA's becomes a nontrivial task, which is the issue we now address. In the following counterfactuals, we consider different schemes to allocate additional police across MSA's. We illustrate the idea by focusing on pair-wise allocations, where the total additional resources to be allocated between a pair of MSA's (with monitoring) is equivalent to 20 percent of their current total police force. We consider the following five allocation rules in order to reduce the overall crime

[43] Evans and Owens (2007) find that each officer funded by the COPS grant led to a 0.7 increase in police force, suggesting that the use of the COPS grant was imperfectly monitored. They also find a 10 percent decrease in property crime for each additional police officer per 1,000 people, that is, an efficiency ratio of 10.

rate between the pair.[44]

Rule 1 (Complete Information): the planner has perfect information on both the MSA-specific observable and unobservable characteristics, and searches for the best allocation that minimizes the overall crime rate between the two MSA's.[45]

Rule 2 (Based on $x$): the planner has information only on the observable characteristics of each MSA and searches for the best allocation that minimizes the expected overall crime rate, where the expectation is taken over MSA-specific unobservables.

Rule 3 (Based on current crime rate): allocate police according to the ratio of the population-size-adjusted crime rates, such that the high-crime-rate MSA receives more extra police per capita. Concretely, letting $pop_j$ and $\mu_j$ be the population size and crime rate in MSA$_j$, the ratio of the police allocation across the two MSA's is given by $\frac{p_1}{p_2} = \frac{\mu_1 pop_1}{\mu_2 pop_2}$.

Rule 4 (Based on current arrest rate): allocate police according to the size-adjusted inverse arrest rate ratio, such that the low-arrest-rate MSA gets more extra police per capita. Let $\pi_j$ be the arrest rate in MSA$_j$, the ratio of police allocation is given by $\frac{p_1}{p_2} = \frac{\pi_2 pop_1}{\pi_1 pop_2}$.

Rule 5 (Based on current GDP): allocate police according to the size-adjusted inverse GDP per capita ratio, such that the low-GDP-per-capita MSA gets more extra police per capita. Let $I_j$ be GDP per capita in MSA$_j$, the ratio of police allocation is given by $\frac{p_1}{p_2} = \frac{I_2 pop_1}{I_1 pop_2}$.

Although they may not be optimal, the last four rules can be seen as reasonable and realistic. For example, Evans and Owens (2007) find that the COPS grants received by local governments are positively correlated with the local crime rate, the population size, the fraction of youths and the fraction of blacks, the latter two being components of $x$ in our model.

As an example, we consider the allocation problem between Philadelphia and Phoenix-Mesa-Scottsdale. Table 9 shows the demographic characteristics of these two MSA's. The first column shows the population size of each MSA as a

[44] Our exercise can be easily adapted to other objectives of the planner.

[45] Information on observable characteristics refers to the joint distribution of the whole vector of $x$'s, not just the marginal distribution of its components.

percentage of the total population size of the 238 MSA's in our data. Compared to Phoenix, educational attainment is higher and the fraction of black and/or Hispanics is lower in Philadelphia; the age and gender composition in these two MSA's are similar.

Table 9 MSA Characteristics

| | Population | Education | | Age | | Race | Gender |
|---|---|---|---|---|---|---|---|
| (%) | | Less than HS | BA or Above | $\leq$25 | >50 | Black/Hisp. | Male |
| Philadelphia | 2.38 | 11.45 | 31.43 | 18.91 | 27.16 | 26.01 | 47.66 |
| Phoenix | 2.01 | 19.61 | 25.34 | 19.69 | 26.09 | 32.89 | 50.12 |

To implement Rule 1 (complete information), we first uncover the unobserved MSA-specific characteristics, based on our parameter estimates, by searching over the space of unobservable characteristics such that the distance between the model predicted and data outcomes is minimized MSA by MSA. In particular, for each MSA, we seek the combinations of MSA-specific arrest efficiency ($\epsilon$), mean leisure value ($\overline{\kappa}$), marginal cost of police ($\nu$) and productivity ($\tau$) to minimize the percentage model-data discrepancy over the crime rate, arrest rate, police force, employment rate and average wage. Table 10 shows the outcomes predicted by our baseline model with the sets of discrepancy-minimizing unobservables, which are closely matched with the data. Philadelphia has fewer police per-capita, a lower crime rate, a higher arrest rate and a larger per-capita GDP than does Phoenix.

Table 10 Current Outcomes (Baseline Model)

| | Police (per 1,000) | Crime (%) | Arrest (%) | GDP per capita ($) |
|---|---|---|---|---|
| Philadelphia | 4.66 | 5.19 | 21.96 | 32,681 |
| Phoenix | 4.80 | 6.89 | 12.74 | 29,073 |

Table 11 Police Allocation and Crime Reduction: Philadelphia v.s. Phoenix

| | % Share of Extra Police for Philadelphia | % Crime Reduction |
|---|---|---|
| Complete Information | 91.0 | 15.97 |
| Based on $x$ | 98.0 | 10.12 |
| Based on Crime Rates | 47.2 | 5.60 |
| Based on Arrest Rates | 40.8 | 5.60 |
| Based on GDP | 51.4 | 7.05 |

The results are shown in Table 11 for each allocation rule. The first column shows the share of extra police allocated to Philadelphia and the second the associated percentage reduction in the overall crime rate. With complete information, 91 percent of the extra police is allocated to Philadelphia and the overall crime rate is reduced by 16 percent. Based on observable characteristics only, 98 percent of the extra police is allocated to Philadelphia, with a 10.1 percent reduction in crime. Based on the three separate outcome variables, the crime rate, the arrest rate and GDP, about 40 to 50 percent of police would be allocated to Philadelphia, with reductions in the overall crime rate one-half of what it would be if the allocation is based on observables alone and one-third if based both on observable and unobservable MSA characteristics. It appears that the effect on the overall crime rate does not change monotonically with the division of extra police. Moreover, as the number of extra police affects the set of market equilibria, a small change in the allocation ratio can lead to very different outcomes. For example, when the allocation of extra police goes from a split slightly favoring Phoenix (based on crime rates) to one slightly favoring Philadelphia (based on GDP), the reduction in crime increases from 5.6 to 7.1 percent.

It should be noted that although the allocation with complete information always dominates, the ranking of the effectiveness of the four sub-optimal allocation rules, i.e., Rules 2-4, depends on the pairs of MSA's under consideration. As shown in appendix Table A7, where we repeat the exercise for four other pairs of MSA's, we find that none of the four sub-optimal rules is always better than the other three. Thus, there is, no "golden rule" in terms of the allocation of police when information about unobservables is not available.

# 6 Conclusion

We have developed and estimated a model of crime in which the number of police, the crime rate, the arrest rate, the employment rate and the wage rate are joint outcomes of a subgame perfect Nash equilibrium. The local government chooses the size of its police force and citizens choose among work, home and crime. MSA's differ in the effectiveness of their police force, the productivity of their economies, their citizens' values of leisure and the marginal cost of police. The model is estimated using MSA-level data from the FBI Uniform Crime Reports and from the Current Population Survey for 2008. The model was shown to match the data well.

The estimates of the model were used to examine the effects on crime of monitored federal transfers to local governments to increase the size of the police force. We found that the effectiveness of such programs differs substantially across MSA's, and that knowledge about unobserved heterogeneity in MSA-specific characteristics is critical for the optimal allocation of police across MSA's. For example, in a pair-wise comparison of the Philadelphia and Phoenix MSA's, it was found that the optimal allocation of additional police between them would produce a reduction in the overall crime rate of 16 percent. On the other hand, an allocation of police based on their levels of crime or arrests would produce only a 5.6 percent reduction in crime. We did not pursue an analysis of a nationwide policy of allocating additional police among all MSA's jointly; although computationally burdensome, such an analysis would be feasible.

A major part of the empirical literature on crime has adopted an econometric strategy that is intended to approximate the solution of a behavioral optimizing model. This approach has been used to provide estimates of the responsiveness of crime rates to deterrence measures, such as apprehension rates, and to criminal justice resources, such as expenditures on policing. However, this approach is limited in terms of the kinds of policy analyses that can be performed. One example of this limitation is provided by the *ex ante* policy analyses that we have pursued in this paper, namely efficiently allocating federal grants to augmenting police forces across MSA's. There are, of course, limitations to our approach as well that would form an agenda for future research. Among the extensions would be to allow for the choice of punishment severity, to allow for dynamics in the choice of criminal behavior and to explicitly account for a government budget constraint.

# References


[1] Angrist, J and G. Imbens (1994), "Identification and Estimation of Local Average Treatment Effects," *Econometrica,* 62: 467-475.

[2] Ackerberg, D. and G. Gowrisankaran (2006): "Quantifying Equilibrium Network Externalities in the ACH Banking Industry," *The RAND Journal of Economics,* 27 (3): 738-761.

[3] Bajari, P., H. Hong and S. Ryan (2010): "Identification and Estimation of a Discrete Game of Complete Information," *Econometrica* 78: 1529-1568.

[4] Becker, G. (1968), "Crime and Punishment: An Economic Approach," *The Journal of Political Economy,* 76: 169–217.

[5] Bjorn, P. and Q. Vuong (1984): "Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with Application to Labor Force Participation," SSWP 537, California Institute of Technology.

[6] Burdett, K., R. Lagos and R. Wright (2003), "Crime, Inequality, and Unemployment," *American Economic Review*, 93: 1764–77.

[7] Burdett, K., R. Lagos and R. Wright (2004), "An On-the-Job Search Model of Unemployment, Inequality, and Crime," *International Economic Review,* 45: 681-706.

[8] Card, D. And L. Giuliano (forthcoming): "Peer Effects and Multiple Equilibria in the Risky Behavior of Friends," *Review of Economic and Statistics*, forthcoming.

[9] Conley, J. and P. Wang (2006), "Crime and Ethics," *Journal of Urban Economics,* 60: 107–123.

[10] Ehrlich, I. (1973), "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *The Journal of Political Economy,* 81: 521-565.

[11] Ehrlich, I. (2010), "The Market Model of Crime: A Short Review and New Directions," in *The Handbook on the Economics of Crime,* B.L. Benson and P.R. Zimmerman eds., Edward Elgar: 3-23.

[12] Engelhardt, B., G. Rocheteau and P. Rupert (2008), "Crime and the labor market: A search model with optimal contracts," *Journal of Public Economics,* 92: 1876–1891.

[13] Evans, W. and E. Owens (2007), "COPS and Crime," *Journal of Public Economics,* 91: 181-201.

[14] Flinn, C. (1986), "Dynamic Models of Criminal Careers." in *Criminal Careers and "Career Criminals",* A. Blumstein et. al eds., National Academy Press: 356-379.

[15] Freeman, S., J. Grogger and J. Sonstelie (1996), "The spatial concentration of crime," *Journal of Urban Economics,* 40: 216–231.

[16] Glaeser, E., B. Sacerdote and J. Scheinkman (1996) "Crime and social interactions," *Quarterly Journal of Economics,* 111: 507–548.

[17] Heckman, J. (1971), "Three essays on the supply of labor and the demand for goods," unpublished Ph.D. dissertation, Department of Economics, Princeton University, Princeton, N.J.

[18] Heckman, J. and E. Vytlacil (1998), "Instrumental Variables Methods for the Correlated Random Coefficient Model," *Journal of Human Resources,* 33: 974-1002.

[19] Huang, C., D. Liang and P. Wang (2004), "Crime and poverty: a search-theoretic approach," *International Economic Review* 45: 909–938.

[20] Jia, P. (2008): "What Happens When Wal-Mart Comes to Town; An Empirical Analysis of the Retail Discounting Industry," *Econometrica* 76: 1236-1316.

[21] Imai, S. and K. Krishna (2004), "Employment, Deterrence, and Crime In A Dynamic Model," *International Economic Review,* 45: 845–872.

[22] Imrohoroglu, A., A. Merlo, and P. Rupert (2000), "On the Political Economy of Income Redistribution and Crime," *International Economic Review,* 41: 1-25.

[23] Imrohoroglu, A., A. Merlo, and P. Rupert (2004), "What Accounts for the Decline in Crime," *International Economic Review,* 45: 707-729.

[24] Lochner, L. (2004), "Education, Work, and Crime: A Human Capital Approach," *International Economic Review,* 45: 811-843.

[25] Lochner, L. and E. Moretti (2004), "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *American Economic Review,* 94: 155-189.

[26] Sah, R. (1991), "Social osmosis and patterns of crime," *Journal of Political Economy,* 94: 1272–1295.

[27] Tauchen, H. (2010), "Estimating the Supply of Crime: Recent Advances," in *The Handbook on the Economics of Crime,* B.L. Benson and P.R. Zimmerman eds., Edward Elgar: 24-52.

[28] Verdier, T. and Y. Zenou (2004), "Racial beliefs, location, and the causes of crime," *International Economic Review*, 45: 731–760.

[29] Zhao, J., M.C. Scheider, and Q. Thurman (2002), "Funding community policing to reduce crime: have COPS grants made a difference?" *Criminology and Public Policy,* 2: 7–32.

## Appendix

### A1. Functional Forms

We categorize each of the observable characteristics. There are four age groups: 16-25, 26-40, 41-50, 51-64; four education groups: less than high school, high school grads, some college, and colleges grads or above; two race groups: black/Hispanics and others; and three number-of-young-kids groups: 0, 1, and 2 or above. The coefficient $\lambda$ for the default group is restricted to be zero. For human capital type:

$$x'\lambda^l = \lambda_0^l + \sum_{a=1}^{4} \lambda_{1a}^l I\left(age_i = a\right) + \sum_{s=1}^{4} \lambda_{2s}^l I\left(edu_i = s\right) + \lambda_3^l I\left(female\right) + \lambda_4^l I\left(black/hispanic\right).$$

For taste for staying at home:

$$\begin{aligned} x'\lambda^\eta &= \lambda_0^\eta + \sum_{a=1}^{4} \lambda_{1a}^\eta I\left(age_i = a\right) + \sum_{s=1}^{4} \lambda_{2s}^\eta I\left(edu_i = s\right) + \lambda_3^\eta I\left(female\right) + \lambda_4^\eta I\left(black/hispanic\right) \\ &\quad + \sum_{k=0}^{2} \lambda_{5k}^\eta I\left(kid_i = k\right) + \sum_{k=0}^{2} \lambda_{6k}^\eta I\left(kid_i = k, female\right) + \sum_{s=1}^{2} \lambda_{7s}^\eta I\left(edu_i = s, age_i = 1\right). \end{aligned}$$

For taste for crime:

$$\begin{aligned} x'\lambda^\kappa &= \lambda_0^\kappa + \sum_{a=1}^{4} \lambda_{1a}^\kappa I\left(age_i = a\right) + \sum_{s=1}^{4} \lambda_{2s}^\kappa I\left(edu_i = s\right) + \lambda_3^\kappa I\left(female\right) + \lambda_4^\kappa I\left(black/hispanic\right) \\ &\quad + \sum_{k=0}^{2} \lambda_{5k}^\kappa I\left(kid_i = k\right). \end{aligned}$$

### A2. Understanding the Dispersion of City Unobservables

Table A1 Response to 1 std dev. Improvement in Each MSA-Level Shock

| % | Productivity($\tau$) | | Arrest Efficiency($\epsilon$) | | Leisure Value($\overline{\kappa}$) | | Police Cost($\nu$) |
|---|---|---|---|---|---|---|---|
| | Gov. Move[a] | Police Fixed[b] | Gov.Move | Pol.Fixed | Gov.Move | Pol.Fixed | Gov. Move |
| Police | -0.43 | 0 | -13.96 | 0 | -0.41 | 0 | 14.83 |
| Crime | -2.81 | -6.30 | -11.91 | -17.73 | -3.61 | -8.27 | -7.83 |
| Arrest | 0.37 | 1.43 | 35.36 | 61.47 | -0.02 | 2.07 | 23.54 |

% Change in national outcomes, compared to the baseline model, if one type of MSA-level unobserved shocks is improved by 1 std dev, holding other shocks fixed.

[a]Local governments adjust police size. [b]Police size fixed at baseline.

## A3. Other Parameter Estimates

Table A2 Home Taste: $\lambda^{\kappa}$ and $\sigma_{\kappa}$

| | | |
|---|---|---|
| Less than HS | 0.04 | (0.02) |
| Some College | 0.10 | (0.03) |
| BA and Above | 0.38 | (0.02) |
| Age<26 | 0.14 | (0.03) |
| 41<Age<51 | 0.70 | (0.03) |
| Age>50 | 1.03 | (0.04) |
| Female | 0.20 | (0.03) |
| Black/Hispanic | -0.03 | (0.02) |
| 1 Kid | -0.98 | (0.08) |
| Multiple Kids | -1.14 | (0.06) |
| 1 Kid, Female | 1.73 | (0.05) |
| Multiple Kids, Female | 2.38 | (0.07) |
| Age<26, Less than HS | 0.95 | (0.04) |
| Age<26, HS | -3.06 | (0.26) |
| Constant | -1.74 | (0.01) |
| $\sigma_{\kappa}$ | 0.85 | (0.01) |

Default group: HS grads, Aged 26-40, male, non-black/Hispanic and no kid.

Table A3 Human Capital: $\lambda^l$ and $\sigma_l$

| | | |
|---|---|---|
| Less than HS | -0.56 | (0.02) |
| Some College | 0.19 | (0.01) |
| BA and Above | 0.84 | (0.02) |
| Age<26 | -0.72 | (0.02) |
| 41<Age<51 | 0.25 | (0.01) |
| Age>50 | 0.09 | (0.01) |
| Female | -0.53 | (0.02) |
| Black/Hispanic | -0.10 | (0.02) |
| Constant | 0.87 | (0.02) |
| $\sigma_l$ | 1.05 | (0.01) |

Default group: HS grads, Aged 26-40, male and non-black/Hispanic.

Table A4 Crime Taste: $\lambda^\eta$ and $\sigma_\eta$

| | | |
|---|---|---|
| Less than HS | 0.47 | (0.01) |
| HS Grads | 0.55 | (0.01) |
| Some college | -0.04 | (0.04) |
| Age<26 | 0.35 | (0.02) |
| 41<Age<51 | -1.19 | (0.10) |
| Age>50 | -0.36 | (0.06) |
| Female | -0.18 | (0.01) |
| Black/Hispanic | 0.16 | (0.02) |
| 1 Kid | -2.88 | (0.11) |
| Multiple Kids | -3.63 | (0.20) |
| Constant | -1.32 | (0.01) |
| $\sigma_\eta$ | 0.69 | (0.01) |

Default group: BA and above, Aged 26-40, male, non-black/Hispanic and no kid.

Table A5 Pr(School Enrollment|Age<26, Edu, Home)

| | | |
|---|---|---|
| Less than HS | 0.97 | (0.04) |
| HS | 0.92 | (0.05) |
| Some College | 0.96 | (0.15) |
| BA and above | 0.44 | (0.04) |

Prob of enrollment by current education for youth who choose the home option.

Table A6 Gov. Preference and Arrest

| | | |
|---|---|---|
| Gov. weight on Crime $\omega_1$ | 28.71 | (1.21) |
| Gov. weight on Arrest $\omega_2$ | 2.03 | (0.03) |
| Arrest Technology $b_0$ | 21.32 | (0.23) |
| Arrest Technology $b_1$ | 35.05 | (2.09) |
| Arrest Technology $\rho$ | 0.06 | (0.001) |

## A4. Counterfactual Experiment

Table A7 Police Allocation Rules and Crime Reduction

| % | Pair 1 | Pair 2 | Pair 3 | Pair 4 |
|---|---|---|---|---|
| Complete Info | 15.20 | 8.42 | 14.70 | 16.01 |
| Based on $x$ | 5.51 | 2.76 | 13.51 | 14.62 |
| Based on Crime Rates | 5.51 | 6.32 | 5.58 | 11.90 |
| Based on Arrest Rates | 5.51 | 6.32 | 5.58 | 5.37 |
| Based on GDP | 11.46 | 6.32 | 11.18 | 7.86 |

% reduction in the overall crime rate of each pair of MSA's in response to a 20% increase in their total police, allocated between the pair according to various rules.

Pair 1: Atlanta-Sandy Springs-Marietta v.s. Philadelphia

Pair 2: Dallas-Plano-Irving v.s. Phoenix-Mesa-Scottsdale

Pair 3: Fort Lauderdale-Pompano Beach-Deerfield Beach v.s. Philadelphia

Pair 4: Philadelphia v.s. Washington-Arlington-Alexandria

**For Online Publication**

## B1 Individual Types

### B1.1 Lower Bound of the Taste for Crime $\eta^*$

$$\begin{aligned}
\max\{V^1|\eta\} &= \max\{(1-\mu)[(1-\pi)E(\ln(\alpha y+b))+\pi\ln(\underline{c})]+\mu\ln(b)+\eta\} \\
&= \ln[\alpha l_{N_l} r + r l_1]+\eta \\
&= \ln(\alpha l_{N_l}+l_1)+\ln(r)+\eta \\
\min\{V^0|\eta\} &= \min\{\max\{\ln(y)+\mu\ln(1-\alpha),\ln(b)+\mu\ln(1-\alpha)+\kappa\}\} \\
&= \min\{\max\{\ln(y),\ln(b)+\kappa\}+\mu\ln(1-\alpha)\} \\
&= \min\{\max\{\ln(y),\ln(b)+\kappa\}\}+\ln(1-\alpha) \\
&\geq \ln(b)+\kappa_1+\ln(1-\alpha) \\
&= \ln(l_1)+\ln(r)+\kappa_1+\ln(1-\alpha)
\end{aligned}$$

$$\eta^* \geq -\ln(\alpha l_{N_l}+l_1)+\ln(l_1)+\kappa_1+\ln(1-\alpha)$$

One will never be a criminal for any $\eta \leq \eta^*$.

### B1.2 Number of Criminal Propensity Types

$$\begin{aligned}
N^* &= \begin{array}{c} N_\eta\left(\sum_{j=1}^{N_\kappa} I\left(\ln(l_1)+\kappa_j > \min\{\ln(l)\}\right)\right) \\ +N_\eta\left(\sum_{i=1}^{N_l} I\left(\ln(l_i) \geq \ln(l_1)+\min\{\kappa\}\right)\right)\end{array} \qquad (16) \\
&= N_\eta \times N_\kappa + N_\eta\left(\sum_{i=1}^{N_l} I\left(\ln(l_i) \geq \ln(l_1)+\kappa_1\right)\right) \\
&\leq N_\eta \times (N_\kappa + N_l - 1).
\end{aligned}$$

The first term in (16) counts the total number of types who choose not to be employed. Conditional on declining a job offer, these individuals' decisions will not vary with $l$. Any $(\eta_n,\kappa_n)$ may fall into this group as long as the first indicator function holds, which is always true given that $\kappa$ is always positive, hence $N_\eta \times N_\kappa$. The second term counts the number of types who will choose between work and crime and hence whose decisions will not vary with $\kappa$. The condition in the indicator function excludes types who will choose home over work and hence are already included in the first term.

### B1.3 Measure of Criminal Propensity Types

$$h_{jn} = \begin{cases} p_{\kappa_n} p_{\eta_n} \sum_{i=1}^{N_l} p_{l_i} I\left(\ln(l_i) < \ln(l_1) + \kappa_n\right) & \text{if } T_n = \ln(l_1) + \kappa_n - \eta_n, \\ p_{\eta_n} p_{l_n} \sum_{i=1}^{N_\kappa} p_{\kappa_i} I\left(\ln(l_n) \geq \ln(l_1) + \kappa_i\right) & \text{if } T_n = \ln(l_n) - \eta_n. \end{cases}$$

The first line is the measure of individuals with $(\eta_n, \kappa_n)$ who declined a job offer. The second line is the measure of those with $(\eta_n, l_n)$ who prefer employment over non-employment.

## B2. Model Fit

### B2.1 Outcomes By Individual Characteristics

Table B1 Outcomes By Individual Characteristics

| | Employment Rate (%) | | Earnings\|Working ($) | |
|---|---|---|---|---|
| | Data | Model | Data | Model |
| Less than HS | 47.0 | 47.4 | 19,199 | 18,626 |
| HS | 73.8 | 74.3 | 31,128 | 30,110 |
| Some College | 77.0 | 77.7 | 35,398 | 33,914 |
| BA and Above | 85.2 | 85.8 | 65,245 | 62,635 |
| Age<26 | 54.1 | 53.5 | 18,502 | 18,574 |
| 41<Age<51 | 82.3 | 82.8 | 49,908 | 48,789 |
| Age>50 | 71.4 | 72.4 | 49,951 | 47,848 |
| Female | 67.6 | 67.1 | 34,203 | 34,128 |
| Black/Hispanic | 68.9 | 68.6 | 31,454 | 31,969 |
| 1 Kid | 75.1 | 76.5 | 44,936 | 41,552 |
| Multiple Kids | 71.1 | 72.0 | 48,430 | 44,259 |
| 1 Kid, Female | 62.8 | 63.0 | 32,999 | 33,522 |
| Multiple Kids, Female | 53.7 | 52.9 | 31,064 | 36,922 |

**B2.2 Correlation between Crime Rates and Distribution of Education Levels**

Figure 1 shows the fit of crime rates conditional on the distribution of educational attainment within an MSA. For the first graph, MSA's are ranked by the fraction of people without a high school degree, and included cumulatively by that rank. The average crime rate among the included MSA's is plotted in the first graph. The model predictions are all within the 95% confidence interval of the data, but the fit is worse at the lower percentiles, i.e., when fewer MSA's are involved in the calculation. The next 3 graphs use similar methods to show the fit for the relationship between crime rates and the fraction of each of the other three education levels. The last figure shows the fit of crime rates among MSA's cumulatively included by their ranks in the mean education among their citizens.[46]

[46] We assign 10, 12, 14 and 17 years of schooling to the four education levels respectively, and calculate the mean education of an MSA based on the discrete distribution of education levels.

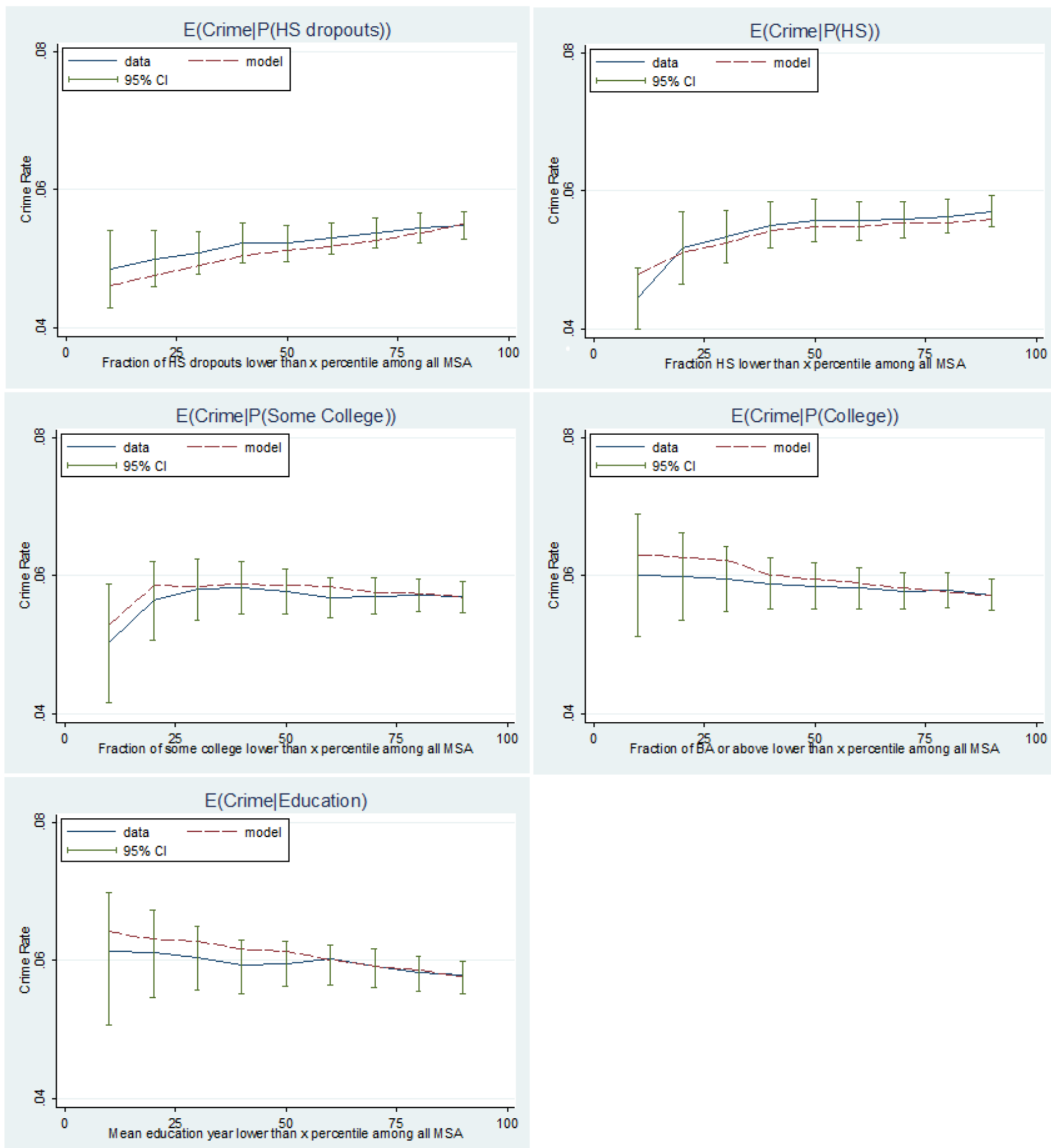


Figure 1: Model Fit: Crime and Educational Distribution