The Effect of Employment Frictions on Crime: Theory and Estimation

By

Bryan Engelhardt

April 2008

COLLEGE OF THE HOLY CROSS, DEPARTMENT OF ECONOMICS
FACULTY RESEARCH SERIES, PAPER NO. 08-05*

Department of Economics
College of the Holy Cross
Box 45A
Worcester, Massachusetts 01610
(508) 793-3362 (phone)
(508) 793-3710 (fax)

http://www.holycross.edu/departments/economics/website

*All papers in the Holy Cross Working Paper Series should be considered draft versions subject to future revision. Comments and suggestions are welcome.

The Effect of Employment Frictions on Crime: Theory and Estimation

By
Bryan Engelhardt[†]
College of the Holy Cross

April 2008

Abstract

I investigate how long it takes for released inmates to find a job, and when they find a job, how their incarceration rate changes. An on-the-job search model with crime is used to model criminal behavior, derive the estimation method and analyze several policies including a job placement program. The results show the unemployed are incarcerated twice as fast as the employed and take on average four months to find a job. Combining these results, it is demonstrated that reducing the average unemployment spell of criminals by two months reduces crime and recidivism by more than five percent.

JEL Classification Codes: C41, E24, J0, J64

Keywords: crime, search, unemployment, wage dispersion

I thank my advisors Beth Ingram, George Neumann, Peter Rupert and Gabriele Camera as well as Guillaume Rocheteau and Gerard van den Berg for their comments on earlier versions of the paper. I also thank the seminar participants at the 2007 Midwest Macroeconomics Meetings and Midwest Economics Association Conference.

[†]Department of Economics, College of the Holy Cross, Worcester, MA 01610-2395, 508-793-3362 (phone), 508-793-3710 (fax), bengelha@holycross.edu, Bryan-Engelhardt@uiowa.edu

1 Introduction

Empirical research has documented the correlation between crime, inequality and unemployment. Theory claims the unemployed and employed earning a low wage face lower costs of committing crime and in turn perpetrate more of it. Therefore, if employment frictions contribute to unemployment and inequality, then how would a policy aimed at reducing these frictions affect unemployment, inequality and crime?

Others have asked similar questions. Imrohoroğlu, Merlo & Rupert (2000) develop a dynamic general equilibrium model with crime in order to investigate how income redistribution policies influence crime and inequality. Engelhardt, Rocheteau & Rupert (2007) use a search and matching model to see how labor market policies affect crime and unemployment.

Here, I take a model that simultaneously captures crime, inequality and unemployment and integrate heterogeneous crime opportunities, workers and firms. After constructing a model from the related literature, I develop a procedure that estimates the model's parameters. Constructing and estimating the model serves several purposes. First, it confirms the empirical link between crime and unemployment demonstrated by Gould, Weinberg & Mustard (2002) and others. Estimation also highlights how heterogeneity is essential in capturing crime, inequality and unemployment simultaneously. Finally, the structural model with the estimated parameters is used to demonstrate how policies aimed at reducing employment frictions, in particular a successful job placement program, would reduce crime, unemployment and inequality.[2]

In constructing an empirically relevant model, I merge several ideas found in the related literature. In particular, the model I develop builds upon the on-the-job search model of crime proposed by Burdett, Lagos & Wright (2004) by integrating heterogeneous types of workers and firms as found in Burdett & Mortensen (1998). I add heterogeneous crime opportunities following Engelhardt et al. (2007). The resulting characteristics of the hybrid model are as follows. First, it takes time for workers to find a job due to labor market search frictions and from these frictions unemployment and wage inequality occur. Adding heterogeneous firms enables the model to accurately characterize the observed wage distribution. Adding heterogeneous workers allows the model to explain why some individuals do not engage in crime, a result dependent on how much individuals

[2] I do not explicitly explain how employment frictions can be reduced except for a small example in Section 4. However, several studies have analyzed job placement programs for former inmates including Chung, Schmidt & Witte (1991) and Visher, Winterfield & Coggeshall (2005).

values their leisure. Finally, incorporating a distribution of crime opportunities accounts for the fact that individuals accept opportunities to commit crime at different rates depending upon the costs, i.e. whether they would lose a good paying job.

After constructing a model, I implement an estimation procedure to test whether crime decisions are influenced by employment frictions. The likelihood function is derived from the model and is able to identify employment frictions and incarceration rates using data taken from the National Longitudinal Survey of Youth (NLSY). The most relevant result is that individuals who are unemployed are caught committing crime and imprisoned two times faster than low wage workers and four times faster than high wage workers. Moreover, individuals released from jail take an average of four and a half months to find a job.

Next, I turn to policy analysis. Consider, for example, a program capable of cutting the average time it takes for criminals to find a job from four months to two months. What I find is that such a program could reduce the equilibrium crime rate by more than five percent. Also, the same policy can reduce the recidivism rate by roughly the same amount. As an alternative tool for fighting crime, I find the elasticity of crime with respect to the duration of incarceration to be -0.18. This is consistent with Levitt (2004) who finds the elasticity to be between -0.1 and -0.4 depending upon the type of crime.

In evaluating how much employment frictions affect crime, I discover several other interesting results. Specifically, I find further support that the duration of unemployment for those previously incarcerated is roughly equal to the average, a finding in line with Grogger (1995). However, the average wage offer for a criminal is 35% less than those never observed to be convicted and incarcerated for a crime. In addition, I analyze how demographics are associated with criminal participation, such as age, race, education and location of residence (urban/rural). In general, I find those who face longer unemployment spells, such as blacks, commit more crime. Finally, the estimation I propose provides a new approach in testing whether a relationship exists between an individual's criminal participation, employment status and wage. For instance, I find those paid the minimum wage are incarcerated 25% less than those who are unemployed, while those paid in the upper decile of the wage distribution are imprisoned 75% less.

Section 2 introduces the model's environment and characterizes the equilibrium. Section 3 develops the estimation procedure, discusses the estimated parameters including demographic effects and analyzes the accuracy of the model. Finally, Section 4 discusses the model's policy implications.

2 Model

In this section, I present the model's environment and outline the equilibrium.

2.1 Environment

The hybrid model is composed of assumptions taken from Burdett & Mortensen (1998), Burdett et al. (2004) and Engelhardt et al. (2007). To begin, there exists a continuum of risk neutral heterogeneous agents and firms who discount the future at rate r. There are two types of firms and two types of agents.

Agents differ by their utility flow when unemployed, which is b_k where $k \in \{c, nc\}$, $b_c < b_{nc}$ and ϕ is the proportion of type c agents. Unemployed agents receive job offers at rate λ_0, observe a wage offer drawn randomly from a wage offer distribution $F(w)$, and if accepted, become employed instantaneously and are paid the wage w over the tenure of the job.

Agent's utility flow when employed is equal to their wage. They lose their jobs at rate δ and receive new job offers at rate λ_1 with a wage drawn randomly from $F(w)$. Given acceptance of a new wage offer, the agent changes jobs instantaneously.

Employed and unemployed agents receive crime opportunities at rate μ.[3] The value of a crime opportunity is drawn from a discrete distribution $\Gamma(g)$ with the finite support $\mathscr{G}$. The timing of the crime opportunities is instantaneous where agents receive a crime opportunity, realize its payoff, and decide whether to take the opportunity. If taken, the utility flow of a crime is instantaneous and equal to its value.

Agents committing crime are instantaneously caught with probability π, consume z while in jail, and are released at rate ρ.

Firms have a linear production function and differ by their marginal (= average) revenue product p_i, where $i \in \{L, H\}$, $p_L < p_H$, and φ is the proportion of low productivity firms. Firms post and commit to pay two types of wages, $\{w_c, w_{nc}\}$, depending upon an agent's criminal history.[4]

[3]Even though the arrival rate of crime is independent of an agents labor force status, the employed individuals could commit less crime due to the fact unemployed agents accept crime opportunities that the employed reject. The potential for such a decision comes from adding heterogeneous crime opportunities. Also, allowing for a state dependent μ will not change the estimated results below.

[4]Assuming firms pay their workers conditional on their criminal history will turn out to be identical to assuming

There are several reasons for incorporating heterogeneous types of agents, crime opportunities and firms into the environment. First, two types of agents allows for the possibility that some individuals may never commit crime and risk imprisonment because of their high value of leisure (or freedom). A distribution of crime opportunities captures the fact individuals commit crime at different rates depending upon the costs. In other words, agents commit crime at different rates because some accept lower value opportunities as their costs of being caught are lower. Finally, a distribution of firm productivities enables the model to fit the wage distribution as demonstrated in Section 3.[5]

2.2 Equilibrium

I will characterize the model's equilibrium in the steady state. To begin, an equilibrium contains a distribution of wage offers, $F(w)$. On the supply side, agents maximize future expected utility by following a set of reservation rules. In particular, they follow a reservation wage strategy for taking a job. In other words, an agent of type $k \in \{c, nc\}$ accepts any wage at or above R_k where the reservation wage is determined at the point where agents are indifferent between unemployment and being employed with a wage R_k. The other type of reservation strategy is the reservation crime value. Following the same logic, the reservation crime value is identified at the point where an agent is indifferent between accepting or declining a crime opportunity. For example, the unemployed accept opportunities $g > g_u$ where g_u is the crime reservation value of the unemployed and g is located on the support $\mathscr{G}$. Finally, an equilibrium contains a mass of individuals incarcerated, unemployed and employed.

Before defining the equilibrium, I provide and discuss several important features of the environment.

To begin, agents with a criminal history can be thought of as operating in an independent

there exists a criminal and non-criminal labor market. Also, the Work Opportunity Tax Credit and other federal government programs offer employers the chance to reduce their tax liabilities conditional on if they hire workers with a criminal history. Therefore, I assume criminals differ in their productivity to the firm, their duration of employment, and in turn, the offered wage.

[5] I include only two types of firms because of the limited number of observations. The alternative specification given the limited number of observations would be to assume a parametric form for firm productivity such as in Bontemps, Robin & van den Berg (1999). I take the non-parametric approach because it is simple, easily interpreted and I argue sufficient within the context of discussing crime. Bowlus, Kiefer & Neumann (1995) take a similar approach and estimate the optimal number of firm types to be five. Given the model I propose and estimate, I demonstrate two is adequate. The sample size imposes a constraint because of the link between firm productivity and crime opportunities, which I explain in detail following Proposition 2.

labor market. The idea is straight forward. Firms cannot decrease profits by using the information about an agent's criminal history. Therefore, due to the additively separable property of the profit function, the wage offer distribution the criminal agents face, $F_c(w)$, can be considered and solved independently of the distribution the non-criminal agents face, $F_{nc}(w)$. The fact implicitly implies that firms observe an agent's criminal history, assume the agent will commit crime again, and therefore pay the agent accordingly.

Proposition 1 describes why agents' valuations of leisure/freedom are important in determining their decisions to commit crime.

Proposition 1 *If an agent's value of leisure is greater than a threshold value $\overline{b}$ then the agent never engages in crime.*

Intuitively, agents who place a high value on their leisure find the cost of imprisonment to be too high and do not commit crime. Analytically, it is straightforward to show given a fixed value of z, π and ρ.

Proposition 1 is important for several reasons. First, it simplifies the model's solution. Second, the assumption that agents differ according to their value of leisure, b, has been used in the literature such as Eckstein & Wolpin (1995) to enable the model to fit the observed wage distribution and duration of unemployment simultaneously. Third, it can explain why some individuals might never engage in crime. Fourth, it enables the model to accurately capture the observed recidivism and crime rates simultaneously.[6] Finally, the result provides an intuitive explanation for why some individuals may never engage in crime: they value their freedom.

Taking from Proposition 1 and the reasons explained above, I assume a fraction of agents never

[6] With homogeneous b, the estimated model either overestimates the crime rate or underestimates the recidivism rate. For example, take the best case scenario by assuming the lower bound on the amount of crime committed. If the model's crime rate is positive and it is at the lower bound then the unemployed must be the only ones committing crime. Now, observe a few features of the data. First, the aggregate number of unemployed individuals is approximately 5%. Second, the crime opportunities an unemployed individual takes is approximately 1 per month. Piehl & DiIulio (1995) contain a detailed discussion of multiple microeconomic data sources for the number of opportunities, all of them asserting roughly 12 per year. Alternatively, the number of crimes a criminal commits per month can be backed out from the recidivism rate given the probability of being caught. In either case, the minimum amount of crime in the model is roughly $(1)(.05) = 5\%$ per month. However, the monthly crime rate given by the FBI is roughly .03% per month. Hence, the model with homogeneous agents produces more than fifteen times too much crime. From Proposition 1, it is possible for only a fraction of the population to commit crime given heterogeneous values of leisure. Thus, the model can produce an appropriate amount of crime given the mass of "criminal" agents is relatively small. I estimate similar values for the number of crimes committed by an individual as found in Piehl & DiIulio (1995) given π is roughly 2.5%.

commit crime or $b_c < \overline{b} < b_{nc}$.[7]

From the assumption $b_c < \overline{b} < b_{nc}$, the wage offer distributions $F_c(w)$ and $F_{nc}(w)$ can be solved using only one type of agent. Although characterizing the "criminal labor market" with homogeneous agents is simplistic, it is appealing because it allows for a more transparent discussion about the effects of employment frictions on crime because the results are not conditional on an agent's type.

The next simplification to the model's equilibrium, and addition I make to the theory, discusses the efficiency of the model's equilibrium.

Proposition 2 *In equilibrium, if firms of the same type are offering wages above and below the threshold to deter workers from a crime, then the equilibrium is not Pareto efficient.*

To outline the idea, assume one agent is working and accepts any crime opportunity worth g or more. Also, another agent is making a higher wage that deters him from accepting an opportunity worth g. Assuming both agents work for the same type of firm, the firm with the worker who takes the g opportunity is indifferent in switching to a wage that deters the worker from taking it (otherwise the other firm would deviate since both firms are identical). Therefore, firms are indifferent from deterring a crime opportunity worth g, while the criminal agents and their victims are strictly better off. Hence, it is not Pareto efficient if firms of the same type are offering wages above and below the threshold to deter agents from a crime opportunity.

Given Proposition 2, I simplify the equilibrium by assuming that identical type firms deter their workers from the same type of crimes. In other words, firms of the same type lose workers at the same rate even though they potentially pay different wages. Individuals paid different wages can commit the same amounts of crime because $\Gamma(g)$ is discrete. Although the assumption might seem restrictive, it can be empirically tested. In addition, it can be relaxed by adding additional firms. From a game theoretic perspective, the assumption can be viewed as allowing only symmetric equilibria.

This leads me to the final simplifying result. Agents do not flow from high productivity firms to low productivity firms.

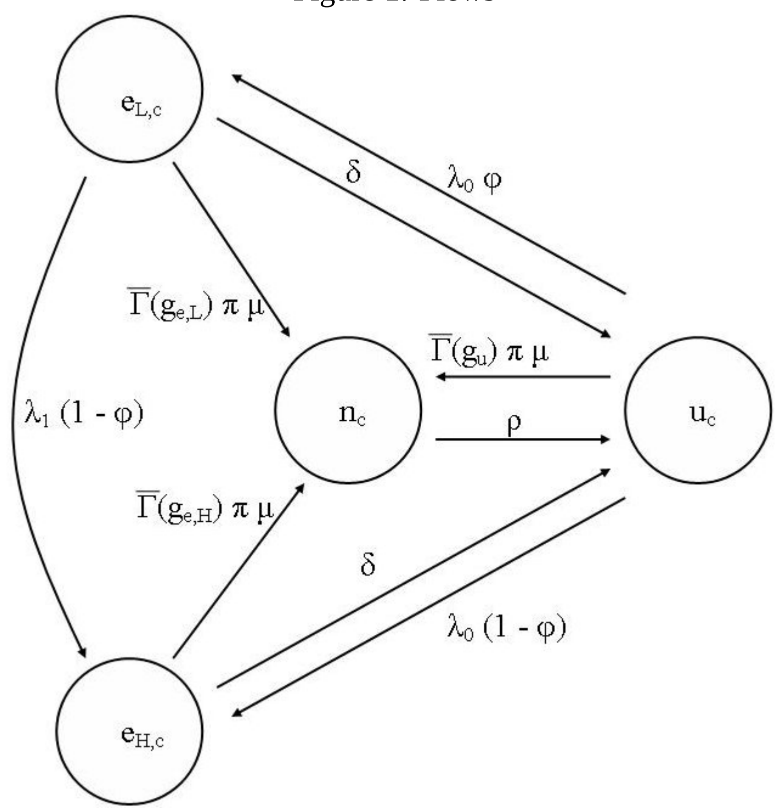
[7]The implications of Proposition 1 can be derived from an alternative set of assumptions. For instance, incorporating human capital would imply an opportunity cost of imprisonment. Therefore, any assumption where agents have heterogeneous values of human capital (and z) implies the same result as heterogeneous values of b since there exists a threshold value, $\bar{z}$, where agents do not commit crime. In Section 3, I look at how education affects crime, unemployment and inequality.

Proposition 3 *High productivity firms pay higher wages than lower productivity firms.*

The proof is given in the Appendix.

Figure 1 demonstrates the flows for an agent choosing to commit crime. At any point in time there is a fraction of agents, $e_{i,k}$, of type k who are employed at type i firms. In addition, there is a fraction of unemployed agents u_k of type k, and a fraction of agents in jail, n_k. To reiterate, agents do not flow from $e_{H,k}$ to $e_{L,k}$ due to Proposition 3. Also, as described in Proposition 2, individuals employed by the same type of firm are deterred from the same type of crimes. Therefore, employed individuals commit crime at a rate $\mu\overline{\Gamma}(g_{e,i})$, where $\overline{\Gamma}(g) = 1 - \Gamma(g)$ and $g_{e,i}$ is the highest crime value an agent will choose not to take given he is employed by firm i. Finally, the unemployed criminal types commit crime at the same rate, $\mu\overline{\Gamma}(g_u)$, where g_u is the highest crime value the unemployed agent will decline.

Figure 1: Flows



At this point, it is important to highlight the reason I incorporate a distribution of crime opportunities into the model. The reason is to allow for different criminal participation rates for those unemployed and employed, or for those paid at the lower and higher end of the wage distribution.

Analytically this means, $\overline{\Gamma}(g_{e,H}) \leq \overline{\Gamma}(g_{e,L}) \leq \overline{\Gamma}(g_u)$, because higher productivity firms pay higher wages and the cost of being caught is lower when unemployed than when employed.

At this stage, the equilibrium can be defined.

Definition 1 *A steady-state equilibrium is defined by*

(i) *a set of reservation wages, R_k, for $k \in \{c, nc\}$ that maximizes agents' expected utility*

(ii) *crime reservation values conditional on unemployment, g_u, employment at low productivity firms, $g_{e,L}$, and employment at high productivity firms, $g_{e,H}$, which maximize agents' expected utility*

(iii) *a fraction of agents unemployed, u_k, employed, $e_{i,k}$, and incarcerated, n_k for $k \in \{c, nc\}$ and $i \in \{L, H\}$ that equate the flows in and out of each state*

(iv) *a crime rate*

$$\mu\overline{\Gamma}(g_u)u_c + \mu\overline{\Gamma}(g_{e,L})e_{L,c} + \mu\overline{\Gamma}(g_{e,H})e_{H,c}, \text{ and}$$

(v) *a wage offer distribution, $F(w)$, that is based on firms maximizing steady-state profits.*

The derivation of the wage distribution, incarceration rate, employment rate and unemployment rate is completed in the Appendix.

Given the above definition and assumptions, the model has the potential for multiple equilibria as first shown by Burdett, Lagos & Wright (2003). The resulting equilibria can be summarized by the way employment and/or higher wages deter agents from committing crime. The number of potential equilibria could be large depending upon the support of $\mathscr{G}$, but I summarize them as

Characterization of Equilibria[8]

1. $\overline{\Gamma}(g_{e,H}) = \overline{\Gamma}(g_{e,L}) < \overline{\Gamma}(g_u)$,
2. $\overline{\Gamma}(g_{e,H}) < \overline{\Gamma}(g_{e,L}) \leq \overline{\Gamma}(g_u)$, and
3. $\overline{\Gamma}(g_{e,H}) = \overline{\Gamma}(g_{e,L}) = \overline{\Gamma}(g_u)$.

The interpretation of the equilibria is critical in understanding the quantitative results. Equilibrium 1 is where employment deters agents from committing crime. In other words, all firms pay a

[8] As noted above, if an agent commits crime at all, then they commit crime when unemployed. It is seen by realizing the costs of crime when an agent is unemployed are no more than when employed.

high enough wage that an individual with a job commits less crime. Equilibrium 2 is where high productivity firms, by paying higher wages, deter their workers from committing as much crime as the unemployed and employed at the lower paying, lower productivity firms. Finally, Equilibrium 3 is where all agents commit the same amount of crime independent of their wage or labor force status. It is critical to realize that if I am unable to reject Equilibrium 3 then I am unable to claim employment or higher wages deter crime and therefore the justification of a job placement program is lost.

3 Estimation

Maximum Likelihood Estimation (MLE) is used to estimate the parameters. Given the estimates, I assess the effect employment frictions have on crime, unemployment and inequality.

3.1 Data

For the estimation I use data taken from the NLSY, a panel data set initiated within the U.S. in 1979. Starting in 1989 and ending in 1993, the NLSY contains weekly data on whether or not an individual is incarcerated, while it contains data on an individual's labor force status for the entire panel. The time individuals spend in each state, whether employed, unemployed or incarcerated, along with their wages when employed, are sufficient to identify the relevant parameters of the model. Therefore, the model is estimated using data starting in 1989 and ending in 1993.

From the assumption $b_c < \overline{b} < b_{nc}$, I break the sample into two subgroups, criminals and non-criminals, or $\{b_c, b_{nc}\}$. I identify an individual as a criminal if I observe him to be incarcerated during the sample period. If I never observe an individual to be incarcerated, I partition him into the non-criminal sub-sample.[9] In assuming observations are missing at random, I exclude them. In addition, anyone exiting the labor market for reasons other than incarceration are excluded as it is likely that the behavior of such individuals, at least in a certain period, deviates substantially from the behavior as described in the model.[10]

[9]This type of partitioning creates latent variable bias when estimating the parameters of the non-criminal group due to the fact criminal types could be in the non-criminal sample. However, I argue the bias is small due to previous evidence that these groups have similar employment opportunities.

[10]The reader can refer to Engelhardt et al. (2007) for a discussion regarding the criminal behavior of agents outside the labor force.

For the non-criminal subgroup, the data begins with a duration of unemployment, t_1. Unemployment is interrupted by an agent becoming employed, $d_{1,e} = 1$. Once employed the wage is observed, $\tilde{w}$. The length of time an agent is employed at his first job is recorded as t_2. At the end of the employment period, the labor force status that the agent transitions to is recorded. They either transition to unemployment, $d_{2,u} = 1$, or become employed with a higher wage, $d_{2,e} = 1$.

For the criminal subgroup, the data also begins with a period of unemployment, t_1. However, prior to this period an individual could have been incarcerated and the time he spent incarcerated is t_0, which is potentially left censored. For the criminal sub-group, unemployment is interrupted by either an agent going to jail, $d_{1,n} = 1$, or becoming employed, $d_{1,e} = 1$. If employed, the wage is observed, $\tilde{w}$. If the agent is incarcerated then the construction of the individual's panel is complete. However, if the agent becomes employed then the duration of employment is recorded as t_2. At the end of the employment period, the reason for exiting employment is recorded. The reasons include incarceration, $d_{2,n} = 1$, unemployment, $d_{2,u} = 1$, or employment, $d_{2,e} = 1$.

The descriptive statistics are found in Table 1. In interpreting the work histories, we see the

Table 1: Work History Sample Means

Variable	Criminals	Non-Criminals	Full Sample
$d_{1,n}$	0.1		0.02
$d_{1,e}$	0.88	0.99	0.97
$d_{2,e}$	0.11	0.19	0.17
$d_{2,u}$	0.4	0.49	0.47
$d_{2,n}$	0.19		0.03
t_0	13.54		2.48
	(10.21)		(6.82)
t_1	3.9	4.64	4.51
	(6.7)	(6.21)	(6.31)
t_2	20.53	26.1	25.08
	(26.18)	(19.44)	(20.81)
$\tilde{w}$	1034.48	1111.45	1097.34
	(560.66)	(705.8)	(684.19)
N	209	931	1140

Note: Standard deviations are in parentheses. Duration and wage statistics are monthly. Unemployment is defined by individuals that claim at least some time is spent searching for work. Employment is defined by those working at least 30 hours per week. The transition probabilities do not sum to one at each stage because the data is right censored.

average wage of the criminal type is smaller than the non-criminal. Also, we find their average duration of unemployment is smaller. However, a smaller duration does not imply the criminal types find jobs faster but rather they find a job *or* are caught committing crime and imprisoned faster than non-criminals find a job. The same interpretation is true for the duration of employment.

It is important to note an attempt to estimate the model with an increased number of firm types is unreasonable as I observe less than fifty individuals incarcerated while employed. Specifically, a limited number of firms are used because each firm requires the estimation of an additional parameter, $\overline{\Gamma}(g_{e,i})$. Therefore, I limit the number of firm types, or parameters $\overline{\Gamma}(g_{e,i})$, to two.

3.2 Likelihood Function

I will break down how the parameters are estimated using MLE. I build the likelihood on the assumption that the model is correct and data on previous convictions is uninformative. In addition, I assume in the model and build the likelihood function under the assumption that an individual does not transition from a criminal to a non-criminal state or vice-versa. The likelihood of the sample is obtained by multiplication of each individual's contribution. To simplify the composition, I outline the criminal likelihood function as the non-criminal likelihood can be deduced by constraining $\mu\pi\overline{\Gamma}(g_u) = \mu\pi\overline{\Gamma}(g_{e,L}) = \mu\pi\overline{\Gamma}(g_{e,H}) = 0$.

In the data I observe criminals exiting jail and entering unemployment, and from the model, the arrival rate of exiting is a Poisson process. As a result, the duration of incarceration is exponential. Hence ρ is estimated by

$$P(t_0) = \rho e^{-\rho t_0}.$$

The time t_0 is the only period that is potentially left censored due to the choice of using flow sampling at the point where individuals enter unemployment. Therefore, the likelihood contribution of an individual panel which is left censored is $P(t_0) = e^{-\rho t_0}$ instead of $P(t_0) = \rho e^{-\rho t_0}$. Going forward, right censoring of the data occurs and is accounted for in the estimation but is suppressed in the composition until the complete likelihood is composed in Equation 1.

At the beginning of the sampling period, individuals are in the unemployed state and they transition out of unemployment according to a Poisson process. Therefore, the duration of unemployment, $P(t_1)$, is exponential. The likelihood function is

$$P(t_1) = (\mu\pi\overline{\Gamma}(g_u) + \lambda_0)e^{-(\mu\pi\overline{\Gamma}(g_u)+\lambda_0)t_1}.$$

The likelihood of transitioning to jail or employment after being unemployed, $P(n_{t_1})$ and $P(e_{t_1})$ respectively, is

$$
\begin{aligned}
P(n_{t_1}) &= \frac{\mu\pi\overline{\Gamma}(g_u)}{\mu\pi\overline{\Gamma}(g_u)+\lambda_0}, \textit{ and} \\
P(e_{t_1}) &= \frac{\lambda_0}{\mu\pi\overline{\Gamma}(g_u)+\lambda_0}.
\end{aligned}
$$

Obviously, for the non-criminal sub-group the probability of transitioning to jail is zero. If a criminal type transitions to jail, then the individual's contribution to the likelihood function ends. Otherwise, an individual transitions to employment and receives a wage offer from the distribution $f_c(w)$, which is the density function of $F_c(w)$ derived in Equation 7 in the Appendix.

Depending upon the equilibrium, agents transition out of employment at different rates. The differences in transition rates arise because employed agents might be employed at low or high productivity firms. On the other hand, Proposition 2 argues that those employed by the same type of firm take the same crime opportunities. In other words, agents employed by low productivity firms transition out of employment at the same rate, and those employed at a high productivity firms transition out of employment at the same rate.

From the model, agents employed by type i firms find their average time of employment to be distributed exponentially. Hence, the duration of employment is

$$
P(t_2|\tilde{w}) = (\mu\pi\overline{\Gamma}(g_{e,i})+\lambda_1(1-F(\tilde{w}))+\delta)e^{-(\mu\pi\overline{\Gamma}(g_{e,i})+\lambda_1(1-F(\tilde{w}))+\delta)t_2},
$$

for $i \in \{L,H\}$ depending on if $\tilde{w} \in [\underline{w}_{L,c}, \overline{w}_{L,c}]$ or $\tilde{w} \in [\underline{w}_{H,c}, \overline{w}_{H,c}]$.[11]

The likelihood of transitioning to a job, jail, or unemployment after being employed, $P(e_{t_2})$, $P(n_{t_2})$, and $P(u_{t_2})$, respectively, is

$$
\begin{aligned}
P(e_{t_2}|\tilde{w}) &= \frac{\lambda_1(1-F(\tilde{w}))}{\mu\pi\overline{\Gamma}(g_{e,i})+\lambda_1(1-F(\tilde{w}))+\delta}, \\
P(n_{t_2}|\tilde{w}) &= \frac{\mu\pi\overline{\Gamma}(g_{e,i})}{\mu\pi\overline{\Gamma}(g_{e,i})+\lambda_1(1-F(\tilde{w}))+\delta}, \textit{ and} \\
P(u_{t_2}|\tilde{w}) &= \frac{\delta}{\mu\pi\overline{\Gamma}(g_{e,i})+\lambda_1(1-F(\tilde{w}))+\delta}.
\end{aligned}
$$

In summary, the likelihood function is constructed from

[11] As described in the Appendix, $\underline{w}_{i,c}$ and $\overline{w}_{i,c}$ is the lowest and highest wage paid by firms of type $i \in \{L,H\}$.

- Wage offers:

$$\tilde{w} \quad = \quad \text{wage received} \quad \sim f(w)$$[12]

- Duration times:

$$\begin{array}{lcll} t_0 & = & \text{duration of jail} & \sim P(t_0) \\ t_1 & = & \text{duration of unemployment} & \sim P(t_1) \\ t_2|\tilde{w} & = & \text{duration of job, conditional on } \tilde{w} & \sim P(t_2|\tilde{w}) \end{array}$$

- Transition indicators:

$$\begin{array}{lcl} d_{1,e} & = & 1 \text{ if unemployment-to-job transition, otherwise} = 0 \\ d_{1,n} & = & 1 \text{ if unemployment-to-jail transition, otherwise} = 0 \\ d_{2,e} & = & 1 \text{ if job-to-job transition, otherwise} = 0 \\ d_{2,n} & = & 1 \text{ if job-to-jail transition, otherwise} = 0 \\ d_{2,u} & = & 1 \text{ if job-to-unemployment transition, otherwise} = 0 \end{array}$$

- Transition Probabilities:

$$\begin{array}{ll} \text{unemployment-to-job transition} & \sim P(e_{t_1}) \\ \text{unemployment-to-jail transition} & \sim P(n_{t_1}) \\ \text{job-to-job transition, conditional on } \tilde{w} & \sim P(e_{t_2}|\tilde{w}) \\ \text{job-to-jail transition, conditional on } \tilde{w} & \sim P(n_{t_2}|\tilde{w}) \\ \text{job-to-unemployment, conditional on } \tilde{w} & \sim P(u_{t_2}|\tilde{w}) \end{array}$$

The resulting likelihood function given by the model, dependent on the observed data (durations and transitions), is

$$l(\theta) = P(t_0)P(t_1)P(n_{t_1})^{d_{1,n}}[P(e_{t_1})f_c(w)P(t_2)\prod_{i=e,u,n} P(i_{t_2})^{d_{2,i}}]^{d_{1,e}}, \tag{1}$$

where $\theta = (\rho,\, \lambda_0,\, \lambda_1,\, \delta,\, \varphi,\, \mu\pi\overline{\Gamma}(g_u),\, \mu\pi\overline{\Gamma}(g_{e,L}),\, \mu\pi\overline{\Gamma}(g_{e,H}),\, \underline{w}_{L,c}, \underline{w}_{H,c},\, \overline{w}_{L,c},\, \overline{w}_{H,c})$.[13]

[12] The wage offer distribution for the criminal type is

$$f_c(w) = \begin{cases} \frac{\varphi(1+\kappa_1)}{2\kappa_1}\frac{1}{\sqrt{(p_1-w_{L,c})(p_1-\underline{w}_{L,c})}} & \text{if} \quad \underline{w}_{L,c} \le w \le \overline{w}_{L,c} \\ \frac{(1-\varphi)(1+\kappa_2)}{2\kappa_2}\frac{1}{\sqrt{(p_2-w_{H,c})(p_2-\underline{w}_{H,c})}} & \text{if} \quad \underline{w}_{H,c} \le w \le \overline{w}_{H,c}, \end{cases}$$

where $\kappa_1 = \frac{\lambda_1\varphi}{\delta+\mu\pi\overline{\Gamma}(g_{e,L})+\lambda_1(1-\varphi)}$ and $\kappa_2 = \frac{\lambda_1(1-\varphi)}{\mu\pi\overline{\Gamma}(g_{e,H})+\delta}$. The wage distribution is derived in the appendix including the closed form solutions for $\overline{w}_{L,c}$ and $\overline{w}_{H,c}$.

[13] Although suppressed in the text, the likelihood function accounts for both left and right censoring of durations. The simplified form is

$$\begin{array}{lcl} l(\theta) & = & \rho^{d_0^n} e^{-\rho t_0} e^{-(\mu\pi\overline{\Gamma}(g_u)+\lambda_0)t_2}(\mu\pi\overline{\Gamma}(g_u))^{d_{1,n}} \\ & & [\lambda_0 f_c(w) e^{-(\mu\pi\overline{\Gamma}(g_{e,i})+\lambda_1(1-F(\tilde{w}))+\delta)t_2}(\mu\pi\overline{\Gamma}(g_{e,i}))^{d_{2,n}}(\lambda_1(1-F(\tilde{w})))^{d_{2,e}}\delta^{d_{2,u}}]^{d_{1,e}}, \end{array}$$

where $d_0^n = 1$ if the duration of incarceration is not left censored.

I propose super-efficient estimators

$$\underline{w}_{L,c} = \min\{\tilde{w}\}, \quad \overline{w}_{L,c} = \tilde{w}_{\varphi}, \quad \underline{w}_{H,c} = \tilde{w}_{1-\varphi}, \quad \text{and } \overline{w}_{H,c} = \max\{\tilde{w}\}$$

where the estimators have been shown to be super-efficient and the theory of local cuts, Christensen & Kiefer (1994), justifies conditioning on these estimates.[14] The notation $\tilde{w}_{\varphi}$ represents the φ percentile of the observed wage distribution, or in other words, φ defines the point between low and high wage workers.

The parameters that are not individually estimated are $(b_c, b_{nc}, z, \mu, \pi, \Gamma(g))$. These estimates are unobtainable due to the inability to measure the value of individual crime opportunities, the number of crimes opportunities and the probability of being caught. However, if data on the total value of crime, the aggregate number of crimes committed and the value of the crime when an individual is caught were known, then the remaining parameters could be estimated. Derivation of the likelihood in such a situation is given in the Appendix as well as an explanation about how the currently estimated parameters and standard errors are left unchanged with the additional data.

3.3 Findings

The estimated parameter values are in Table 2.

The parameter estimates for the labor market frictions provide adequate reason to estimate the model as they are significantly different than what Burdett et al. (2004) use to evaluate a similar model. For instance, the job arrival rate implies that on average an individual, either previously incarcerated or not, takes roughly eighteen weeks to find employment. The estimate is nearly three times greater than the values found in Burdett et al. (2004), or Bowlus et al. (1995) who use the same data set but earlier in the panel. Also, the job separation rate is at least two times faster (higher) than estimates from similar models (van den Berg & Ridder (1998),Bontemps, Robin & van den Berg (2000)). On the other hand, the on-the-job arrival rate is close to the related literature.

Using the parameter estimates from the full sample, I estimate the steady state unemployment rate for non-criminals to be 8.5% while the unemployment rate for criminals is 12.5%. In addition, the average wage offer as calculated from the model using the parameter estimates is \$1,203 per

[14] These sample extremes serve as estimators of the unknown productivities and reservation wages as discussed in Bowlus et al. (1995). Asymptotically to order $N^{1/2}$ and ignoring the variability in estimates of the reservation wages and productivities, Kiefer & Neumann (1991) show that the bias from these estimates is ignorable for sample sizes over 200.

Table 2: Parameters Estimates by b_c, b_{nc}, and the Full Sample

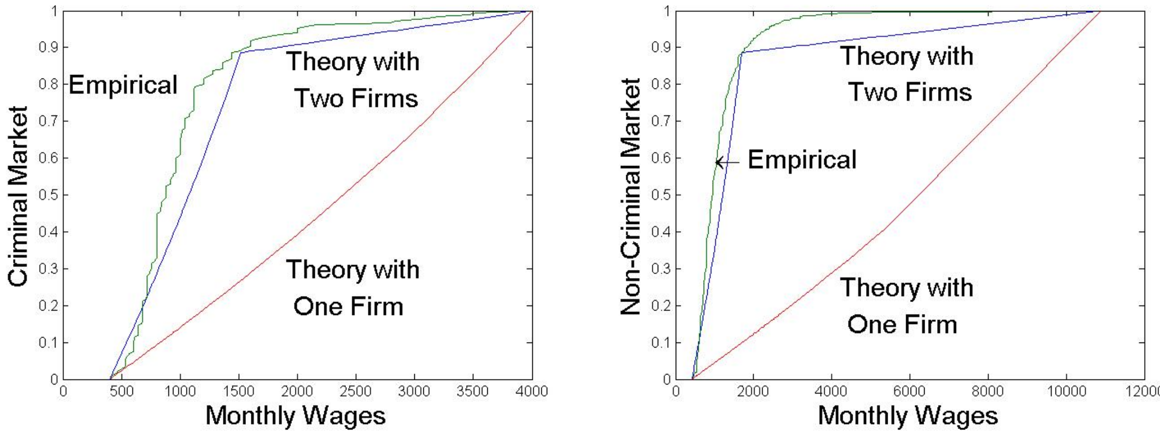
Parameters	Criminals	Non-Criminals	Full Sample
$\hat{\delta}$	0.022	0.020	0.020
	[0.017, 0.028]	[0.018, 0.021]	[0.018, 0.022]
$\hat{\lambda}_0$	0.226	0.213	0.215
	[0.190, 0.277]	[0.200, 0.231]	[0.202, 0.232]
$\hat{\lambda}_1$	0.013	0.011	0.011
	[0.008, 0.018]	[0.010, 0.014]	[0.010, 0.014]
$\hat{\varphi}$	0.794	0.880	0.880
	[0.761, 0.908]	[0.785, 0.897]	[0.783, 0.897]
$\hat{\rho}$	0.051		0.051
	[0.045, 0.057]		[0.045, 0.057]
$\widehat{\mu\pi\overline{\Gamma}}(g_u)$	0.026		0.026
	[0.020, 0.033]		[0.020, 0.033]
$\widehat{\mu\pi\overline{\Gamma}}(g_{e,L})$	0.010		0.012
	[0.007, 0.015]		[0.007, 0.015]
$\widehat{\mu\pi\overline{\Gamma}}(g_{e,H})$	0.011		0.006
	[0.003, 0.020]		[0.003, 0.019]

Note: Arrival rates are monthly. In the brackets are the 95% confidence interval from bootstrapping with 500 draws.

month for the criminal type and \$1687 for the non-criminal. Since the parameters are taken from the full sample and identical for each type, the difference in inequality and unemployment arises solely from the fact that criminals are separated from their jobs at a faster rate due to incarceration. In support of these results, I note Grogger (1995) takes a multi-stage regression approach and finds that those previously incarcerated face only a slightly tougher job market. Also, the difference in the average wage between the criminal and non-criminal agent is greater than what is observed in the data. However, I argue that using two types of firms allows the model to fit the wage inequality adequately as shown in Figure 2.

The key to the policy analysis comes from the estimation of $\mu\pi\overline{\Gamma}(g_u)$, $\mu\pi\overline{\Gamma}(g_{e,L})$ and $\mu\pi\overline{\Gamma}(g_{e,H})$. They identify the equilibrium described in Section 2. As shown in Table 2, the estimate for $\mu\pi\overline{\Gamma}(g_u)$ is significantly greater than $\mu\pi\overline{\Gamma}(g_{e,L})$ and $\mu\pi\overline{\Gamma}(g_{e,H})$. Therefore, I can reject Equilibrium 3 with confidence. For the full sample, I find evidence that agents are deterred from crime when paid a higher wage, or $\mu\pi\overline{\Gamma}(g_{e,H}) < \mu\pi\overline{\Gamma}(g_{e,L})$. All in all, I find evidence that employment, and potentially higher wages, deter individuals from crime. In other words, economic incentives

Figure 2: Wage Distribution



reduce criminal participation, or $\mu\pi\overline{\Gamma}(g_{e,H}) < \mu\pi\overline{\Gamma}(g_{e,L}) < \mu\pi\overline{\Gamma}(g_u)$.[15]

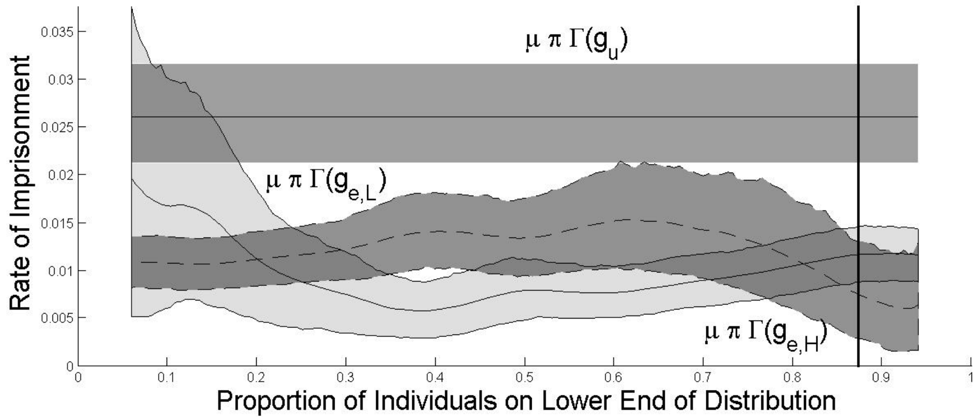
The criminal sub-sample are unable to reject Equilibrium 1 because of how $\mu\pi\overline{\Gamma}(g_{e,L})$ and $\mu\pi\overline{\Gamma}(g_{e,H})$ rely on the estimate of φ. However, φ is estimated through the wage dispersion. It might appear unreasonable to hinge estimates for criminal participation on considerations of the shape of the wage distribution. So, I plot $\mu\pi\overline{\Gamma}(g_{e,L})$ and $\mu\pi\overline{\Gamma}(g_{e,H})$ for $\varphi \in [.05, .95]$ where the end points are excluded due to an inadequate number of observations.

Figure 3 demonstrates how φ affects criminal participation. The x-axis is labeled "Proportion of Individuals on Lower End of (Wage) Distribution" because of the interpretation of the model. The interpretation is low productivity firms, or firms who pay lower wages, are differentiated at the point φ. Therefore, φ captures both the fraction of low productivity firms as well as what is referred to as the lower end of the wage distribution. Hence, the model is estimated with φ from low to high, and the corresponding incarceration rates are plotted.

It is interesting to see how wages affect criminal participation. Only at the upper and lower end of the distribution do I find evidence that higher wages deter crime, although the confidence interval is too wide at the lower end to provide any significant evidence.

[15]The model assumes the arrival rate of crime opportunities, μ, is identical for the unemployed and employed. However, one might think the unemployed have more time and therefore have a higher arrival rate. If so, the estimates of $\mu\pi\overline{\Gamma}(g_{e,H})$, $\mu\pi\overline{\Gamma}(g_{e,L})$ and $\mu\pi\overline{\Gamma}(g_u)$ would not change but rather the relative weight of μ and the acceptance probability $\overline{\Gamma}(g)$. The relative size of each factor is not identified and therefore no comparison is made between the arrival rates of the unemployed and employed.

Figure 3: Criminal Participation with Full Sample



Note: I have plotted the 15^{th} and 85^{th} percentiles as confidence bands from the bootstrap sample.

3.4 Demographics

Other research has found that employment frictions are dependent upon occupation, age, and/or education. In addition, crime is consistently found to be correlated to race, gender and an individual's location of residence. I use the same approach as van den Berg & Ridder (1998) to analyze how these variables are related to employment frictions and crime. Specifically, I introduce "heterogeneity by assuming that there are separate labor markets (or segments of the labor market, or sub-markets) for different types of individuals," (van den Berg & Ridder (1998) p. 1203). For example, there exist separate markets for those with different education levels. An alternative argument would be urban/rural areas tend to have thicker/thinner labor markets. The result is the model can be estimated separately by demographic.

The issue with this approach is the large number of observations needed in each category. For instance, I observe very few females or college graduates committing crime. Also, the sample ages in the initial period range from 24 to 32. Therefore, I limit the estimation by demographic in two ways. First, I look at demographics which can be split into two sub-groups, i.e. {Black,White}. Second, I restrict attention to demographics where each group retains at least 25% of the criminal observations. Table 3 contains the demographic breakdown of the sample.

Table 4 contains the estimation results by demographic. The results show, almost universally, that the unemployed are caught committing crime at a faster rate. Also, they demonstrate that rural individuals, black individuals, and those without an education find jobs at a slower rate and lose

Table 3: Demographic Sample Means

Variable	Criminals	Non-Criminals	Full Sample
Male	0.92	0.54	0.61
Married	0.13	0.33	0.29
Black	0.53	0.33	0.37
White	0.39	0.61	0.57
Urban	0.74	0.76	0.76
Rural	0.26	0.24	0.24
High School Diploma	0.51	0.74	0.7
Older than 28	0.43	0.44	0.44

their jobs faster. In addition, individuals facing a tougher job market tend to commit more crime when unemployed with the exception of high school graduates. The findings align themselves with other empirical studies. For instance, Grogger (1998) and many others find race is correlated with criminal participation. Also, Uggen (2000) finds older criminals have a lower likelihood of criminal participation when employed which is what I estimate for those employed at the upper end of the wage distribution. Surprisingly, I find those living in rural areas have a slightly higher incarceration rate which runs contrary to many findings. However, the results on criminal participation have large confidence bands due to the limited number of observations. Therefore, the results should be interpreted accordingly.

3.5 Model Evaluation

I have provided evidence that economic incentives play a role in criminal behavior. At this point, I evaluate how well the model predicts recidivism rates as well as the elasticity between crime, incarceration and the likelihood of apprehension.

Policy makers and researchers alike concern themselves with recidivism rates. Although not previously discussed, the model is set up to evaluate recidivism.[16] The prediction of the recidivism

[16] Recidivism can be defined as given a convicted criminal just left jail, what is the probability the criminal will return to jail within "t" periods. The calculation can be considered as a function of the Markov transition matrix between $(u_c, e_{L,c}, e_{H,c})$ where

$$P = \begin{pmatrix} (1-\lambda_0) & \lambda_0\varphi & \lambda_0(1-\varphi) \\ \delta & (1-\delta)(1-\lambda_1)+(1-\delta)\lambda_1\varphi & (1-\delta)\lambda_1(1-\varphi) \\ \delta & 0 & (1-\delta) \end{pmatrix}.$$

In words, the probability of going to jail in the t^{th} period, $\Phi(t)$, is dependent on if he gets caught, π, the opportunity to commit a crime, μ, if he takes it dependent on his employment state, $(\overline{\Gamma}(g_u)P^t_{1,1} + \overline{\Gamma}(g_{e,L})P^t_{1,2} + \overline{\Gamma}(g_{e,H})P^t_{1,3})$, and he

Table 4: Parameters Estimates by Demographic

	Race		Residence		High School Diploma		Age	
	Black	White	Urban	Rural	No	Yes	Under 28	Over 28
$\hat{\delta}$	0.026	0.017	0.020	0.020	0.023	0.019	0.02	0.020
	[0.023, 0.029]	[0.015, 0.019]	[0.018, 0.022]	[0.017, 0.024]	[0.020, 0.027]	[0.017, 0.021]	[0.018, 0.022]	[0.017, 0.022]
$\hat{\lambda}_0$	0.171	0.254	0.221	0.201	0.183	0.233	0.223	0.206
	[0.151, 0.191]	[0.234, 0.279]	[0.205, 0.24]	[0.175, 0.227]	[0.164, 0.205]	[0.216, 0.254]	[0.205, 0.247]	[0.186, 0.227]
$\hat{\lambda}_1$	0.015	0.011	0.014	0.009	0.012	0.011	0.013	0.011
	[0.011, 0.018]	[0.009, 0.013]	[0.010, 0.015]	[0.008, 0.016]	[0.010, 0.016]	[0.010, 0.014]	[0.010, 0.015]	[0.009, 0.015]
$\hat{\varphi}$	0.868	0.857	0.793	0.918	0.848	0.870	0.808	0.875
	[0.797, 0.923]	[0.766, 0.896]	[0.785, 0.904]	[0.735, 0.918]	[0.772, 0.880]	[0.793, 0.902]	[0.760, 0.923]	[0.775, 0.900]
$\hat{\rho}$	0.042	0.067	0.048	0.058	0.051	0.05	0.057	0.044
	[0.035, 0.051]	[0.055, 0.080]	[0.041, 0.056]	[0.046, 0.074]	[0.042, 0.060]	[0.043, 0.059]	[0.048, 0.066]	[0.036, 0.052]
$\mu\pi\hat{\overline{\overline{\Gamma}}}(g_u)$	0.033	0.015	0.024	0.035	0.021	0.031	0.032	0.018
	[0.024, 0.045]	[0.011, 0.021]	[0.019, 0.032]	[0.021, 0.060]	[0.014, 0.03]	[0.022, 0.044]	[0.025, 0.042]	[0.009, 0.030]
$\mu\pi\hat{\overline{\Gamma}}(g_{e,L})$	0.011	0.011	0.011	0.012	0.011	0.010	0.008	0.016
	[0.006, 0.017]	[0.005, 0.017]	[0.007, 0.016]	[0.004, 0.020]	[0.006, 0.019]	[0.005, 0.016]	[0.005, 0.014]	[0.008, 0.025]
$\mu\pi\hat{\overline{\Gamma}}(g_{e,H})$	0.015	0.002	0.009	0.007	0.022	0.002	0.011	0.005
	[0.005, 0.043]	[0.001, 0.016]	[0.002, 0.018]	[0.001, 0.074]	[0.006, 0.057]	[0.001, 0.010]	[0.002, 0.025]	[0.001, 0.026]

Note: Arrival rates are monthly. In the brackets are the 95% confidence intervals from bootstrapping with 500 draws.

rates for 6, 12, 24, and 36 months are in Table 5 using the estimated parameters from the full sample. I am unable to make an exact comparison between the model and the data because the two contain different measures of recidivism. Thus, I include several different measures found in the data. What I see is that the estimated model, which captures those who returned to prison for new and old offenses, generally lies between the U.S. recidivism rates of those reconvicted and those returned to prison for a new offense.

Table 5: Prediction of Recidivism Rates

	Model with Full Sample	U.S. Data: Rearrested	U.S. Data: Reconvicted	U.S. Data: Returned to prison for new offense
Rec_6 (%)	11.5	29.9	10.6	5.1
Rec_{12} (%)	18.7	44.1	21.5	10.6
Rec_{24} (%)	30.2	59.2	36.4	19.2
Rec_{36} (%)	39.9	67.4	46.2	25.8

Note: U.S. data are reported by the Bureau of Justice Statistics for the time period 1994-1997.

Besides recidivism, the model is capable of predicting the elasticity of crime with respect to the average time spent incarcerated ($\frac{1}{\rho}$) and the number incarcerated (n_c) from a change in ρ. The elasticities can be calculated from Table 6.[17] In general, changes in ρ can have two effects on deterring crime. This exercise captures the incapacitation effect (keeping criminals off the streets) and not the crime deterrence effect. The reason I do not evaluate the deterrence effect is because not all the parameters are identified in the estimation procedure. On the other hand, a sufficiently small change in ρ only alters the incapacitation effect and not the equilibrium/deterrence effect because $\Gamma(g)$ is discrete.

Table 6: Changes in the Duration of Incaceration (ρ)

	ρ = 0.031	0.051	0.071
Unemployed Criminals (%)	8.9	10	10.6
Incarcerated Criminals (%)	29.2	20.1	15.3
Employed Criminals (%)	61.9	69.9	74.1
Crime Index	88.54	100	105.99

Note: Crime is indexed because π is not uniquely identified.

did not go to jail in the previous periods. Analytically it is $\Phi(t) = \pi\mu(\overline{\Gamma}(g_u)P^t_{1,1} + \overline{\Gamma}(g_{e,L})P^t_{1,2} + \overline{\Gamma}(g_{e,H})P^t_{1,3})(1-Rec_t)$ where the recidivism rate, Rec_t, is he goes to jail before the t^{th} period, or $Rec_t = \sum_{i=1}^{t}\Phi(i)$.

[17] Estimates for the recidivism rate and wage distribution are excluded because they are not a function of ρ.

I find the elasticity of crime with respect to the average time spent incarcerated, $\frac{1}{\rho}$, to be -0.18. My findings are an improvement from previous models such as Burdett et al. (2004) as my results align themselves with several other empirically based studies outlined in Levitt (2004) who argues "Typical estimates of elasticities of crime with respect to expected punishment range from -0.1 and -0.4" (p. 178).

Alternatively, the elasticity of crime with respect to the prison population has been debated. The question's relevance is rooted in the costs prisons incur on state and federal budgets. My estimate of the elasticity of crime with respect to the population incarcerated due to a rise in ρ is -0.25 and is in line with Levitt (1996) who states "The elasticities with respect to prison populations range from -0.147 to -0.703." (p. 178).

The final comparison I make is how crime rates change with respect to the probability of being caught. I find the elasticity of crime with respect to the probability of apprehension to be -0.19 as seen in Table 7. How the estimate compares to the literature is not completely clear, as it is hard to measure. However, Levitt (1997) finds the elasticity of crime with respect to the number of police to be between -0.05 and -1.98. Therefore, if the apprehension technology is linear in the quantity of police, then my estimate is within the range of the crime literature albeit the range is large. Finally, it is important to point out that an increase in π destroys jobs. Table 7 shows an increase in the probability of being caught reduces the amount of employed criminals.

Table 7: Changes in the Likelihood of Apprehension (π)

	Likelihood of Apprehension Index		
	50	100	150
Unemployed Criminals (%)	9.4	10	10.4
Incarcerated Criminals (%)	10.9	20.1	27.8
Employed Criminals (%)	79.6	69.9	61.8
Crime Index	109.01	100	92.26
Rec_{12}	9.79	18.7	26.81
Rec_{36}	22.42	39.93	53.58
Average Wage of Criminals	1306.33	1202.67	1123.46

Note: Crime and the likelihood of apprehension are indexed because π is not uniquely identified.

To reiterate, the model shows it is able to accurately predict the appropriate elasticities of crime with respect to time incarcerated, the size of the prison population and the likelihood of apprehension. In addition, the model is in the range of U.S. recidivism rates while accurately capturing crime, inequality, and unemployment. The results are innovative within the crime literature because the predictions are based on estimates from a structural model.

4 Policy Discussion

The main result highlighted in the introduction is that reductions in labor market frictions can reduce crime. Although the model does not explain how frictions are reduced, Wilson, Gallagher, Coggeshall & MacKenzie (1999), Visher et al. (2005) and others analyze the effectiveness of specific job placement programs. Alternatively, I calculate in Table 8 the effects of a *successful* program. Specifically, how does crime change given a placement program that works. I find the elasticity of crime with respect to the average time unemployed ($\frac{1}{\lambda_0}$) to be 0.11. Also, recidivism falls as it becomes easier to find a job. The reason is individuals commit half as much crime when employed. Therefore, if released inmates find jobs faster, then they commit fewer crimes and do not return as quickly. In effect, reducing employment frictions by half could reduce crime and recidivism by more than five percent. Note that the effect is purely employment driven, as the average wage is constant.

The elasticity of a job placement program might seem "small." However, the policy is primarily affecting a small part of the population, the unemployed. In addition, U.S. law enforcement observes roughly ten million crimes per year. Therefore, a five percent reduction would eliminate more than a half of a million crimes annually. Finally, other anti-crime policies have a "small" elasticity such as an increase in incarceration. Therefore, the costs are essential in evaluating the success of a job placement program.

Table 8: Changes in the Unemployed Job Arrival Rate (λ_0)

	λ_0		
	0.108	0.215	0.43
Unemployed Criminals (%)	17.4	10	5.4
Incarcerated Criminals (%)	21.9	20.1	18.9
Employed Criminals (%)	60.7	69.9	75.6
Crime Index	109	100	94.39
Rec_{12} (%)	21.6	18.7	16.1
Rec_{36} (%)	44.5	39.9	36.9
Average Wage of Criminals	1202.67	1202.67	1202.67

Note: Crime is indexed because π is not uniquely identified.

Job placement programs can take many different forms. Consider the costs and benefits of a residential re-entry center (RRC). RRC's provide a structured environment for convicts being released from jail. In particular, they limit the time individuals are outside of the center. The time individuals are allowed outside of the center is closely monitored and used mainly for job searching or employment.

The cost-benefit analysis of an RRC using the estimated model is insightful. The costs of an RRC can range widely between one and three thousand dollars per person per month depending upon the location and environment. The benefits can range widely as well. As shown above, the benefit an RRC provides in reducing crime depends upon an individual's labor force status because the unemployed commit twice as much crime as those employed. In evaluating the benefits of an RRC, assume they deter agents from committing crime, and the probability of being caught and incarcerated is 2.5%. In addition, the average cost of a crime (excluding murder) is roughly $4,255.[18] The bottom line is benefits from an RRC, or the reduction in crime, is worth $4,425 per month for those unemployed, as they commit roughly 1.03 crimes per month. In addition, the benefits for an employed criminal type is between $1,020-2,040, as they would commit on average between 0.24 and 0.48 crimes per month depending upon their wage. Therefore, the benefits discussed above are greater than the costs when the resident of an RRC is unemployed but not necessarily while employed.

The benefits might not outweigh the costs of holding an employed individual in an RRC. However, what is the effect of an increase in the job finding rate for an employed individual? The purpose would be to reduce the crime rate of the employed by finding them higher paying jobs. The results are found in Table 9. First, notice the expected wage rises as individuals are finding better paying jobs at a faster rate. Second, the policy has very little affect on the short run recidivism rate, as those exiting jail take time to find their first job let alone a second. Third, the equilibrium crime rate falls only slightly, as the number of individuals finding a high productivity firm, or "low crime" job, is small. As a result, halving the time it takes to receive a new job offer reduces the crime rate by roughly 1%.

5 Conclusion

I have proposed an on-the-job search model of crime that incorporates heterogeneous agents, firms and crime opportunities. The heterogeneity allows for a more efficient equilibrium, a better esti-

[18] The average cost is taken from Cohen (1988) and adjusted for inflation using the CPI.

Table 9: Changes in the Employed Job Arrival Rate (λ_1)

	λ_1		
	0.006	0.011	0.022
Unemployed Criminals (%)	10.1	10	10
Incarcerated Criminals (%)	20.2	20.1	19.8
Employed Low Wage Criminals (%)	58.3	57	54.7
Employed High Wage Criminals (%)	11.5	12.9	15.5
Crime Index	100.68	100	98.72
Rec_{12} (%)	18.7	18.7	18.7
Rec_{36} (%)	40	39.9	39.8
Average Wage of Criminals	933.59	1202.67	1655.25

Note: Crime is indexed because π is not uniquely identified.

mate of wage dispersion, for agents to commit crime at different rates conditional on employment and wages, and results in a proportion of agents declining criminal opportunities based on their individual value of leisure.

Furthermore, I have developed a procedure to estimate the model. The major result being economic incentives, in particular employment frictions and wages, affect crime.

Given the incentives between crime, unemployment and wage inequality, I argue policies aimed at reducing employment frictions can improve the labor market for criminals and in turn decrease crime. I find that a successful job placement program, one capable of cutting the average length of unemployment by two months, can reduce crime and recidivism by more than five percent.

A Appendix

A.1 Wage dispersion

At this point, I develop the remaining part of the model required in estimation. Specifically, I develop $F_c(w)$. The development of $F_c(w)$ follows from the work of Mortensen (1990) and Burdett et al. (2004). The distribution of wages posted for the non-criminal market, $F_{nc}(w)$, is derived in Mortensen (1990) and therefore excluded. As the "non-criminal labor market" has been previously derived, I will drop the c subscript on R, w and $F(w)$. In other words, I am analyzing the criminal market with two types of firms and one type of agent. Proposition 1 justifies why only one type of agent is needed to estimate the "criminal labor market."

To begin, I provide Lemma 1 in order to characterize $F(w)$. Lemma 1 is stated generically with the understanding that some cases are vacuous. Let $\underline{w}$ be the lower support of $F(w)$, $\underline{w}_i$ and $\overline{w}_i$ is the lowest and highest wage paid by a firm of type i, $\Pi_i(\cdot)$ is firm i profit, $C(g)$ is the reservation wage necessary to deter an agent from crime g, and $L(w)$ is the amount of labor a firm retains in the steady-state when offering a wage w.

Lemma 1 *1. $F(w)$ has no mass points, 2. $\underline{w}_L = R$ or $\underline{w}_L = C(g)$ for a $g \in \mathscr{G}$, 3. There are no gaps in $F(w)$ except on the intervals $(C(g) - \varepsilon_j, C(g))$, where $\varepsilon_j > 0$ for all $g \in \mathscr{G}$, 4. $\overline{w}_L = \underline{w}_H$ or $\underline{w}_H = C(g)$, and 5. $L(w)$ is increasing in w.*

Proof.

1. Suppose $F(w)$ has a mass point at w'. By offering $w' + \varepsilon$ the firm would increase their labor market supply by a discrete amount, implying $w' + \varepsilon$ has a larger profit and firms deviate to the slightly higher wage until no mass point exists.

2. Suppose all firms offer $R + \varepsilon \leq w$ where $\varepsilon > 0$. A firm paying $R + \varepsilon$ would deviate to paying R because they would be paying the worker less while they would lose workers at the same rate, therefore increasing profits. The identical argument can be made but replacing R with $C(g)$.

3. Suppose a firm offers $w' - \varepsilon$ arbitrary close to $C(g)$, then by offering $C(g)$ the firm could decrease their job destruction rate by a discrete amount because they would not be losing

workers to prison, implying an increase in profits. Therefore, no firms offer wages within $(C(g)-\varepsilon_j, C(g))$.

4. Same argument as Lemma 1.3.

5. First, considering $L(w) \in [\underline{w}_L, \overline{w}_L]$ or $[\underline{w}_H, \overline{w}_H]$ separately allows one to apply the proof of Mortensen (1990) Proposition 3. Second, $L(C(g)-\varepsilon_j) < L(C(g))$ holds because for both to exist then $\Pi_i(C(g)-\varepsilon_j) \leq \Pi_i(C(g))$ and $C(g)-\varepsilon_j < C(g)$ implying $L(C(g)-\varepsilon_j) < L(C(g))$.

■

I refer several times in the body of the text to the fact that higher productive firms pay more than lower productive firms. Here, I restate Proposition 3 as

Lemma 2 $p_L < p_H \Rightarrow w_L \leq w_H$.

Proof.
The proof is in two steps. First, I show it holds for wages offered within $[C(g), C(g')]$ where $g < g'$ on the support $\mathscr{G}$. Second, I show it holds across any crime wage $C(g)$ for $g \in \mathscr{G}$.

1. For firms L and H offering wages in $[C(g), C(g')]$ then $\overline{w}_L \leq \underline{w}_H$ because $L(w) \in [\underline{w}_L, \overline{w}_L]$ or $[\underline{w}_H, \overline{w}_H]$ is continuous and increasing in w implying Mortensen (1990) Proposition 3 applies.

2. Suppose there are two firms (L and H), paying above and below $C(g)$, then $\exists\ \underline{w}_{C(g)} < C(g)$ and $C(g) < \overline{w}_{C(g)}$ such that $\Pi_L(\underline{w}_{C(g)}) = \Pi_L(\overline{w}_{C(g)})$ and $\Pi_H(\underline{w}_{C(g)}) = \Pi_H(\overline{w}_{C(g)})$, but if

   $(p_L - \underline{w}_{C(g)})L(\underline{w}_{C(g)}) = (p_L - \overline{w}_{C(g)})L(\overline{w}_c) \Rightarrow$

   $\overline{w}_{C(g)}L(C) - \underline{w}_{C(g)}L(\underline{w}_{C(g)}) = p_L(L(\overline{w}_{C(g)}) - L(\underline{w}_{C(g)})) \Rightarrow$

   $\overline{w}_{C(g)}L(\overline{w}_{C(g)}) - \underline{w}_{C(g)}L(\underline{w}_{C(g)}) < p_H(L(\overline{w}_{C(g)}) - L(\underline{w}_{C(g)})) \Rightarrow$

   $(p_H - \underline{w}_{C(g)})L(\underline{w}_{C(g)}) < (p_H - \overline{w}_{C(g)})L(\overline{w}_{C(g)})$

   $\Rightarrow$ no H firm would offer $\underline{w}_{C(g)}$.

■

I have proven $F(w)$ is continuous on the support except below the points $C(g)$ for all $g \in \mathscr{G}$. From Lemma 2, I will break down the distribution of $F(w)$ into parts

$$F^i(w_i) = F(w|\underline{w}_i \leq w \leq \overline{w}_i),\ for\ i \in \{L,H\} \tag{2}$$

It is key to derive the wages being paid, $G(w)$, in order to back out $F(w)$. Therefore, define $G(w)$ as

$$G^i(w_i) = G(w|\underline{w}_i \leq w \leq \overline{w}_i),\ for\ i \in \{L,H\} \tag{3}$$

where $G(w)$ is defined explicitly for the two firms as[19]

$$\begin{aligned} G^L(w_L) &= \frac{F^L(w_L)}{(1+\kappa_1(1-F^L(w_L))}, \textit{ and} \\ G^H(w_H) &= \frac{F^H(w_H)}{1+\kappa_2(1-F^H(w_H))}, \end{aligned} \tag{4}$$

where $\kappa_1 = \frac{\lambda_1\varphi}{\delta+\mu\pi\overline{\Gamma}(g_{e,L})+\lambda_1(1-\varphi)}$ and $\kappa_2 = \frac{\lambda_1(1-\varphi)}{\mu\pi\overline{\Gamma}(g_{e,H})+\delta}$.

The rest of the necessary steps in attaining $F(w)$ in closed form can be summarized in two steps. The first is to derive $L(w)$ from $L^L(w_L) = e_{L,c}\frac{\frac{dG^L(w_L)}{dw}}{\frac{dF^L(w_L)}{dw}}$ and $L^H(w_H) = e_{H,c}\frac{\frac{dG^H(w_H)}{dw}}{\frac{dF^H(w_H)}{dw}}$. The second and final step in attaining $F(w)$ plugs $L(w)$ into the profit function of a firm, $\Pi_i(w_i)$, then sets the equilibrium condition that firms of the same type make the same profit, $\Pi_i(\underline{w}_i) = \Pi_i(w_i)$, along with $F^i(\underline{w}_i) = 0$. The result is

$$\begin{aligned} F^L(w_L) &= \frac{1+\kappa_1}{\kappa_1}\left(1-\sqrt{\frac{(p_L-w_L)}{(p_L-\underline{w}_L)}}\right), \\ F^H(w_H) &= \frac{1+\kappa_2}{\kappa_2}\left(1-\sqrt{\frac{(p_H-w_H)}{(p_H-\underline{w}_H)}}\right), \end{aligned} \tag{5}$$

and $\overline{w}_L$ and $\overline{w}_H$ are derived using $F^L(\overline{w}_L) = 1$ and $F^H(\overline{w}_H) = 1$, or

[19] Setting the time derivatives equal to zero gives you equation 4

$$\frac{d}{dt}G^L(w_L)e_{L,c} = \lambda_0 u_k\varphi F^L(w_L) - (\delta+\mu\pi\overline{\Gamma}(g_{e,L})+\lambda_1(1-\varphi)+\lambda_1\varphi(1-F^L(w_L)))e_{L,c}G^L(w_L),$$

$$\frac{d}{dt}G^H(w_H)e_{H,c} = (\lambda_0 u_k+\lambda_1 e_{L,c})(1-\varphi)F^H(w_H) - (\delta+\mu\pi\overline{\Gamma}(g_{e,H})+\lambda_1(1-\varphi)(1-F^H(w_H)))e_{H,c}G^H(w_H),$$

where the steady state flows are:

$$\begin{aligned} u_k &= (\delta+\mu\pi\overline{\Gamma}(g_{e,L})+\lambda_1(1-\varphi))\rho(\delta+\mu\pi\overline{\Gamma}(g_{e,H}))/\Omega, \\ e_{L,c} &= \lambda_0\varphi\rho(\delta+\mu\pi\overline{\Gamma}(g_{e,H}))/\Omega, \\ e_{H,c} &= (\delta+\mu\pi\overline{\Gamma}(g_{e,L})+\lambda_1)(1+\varphi)\rho\lambda_0/\Omega, \end{aligned}$$

and $\Omega = u_k + n_k + e_{L,c} + e_{H,c}$.

$$\overline{w}_L = p_L - (p_L - \underline{w}_L)(\tfrac{1}{1+\kappa_1})^2,$$
$$\overline{w}_H = p_H - (p_H - \underline{w}_H)(\tfrac{1}{1+\kappa_2})^2. \tag{6}$$

Therefore, the necessary equation in estimation is

$$F_c(w) = \begin{cases} 0 & \text{if} \quad w < \underline{w}_L \\ \varphi F^L(w_L) & \text{if} \quad \underline{w}_L \leq w \leq \overline{w}_L \\ \varphi & \text{if} \quad \overline{w}_L \leq w \leq \underline{w}_H \\ \varphi + (1-\varphi)F^H(w_H) & \text{if} \quad \underline{w}_H \leq w \leq \overline{w}_H \\ 1 & \text{if} \quad \overline{w}_H < w \end{cases}, \tag{7}$$

which is continuous by Lemma 1 in the support as defined in Equation 7 and differential except at the points $\underline{w}_i$ and $\overline{w}_i$ for $i \in \{L, H\}$.

A.2 Candidate Likelihood Function

In this section, I derive a candidate likelihood function that identifies all of the parameters including $(b_c,\, b_{nc},\, z,\, \pi,\, \Gamma(g))$.[20] In addition, I discuss how the parameters I have already estimated are left unchanged when estimating the unknown parameters.

The likelihood function requires two additional types of data. The data necessary to estimate the remaining parameters are

1. the value of crime when an individual is caught,
2. the aggregate number of crimes committed, $\mathscr{B}$.

In effect, $\Gamma(g)$ can be deduced using data about the value of the crime when an individual is caught. The discrete pdf $\gamma(g)$ of $\Gamma(g)$ is incorporated into the likelihood as

$$\begin{aligned} l(\theta) \;\; = \;\; & \rho^{d_0^n} e^{-\rho t_0} e^{-(\mu\pi\overline{\Gamma}(g_u)+\lambda_0)t_2} (\mu\pi\overline{\Gamma}(g_u)\gamma(g))^{d_{1,n}} \\ & [\lambda_0 f_c(w) e^{-(\mu\pi\overline{\Gamma}(g_{e,i})+\lambda_1(1-F(\tilde{w}))+\delta)t_2} (\mu\pi\overline{\Gamma}(g_{e,i})\gamma(g))^{d_{2,n}} (\lambda_1(1-F(\tilde{w})))^{d_{2,e}} \delta^{d_{2,u}}]^{d_{1,e}}, \end{aligned} \tag{8}$$

where $\gamma(g)$ is identified using a clustering type method up to the relative frequency of each occurrence on the domain $\mathscr{G}$.

Next, π can be deduced by the aggregate moment

[20] μ can be normalized to a sufficiently large number without loss of generality. Therefore, it is excluded from the discussion.

$$\pi = \frac{\mu\pi\hat{\overline{\Gamma}}(g_u)u_c + \mu\pi\hat{\overline{\Gamma}}(g_{e,L})e_{L,c} + \mu\pi\hat{\overline{\Gamma}}(g_{e,H})e_{H,c}}{\mathscr{B}},$$

where the parameters in the numerator are identified by Equation 8. This argument is similar to Flinn (2006), who uses an aggregate measure of firm profits to estimate one parameter of the model, specifically bargaining power.

Finally, $(b_c,\ b_{nc},\ z)$ can be deduced from the previously identified parameters and the super-efficient estimators $\underline{w}_{L,c}$ and $\underline{w}_{H,c}$. The reservation wages $[C(g_{e,L}), C(g_{e,H})]$ or $[R, C(g_{e,H})]$ are deduced from the model. To find them, realize the flow Bellman equations of the incarcerated, unemployed and those employed at the low and high wage firms are

$$rJ = z + \rho(V_0 - J), \tag{9}$$

$$rV_0 = b + \mu \sum_{g>g_u} \gamma(g)(g + \pi(J - V_0)) + \lambda_0 \Delta(C_L), \tag{10}$$

$$rV_1 = C(g_{e,L}) + \delta(V_0 - V_1) + \mu \sum_{g>g_{e,L}} \gamma(g)(g + \pi(J - V_1)) + \lambda_1 \Delta(C_L), \tag{11}$$

$$rV_2 = C(g_{e,H}) + \delta(V_0 - V_2) + \mu \sum_{g>g_{e,H}} \gamma(g)(g + \pi(J - V_2)) + \lambda_1 \Delta(C_H), \tag{12}$$

respectively, where

$$\Delta(C_i) = \int_{R_c} \frac{1 - F(x)}{r + \delta + \mu\pi(1 - \Gamma(g_{e,i})) + \lambda_1(1 - F(x))} dx, \text{ or}$$

$$\Delta(C_i) = \int_{C(g_{e,i})} \frac{1 - F(x)}{r + \delta + \mu\pi(1 - \Gamma(g_{e,i})) + \lambda_1(1 - F(x))} dx,$$

depending upon if $\mu\pi\hat{\overline{\Gamma}}(g_u) = \mu\pi\hat{\overline{\Gamma}}(g_{e,L})$ or not, respectively. Also, the threshold values of crime opportunities are

$$V_1 = J + \frac{g_{e,L}}{\pi},$$

$$V_2 = J + \frac{g_{e,H}}{\pi}.$$

As a result, Equation 9-12 can be reduced to two equations and two unknowns, (b_c, z), given the estimated parameters and setting $C(g_{e,L}) = \underline{w}_{L,c}$ and $C(g_{e,H}) = \underline{w}_{H,c}$. I leave it to the reader to show they identify both (b_c, z) given $g_{e,L} \neq g_{e,H}$. Also, it should be noted that identification of the remaining two parameters is dependent upon the type of equilibrium deduced in the first stage. For example, the reservation wage should be used as one of the equilibrium conditions to deduce (b_c, z) if $\overline{\Gamma}_u = \overline{\Gamma}_{e,L}$. Finally, b_{nc} is deduced from the non-criminal reservation wage.

To conclude, the reader should realize the estimates of these parameters will not affect the ones already estimated in the body of the paper. To see this, notice estimates for $\Gamma(g)$ are independent of the rest of the likelihood. In addition, π is estimated from one aggregate moment. Finally, $(b_c,\ z)$ are found using the restrictions of the model. Therefore, the estimates of $(\hat{\rho},\ \hat{\lambda}_0,\ \hat{\lambda}_1,\ \hat{\delta},\ \hat{\varphi},\ \widehat{\mu\pi\overline{\Gamma}}(g_u),\ \widehat{\mu\pi\overline{\Gamma}}(g_{e,L}),\ \widehat{\mu\pi\overline{\Gamma}}(g_{e,H}))$ found in Section 3 are independent of the estimates of $(b_c,\ b_{nc},\ z,\ \pi,\ \Gamma(g))$.

References

Bontemps, C., Robin, J. & van den Berg, G. J. (1999), ‘An Empirical Equilibrium Job Search Model with Search on the Job and Heterogeneous Workers and Firms’, *International Economic Review* **40**, 1039–74.

Bontemps, C., Robin, J. & van den Berg, G. J. (2000), ‘Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation’, *International Economic Review* **41**, 305–58.

Bowlus, A. J., Kiefer, N. M. & Neumann, G. R. (1995), ‘Estimation of Equilibrium Wage Distributions with Heterogeneity’, *Journal of Applied Econometrics* **10**, S119–S131.

Burdett, K., Lagos, R. & Wright, R. (2003), ‘Crime, Inequality, and Unemployment’, *American Economic Review* **93**, 1764–77.

Burdett, K., Lagos, R. & Wright, R. (2004), ‘An On-the-Job Search Model of Crime, Inequality, and Unemployment’, *International Economic Review* **45**, 681–706.

Burdett, K. & Mortensen, D. T. (1998), ‘Wage Differentials, Employer Size and Unemployment’, *International Economic Review* **39**, 257–73.

Christensen, B. J. & Kiefer, N. M. (1994), ‘Local Cuts and Separate Inference’, *Scandinavian Journal of Statistics* **21**, 389–401.

Chung, C., Schmidt, P. & Witte, A. D. (1991), ‘Survival Analysis: A Survey’, *Journal of Quantitative Criminology* **7**, 59–98.

Cohen, M. (1988), ‘Pain, Suffering, and Jury Awards: A Study of the Cost of Crime to Victims’, *Law and Society Review* **22**, 537–556.

Eckstein, Z. & Wolpin, K. I. (1995), ‘Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model’, *The Review of Economic Studies* **62**, 263–286.

Engelhardt, B., Rocheteau, G. & Rupert, P. (2007), Crime and the Labor Market: A Search Model with Optimal Contracts. Federal Reserve Bank of Cleveland, Working Paper 0715.

Flinn, C. J. (2006), ‘Minimum Wage Effects on Labor Market Outcomes Under Search, Matching, and Endogeneous Contact Rates’, *Econometrica* **74**, 1013–1062.

Gould, E. D., Weinberg, B. A. & Mustard, D. B. (2002), ‘Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997’, *The Review of Economics and Statistics* **84**, 45–61.

Grogger, J. (1995), ‘The Effect of Arrests on the Employment and Earnings of Young Men’, *The Quarterly Journal of Economics* **110**, 51–72.

Grogger, J. (1998), ‘Market Wages and Youth Crime’, *Journal of Labor Economics* **16**, 756–91.

Imrohoroğlu, A., Merlo, A. & Rupert, P. (2000), ‘On the Political Economy of Income Redistribution and Crime’, *International Economic Review* **41**, 1–25.

Kiefer, N. M. & Neumann, G. R. (1991), Estimation of Equilibrium Wage Distributions with Heterogeneity. Princeton University, Mimeo.

Levitt, S. D. (1996), ‘The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation’, *The Quarterly Journal of Economics* **111**, 319–51.

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

Levitt, S. D. (2004), ‘Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not’, *Journal of Economic Perspectives* **18**, 163–90.

Mortensen, D. T. (1990), *Equilibrium Wage Distributions: A Synthesis*, J. Hartog, G. Ridder, and J. Theeuwes edn, North-Holland, New York, pp. 279–96.

Piehl, A. & DiIulio, J. (1995), ‘Does Prison Pay? Revisited’, *The Brookings Review* (Winter), 21–25.

Uggen, C. (2000), ‘Work as a Turning Point in the Life Course of Criminals: A Duration Model of Age, Employment, and Recidivism’, *American Sociological Review* **65**, 529–546.

van den Berg, G. J. & Ridder, G. (1998), ‘An Empirical Equilibrium Search Model of the Labor Market’, *Econometrica* **66**, 1183–1221.

Visher, C. A., Winterfield, L. & Coggeshall, M. B. (2005), 'Ex-offender Employment Programs and Recidivism: A Meta-Analysis', *Journal of Experimental Criminology* **1**, 295–316.

Wilson, D. B., Gallagher, C. A., Coggeshall, M. B. & MacKenzie, D. L. (1999), 'A Quantitative Review and Description of Corrections-Based Education, Vocation, and Work Programs.', *Corrections Management Quarterly* **3**, 8–18.