Journal of Urban Economics 68 (2010) 231–246

Contents lists available at ScienceDirect

Journal of Urban Economics

www.elsevier.com/locate/jue

Thieves, thugs, and neighborhood poverty

David Bjerk

Robert Day School of Economics and Finance, Claremont McKenna College, 500 East Ninth Street, Claremont, CA 91711, United States

ARTICLE INFO

Article history:
Received 12 November 2009
Revised 16 June 2010
Available online 23 June 2010

Keywords:
Crime
Segregation
Neighborhood effects
Instrumental variables
Poverty

ABSTRACT

This paper develops a model of crime analyzing how such behavior is associated with individual and neighborhood poverty. The model shows that even under relatively minimal assumptions, a connection between individual poverty and both property and violent crimes will arise, and moreover, "neighborhood" effects can develop, but will differ substantially in nature across crime types. A key implication is that greater economic segregation in a city should have no effect or a negative effect on property crime, but a positive effect on violent crime. Using IV methods, I show this implication to be consistent with the empirical evidence.

> I don't care if I got money, or work Monday through Friday. I just go shoot a motherf*@#er on the weekends. If that's what need to be done to keep my hood and my young ones around here safe, then that's what to get done (quoted by Landesman, 2007).

1. Introduction

High rates of crime and violence in poor neighborhoods have been described by numerous scholars and journalists (Wilson, 1987; Krivo and Peterson, 1996; Kotlowitz, 1991; Patterson, 1991; Messner and Tardiff, 1986, to name just a few). However, the quote above from a man residing in a high-poverty housing project in south Los Angeles emphasizes that not only is crime a large part of life in high-poverty neighborhoods, but also that violent crimes may often serve a quite different purpose than basic property crime. Namely, while the motivation for basic property crimes is generally monetary, becoming involved in violent crime may have a defensive motivation as well.

While this defensive motivation for violence has long been recognized by sociologists (Massey, 1995; Anderson, 1999) and more recently by economists (Silverman, 2004; O'Flaherty and Sethi, 2010, this issue), the mechanisms through which such motivations are exacerbated by individual and neighborhood poverty are less well understood. This paper attempts to explicitly model some of the key distinctions between participation in violent crimes versus basic property crimes, as well as the ways such participation decisions might be affected by an individual's own economic circumstances as well as the economic circumstances of his neighbors. The paper also considers the broader implications of how these forces interact with the degree to which a city is segregated by economic status to affect overall crime rates.

While there may be numerous, and possibly quite complex, paths through which poverty and neighborhood characteristics may affect criminal behavior, the models developed here trace out the implications of two quite simple assumptions regarding behavior and criminal interactions. The first is simply that individuals incur diminishing marginal utility in money. The second is with respect to how violent crimes differ from basic property crimes. Specifically, I first develop a model of participation in basic property crimes where I assume that by choosing to become a thief at any given period in time, an individual will simply have additional consumption beyond his legal income in that period. Alternatively, in the model of participation in violent crimes, I assume that individuals can choose to become either a thug (i.e., an individual who engages in violence) or a law-abider (i.e., an individual who does not engage in violence). This will mean that when two law-abiders encounter one another in their neighborhood, they pass each other without incident. On the other hand, when a law-abider encounters a thug, the thug will attack him and take some of his money. However, when two thugs encounter one another, violence can still ensue but not necessarily with certainty. More importantly though, even when violence does ensue in such encounters, I assume that neither is able to take money from the other. In this way, choosing to be a thug can not only serve offensive purposes, allowing thugs to take money from law-abiders in their neighborhood, but also be defensive in terms of ensuring ones' own money will not be taken by other thugs and potentially preventing an attack from occurring in the first place.

E-mail address: david.bjerk@cmc.edu

0094-1190/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jue.2010.06.002

As will be shown below, the above assumptions are sufficient to ensure that poor individuals are more likely to become both thieves and thugs than an otherwise similar population of non-poor individuals. With respect to becoming a thief, the intuition is quite straightforward. The diminishing marginal utility of money assumption will mean a poor individual will value a fixed amount of money above and beyond his legal income more than will a rich individual, meaning that poor individuals will have a stronger incentive than non-poor individuals to steal, all else equal.

However, without additional assumptions, an individual will be no more or less likely to become a thief if he lives in an extremely poor neighborhood versus a richer neighborhood, meaning no "neighborhood effects" will arise when it comes to basic property crimes. However, if it is further assumed that the return to theft is greater when a greater fraction of one's neighbors are non-poor, "neighborhood effects" do arise, with individuals becoming more likely to become thieves the smaller the fraction of their neighbors that are poor. Therefore, all else equal, this model suggest that the degree to which a city's poor residents are residentially segregated from the city's non-poor residents will either have no effect or even a negative effect on the city's overall rate of basic property crime.

On the other hand, beyond the two basic assumptions discussed above, no further assumptions are necessary for neighborhood effects to arise with respect to violent crimes, as an individual's motivation to become a violent person (i.e., a thug) depends on the likelihood he will run into other thugs in his neighborhood, which in turn can depend on the level of poverty in his neighborhood. In fact, the model actually shows that the effect of neighborhood poverty on individual incentives to become a violent person becomes increasingly stronger as the neighborhood poverty rate rises, which in turn implies that the more a city is segregated by poverty status, the greater will be the overall rate of violent crime—the opposite of what was true with respect to basic property crimes.

The latter part of the paper then empirically examines these two key implications of the model at the citywide level—namely that all else equal, greater (exogenous) economic segregation should have either a negligible or negative impact on basic property crimes such as burglary, larceny and motor vehicle theft, but should have a positive impact on interpersonal violent crimes such as robbery and aggravated assault. Using MSA level data for the United States from the year 2000, I find support for these implications. In particular, after I instrument for the economic segregation for each MSA using data on how public housing is allocated in the MSA, the fraction of local public funds in the MSA coming from the state or federal government, the number of municipal governments in the MSA, and the number of larger rivers in the MSA, I find greater economic segregation has a negative but somewhat imprecisely estimated effect on burglary, a negligible effect on larceny and motor vehicle theft, and a positive and significant effect on robbery and aggravated assault.

2. Related literature

The theoretical model developed below relates to three particular streams of literature. The first is that of neighborhood effects, where individual behavior is directly affected by the characteristics and/or actions of his neighbors. For example, preferences for peer conformity may alter an individual's taste for engaging in crime (Glaeser et al., 1996; Brock and Durlauf, 2001), or an individual's information about payoffs to crime may evolve differently depending on the number of criminals in his neighborhood (Heavner and Lochner, 2002; Calvo-Armengol and Zenou, 2004; Calvo-Armengol et al., 2007; Patacchini and Zenou, 2008). Somewhat relatedly, one individual's criminal behavior may create a positive externality for other potential criminals in that one person's criminal behavior taxes fixed police resources, which in turn lowers the probability of detecting/arresting other criminals (Ferrer, 2010), or similarly, one neighborhoods' level of private policing may affect the relative payoffs to crime in another neighborhood (Helsley and Strange, 2005). While these papers describe how "neighborhood effects" can arise given these assumptions, where a given individual's optimal behavior may be depend on which neighborhood he lives in, the model developed below takes a step back to consider how individual and neighborhood level poverty in and of themselves can affect criminal behavior even in absence of the types of assumptions discussed above.

The second area of research this model builds on is the literature on the relationship between crime and segregation. In this vein, Verdier and Zenou (2004) develop a model of labor market discrimination, crime, and *racial* segregation. While their model can imply a correlation between crime and neighborhood income, it does not lead to "neighborhood effects" per say in that the actual characteristics of the an individual's neighbors do not directly affect his own criminality. O'Flaherty and Sethi (2007) also model the relationship between racial segregation and a particular type of crime—namely robbery. An important contribution of this model is that criminal activity (namely robbery) and racial segregation are simultaneously determined, with both affecting the other. Moreover, individual criminal behavior is affected by the racial/income characteristics of his neighbors. However, a key implication of O'Flaherty and Sethi's model is that while higher robbery rates can lead to greater segregation as the wealthier whites move away from the poorer blacks, greater segregation should lead to lower robbery rates, as robbers would expect to meet more resistance to robbery attempts when a city is more segregated. Therefore, this model suggests that the simple correlation between robbery and segregation could be either positive or negative. However, according to the model, *exogenous* sources of segregation should lead to lower rates of robbery. As will be shown below, this is the opposite prediction from that generated in the model developed here, suggesting that finding exogenous sources of segregation will be a key aspect to any attempts to empirically distinguish between O'Flaherty and Sethi's model and the one developed below.

The model developed below arguably most closely relates to the work describing how violence may play a strategic role. In particular, the sociological work of Anderson (1999) and Massey (1995) discusses how individuals adapt to high poverty isolated neighborhoods by becoming obsessively concerned with "respect" in order to lower the risk of their own criminal victimization, where such respect is maintained primarily through strategic use of force. Relatedly, Jankowski (1991) argues that one of the central motivations for joining a gang is often self-protection, even if by joining a gang an individual commits to perpetrating violent acts against others. Silverman (2004) and O'Flaherty and Sethi (2010, this issue) develop explicit models of such strategic violence, where individuals use violence as way of preempting or deterring violence being done to them. While these papers consider the strategic role of violence, they generally do not make explicit why such behaviors are connected to an individual's own income and the income distribution in the individual's neighborhood, or how the overall amount of violence may be affected by the degree to which the poor are residentially segregated from the non-poor in the overall community. The model below attempts to develop these connections more formally.

3. Model of crime, poverty, and neighborhood composition

This section develops two separate, but related models of criminal participation—one for participation in basic property crimes, the other for participation in interpersonal violent crimes. The models are separate in that an individual's decision whether or not to participate in one type of crime is not directly related to his decision about whether or not to participate in the other type of crime. The two models are related in that each considers a community made up of a large number of individuals who live for an infinite number of periods, where each individual can be classified as having either low income or high income, where income is exogenous to the individual. In the absence of committing any crime, assume low-income individuals have ω_ℓ dollars available for consumption each period, while high-income individuals have ω_h dollars available for consumption each period, and individuals cannot save or borrow across periods. Both models also assume individuals value consumption in any given period according the a utility function u, where u is an increasing strictly concave function in consumption, meaning individuals incur diminishing marginal utility in money. Finally, in each model, suppose the overall community can be divided up into a collection of neighborhoods, where each individual lives in one and only one neighborhood. Let λ_k denote the fraction of residents in a given neighborhood k who have low-income (i.e. are poor), and let λ denote the community-wide fraction of residents who have low-income.

3.1. Participation in basic property crimes

Let us first consider an individual's decision to become a thief, or to engage in a property crime that does not involve a direct confrontation with other individuals (e.g., burglary, larceny, motor vehicle theft). By becoming a thief, an individual i adds b units of additional consumption above and beyond the consumption possible though consuming his legal income that period, but also incurs a utility cost of ϵ_p^i. In words, ϵ_p^i represents each individual i's disutility (or possibly his utility if $\epsilon_p^i < 0$) associated with being a thief and committing property crimes, such guilt or pleasure, as well as the expected disutility associated with the possibility of receiving a jail sentence. This parameter will be referred to as each individual's "criminal propensity for property crimes," with a lower ϵ_p^i indicating a higher criminal propensity. In order to focus only on the role income, assume that ϵ_p^i is an i.i.d. random draw from a normal distribution with mean μ_p and variance σ_p^2, and therefore is independent across individuals. Note that this means that this model explicitly considers an environment where individual's preferences and/or judicial costs for committing or not committing crime are not directly influenced by their neighbors' preferences or neighborhood characteristics, which has been the focus of much of the previous literature on neighborhood effects and crime (e.g. Glaeser et al., 1996; Brock and Durlauf, 2001; Heavner and Lochner, 2002; Calvo-Armengol and Zenou, 2004; Calvo-Armengol et al., 2007; Patacchini and Zenou, 2008; Ferrer, 2010; Helsley and Strange, 2005; Wilson, 1987; Krivo and Peterson, 1996).

Given the discussion from above, an individual chooses to become a thief if and only if $u(\omega^i + b) - \epsilon_p^i \geqslant u(\omega^i)$, meaning the equilibrium fraction of individuals of income level ω_j (for $j \in \{\ell, h\}$) living in neighborhood k who choose to become thieves in any given period equals

$$\pi_j^* = \Phi_p\big(u(\omega_j + b) - u(\omega_j)\big), \tag{1}$$

where Φ_p is the cumulative distribution for the normally distributed random variable ϵ_p^i.

Because of the strict concavity of the u function, it is straightforward to see that $\pi_h^* < \pi_\ell^*$. In words, because the utility associated with any fixed monetary payoff from stealing is lower for high-income individuals, high-income individuals are less likely to become thieves. Hence, the greater the overall fraction of a neighborhood who are of low-income, the greater the fraction of the neighborhood who become thieves. This argument also holds at the community-wide level. Therefore, in this simple model, the rate of basic property crimes committed in a neighborhood (or a whole community), should be increasing in the fraction of neighborhood (community) made up of poor individuals. A second thing to notice about π_j^* as given in Eq. (1) is that it does not depend on λ_k. In words, the likelihood that an individual becomes a thief does not depend on the income of his neighbors, meaning there are no "neighborhood effects" with respect to basic property crimes without making further assumptions. Therefore, after controlling for the overall fraction of the community made up of poor individuals, the level of economic segregation in the community as a whole should have no direct effect on the overall rate of basic property crimes.

One reasonable extension is to assume that the monetary benefit to being a thief is greater when fewer of one's neighbors are poor, or if the monetary benefit to being a thief is given by $b(\lambda_k)$, then $b'(\lambda_k) < 0$. In this case, Eq. (1) would become

$$\pi_j^*(\lambda_k) = \Phi_p\big(u(\omega_j + b(\lambda_k)) - u(\omega_j)\big).$$

Since $b(\lambda_k)$ is decreasing in λ_k, the above equation implies that the fraction of individuals of any given income level j who choose to become thieves is decreasing in λ_k. Therefore, when the monetary benefit to being a thief depends on the economic status of one's neighbors, there will exist neighborhood effects with respect to becoming a thief. Moreover, note that the change in expected criminality from moving an individual of income level ω_j from a neighborhood k to a richer neighborhood k' (meaning $\lambda_k < \lambda_{k'}$ and $b(\lambda_k) > b(\lambda_{k'})$), will equal

$$\Delta\pi_j^* = \Phi_p(u(\omega_j + b(\lambda_k)) - u(\omega_j)) - \Phi_p(u(\omega_j + b(\lambda_{k'})) - u(\omega_j)).$$

Further note that the concavity of the u function implies $[u(\omega_j + b(\lambda_k)) - u(\omega_j)] - [u(\omega_j + b(\lambda_{k'})) - u(\omega_j)]$ will be larger when $\omega_j = \omega_\ell$ than when $\omega_j = \omega_h$. Therefore, a sufficient condition for $\Delta\pi_\ell^* > \Delta\pi_h^*$ is for Φ to be weakly convex when evaluated at or before $u(\omega_\ell + b(\lambda_k)) - u(\omega_\ell)$. Given Φ_p is the cdf of a normal distribution, this would be true for example if $\pi_\ell^*(0) \leqslant 0.5$, or if less than half of the poor individuals would choose to become thieves even if they were the only poor person in their neighborhood.

Recalling that $\Delta\pi_j^*$ denotes the expected change in criminality with respect to basic property crimes from moving an individual of income level ω_j from a richer to a poorer neighborhood, we can infer that an important implication of $\Delta\pi_\ell^*$ being greater than $\Delta\pi_h^*$ is that there will be bigger increase in expected criminality when moving a poor individual from a poorer neighborhood to a richer neighborhood than would be offset by the decrease in expected criminality from moving a rich individual from the richer neighborhood to the poorer neighborhood. This in turn implies that when the monetary benefit to committing a given basic property crime is inversely related to the fraction of the neighborhood that is poor, *less* segregation will actually *increase* this type of basic property crime and vice versa.

In summary, the simple model laid out in this section shows that in the absence of assuming preferences, information regarding payoffs to crime, or policing depend on the behavior or characteristics of one's neighbors, *greater segregation by income will either have no effect, or a negative effect on community-wide basic property crimes*, depending on whether the monetary benefit of basic property crime becomes greater the neighborhood poverty rate decreases.

3.2. Participation in interpersonal violent crime

Now consider crimes against persons, such as muggings, robberies, and assaults. In modeling these crimes, assume each individual decides whether to be a "thug" or a "law-abider," then proceeds to encounter other individuals in his neighborhood at a rate of one person per period. By choosing to be a law-abider, an individual commits to acting passively when encountering anyone in his neighborhood. Alternatively, by choosing to be a thug, an individual commits to violently attacking any law-abider he encounters in his neighborhood and having a violent interaction with another thug with probability $p \in [0,1]$.

Therefore, when a law-abider and a thug encounter each other, the one-sided violence will allow the thug to successfully rob the law-abider, thereby increasing the thug's consumption in that period by b, while decreasing the law-abider's consumption that period by b and further imposing a cost of c on the law-abider due to pain and suffering. We can think of these interactions as "robberies." Note that I assume that b does not depend on the income of one's victim. While the model is robust to loosening this assumption a little bit, I feel that such an assumption is generally justified. After all, it is not clear that poor individuals carry less cash on them than do rich individuals, especially since poor individuals are less likely to store their wealth in bank accounts or credit cards. A similar assumption and justification is made by O'Flaherty and Sethi (2007).

On the other hand, when two thugs encounter each other violence arises with probability p, and when it does, both individuals will still incur a cost of c due to pain and suffering *but no money will change hands*. We can think of these interactions as "aggravated assaults." Finally, when two law-abiders encounter each other, no violence takes place, meaning no money changes hands and no pain and suffering arises.

The above assumptions can be motivated two ways. First, choosing to be a thug can be interpreted as an individual learning the fighting skills and/or obtaining the weapons necessary to take possessions from law-abiders, who do not have such skills and/or weapons. However, since other thugs also have fighting skills and/or weapons, thugs cannot take possessions from each other, but will incur substantial pain and suffering when they fight. A second, complementary interpretation is that choosing to be a thug is equivalent to joining a street gang, where gang members take property from the non-gang members they encounter in their neighborhood, while at the same time must periodically engage in violence when encountering rival gang members, but do not lose their own property in such altercations. Such motivation is consistent with some of the ethnographic literature on gangs. For example, in summarizing the work of Savitz et al. (1980), Spergel (1990) states "(j)oining a gang may also result from rational calculations to achieve personal security, particularly for males, in certain neighborhoods."

Finally, analogous to the basic property crime model, by choosing to be a thug an individual i must further incur a utility cost ϵ_v^i each period, where once again ϵ_v^i is drawn from a normal distribution with mean μ_v and variance σ_v^2 and is independent across individuals (but is fixed for a given individual across periods). As before, this criminal propensity parameter ϵ_v^i captures the effort and any feelings of guilt (or pleasure) associated with being a thug and engaging in violence, as well as the expected disutility of being arrested and punished for being a thug.

The above assumptions mean that the expected utility for any given period for an individual i of income level ω_j living in neighborhood k associated with becoming a thug is given by

$$\pi[u(\omega_j) - pc] + (1-\pi)u(\omega_j + b) - \epsilon_v^i,$$

where π is the the likelihood he encounters a thug as opposed to a law-abider in his neighborhood (which arises endogenously as shown below). Alternatively, the expected utility from being a law-abider for any given period for an individual i of income level ω_j living in neighborhood k is given by

$$\pi[u(\omega_j - b) - c] + (1-\pi)u(\omega_j).$$

Given the above expected utilities, we can derive that optimal behavior for an individual i of income level j living in neighborhood k is to become a thug if and only if

$$\pi[u(\omega_j) - u(\omega_j - b) + (1-p)c] + (1-\pi)[u(\omega_j + b) - u(\omega_j)] \geqslant \epsilon_v^i. \tag{2}$$

Analogous to basic property crimes, the above expression indicates that it will generally be those with a low ϵ_v^i, meaning those with high criminal propensities for violence, who will choose to become thugs.

In order to further simplify Eq. (2), define $\delta_t(\omega_j)$ to equal $u(\omega_j) - u(\omega_j - b)$. In words, $\delta_t(\omega_j) + (1-p)c$ is the opportunity cost incurred by not being a thug when encountering a thug, for an individual with income ω_j. Similarly, define $\delta_a(w_j)$ to equal $u(\omega_j + b) - u(\omega_j)$, meaning $\delta_a(\omega_j)$ is the opportunity cost incurred by not being a thug when encountering a law-abider, for an individual with income ω_j.

Given these definitions, Eq. (2) becomes

$$\pi(\delta_t(\omega_j) + (1-p)c) + [1-\pi]\delta_a(\omega_j) \geqslant \epsilon_v^i. \tag{3}$$

This equation highlights the important components with respect to the decision individuals make regarding whether or not to become a thug in this environment. Namely, the fraction of individuals in a neighborhood choosing to become thugs is increasing in both the monetary benefit that can be obtained from doing so (i.e., $\delta_a(\omega_j)$), as well as the monetary and pain and suffering cost that can be avoided by doing so (i.e., $\delta_t(\omega_j) + (1-p)c$). This latter benefit to being a thug is one thing that makes the decision to become a thug different from the decision to become a thief. Moreover, also unlike the decision regarding whether or not to become a thief, the decision to become a thug depends on the overall fraction of other individuals in the neighborhood who are thugs (i.e. π).

From Eq. (3), we can now derive this fraction of individuals of income level ω_j living in neighborhood k choosing to be a thug to be

$$\pi_j = \Phi_v(\pi(\delta_t(\omega_j) + (1-p)c) + [1-\pi]\delta_a(\omega_j)), \tag{4}$$

where Φ_v again denotes the cdf of a normal distribution. For simplicity, I will refer to the fraction of individuals of a given group who choose to be a thug as the violent criminal participation rate for this group.

A Nash Equilibrium in this environment will be when all individual criminal participation decisions are optimal given everyone else's criminal behavior. Intuitively, a Nash Equilibrium will be an overall neighborhood violent crime participation rate π such that when each individual makes his criminal participation decision that maximizes his expected utility given this π, the resulting overall fraction of individuals in neighborhood k who choose to be thugs equals π. This leads to Proposition 1.

Proposition 1. *Given there is sufficient variation in violent criminal propensity over the population (namely $\sigma_v > \frac{\delta_t(\omega_\ell)+(1-p)c-\delta_a(\omega_\ell)}{\sqrt{2\pi}}$), then for any $\lambda_k \in [0,1]$, there exists a unique Nash Equilibrium characterized by violent criminal participation rates for each income group $\{\pi_\ell^*(\lambda_k), \pi_h^*(\lambda_k)\}$, and an overall violent criminal participation rate $\pi^*(\lambda_k) = \lambda_k\pi_\ell^*(\lambda_k) + (1-\lambda_k)\pi_h^*(\lambda_k)$, such that the following two equations hold:*

$$\pi_\ell^*(\lambda_k) = \Phi_v(\pi^*(\lambda_k)(\delta_t(\omega_\ell) + (1-p)c) + [1-\pi^*(\lambda_k)]\delta_a(\omega_\ell)),$$
$$\pi_h^*(\lambda_k) = \Phi_v(\pi^*(\lambda_k)(\delta_t(\omega_h) + (1-p)c) + [1-\pi^*(\lambda_k)]\delta_a(\omega_h)).$$

Proof. In Appendix A. □

Note that the assumption regarding the variance in criminal propensity (i.e. $\sigma_v > \frac{\delta_t(\omega_\ell)-\delta_a(\omega_\ell)}{\sqrt{2\pi}}$) is essentially just a regularity condition that assumes the distribution of criminal propensity is relatively diffuse, thereby ensuring that the externality of one individual's behavior on the optimal behavior of others is relatively small.

Given the existence and uniqueness of an equilibrium, a first thing to note is that the strict concavity of the u function implies that $(\delta_t(\omega_\ell) + (1-p)c) > (\delta_t(\omega_h) + (1-p)c)$ (or equivalently $\delta_t(\omega_\ell) >$

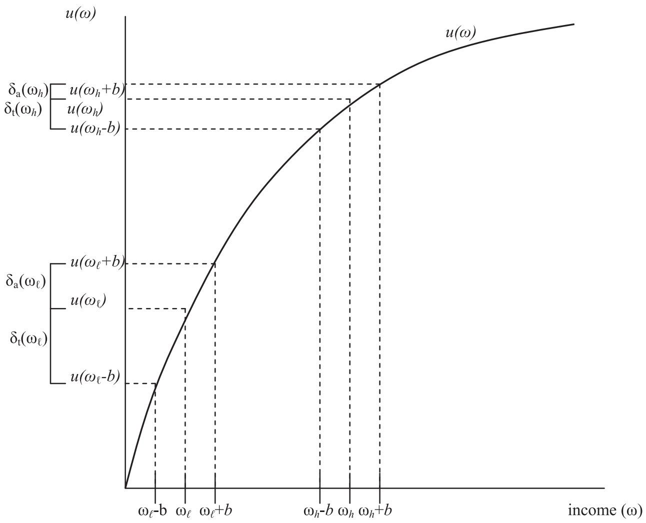

Fig. 1. Graphical depiction showing that $\delta_t(\omega_l) > \delta_p(\omega_l) > \delta_t(\omega_h) > \delta_p(\omega_h)$ will be true with strictly concave u function.

$(\delta_t(\omega_h))$, and $\delta_a(\omega_\ell) > \delta_a(\omega_h)$.[1] In words, low-income individuals will have a greater incentive to become thugs than will higher-income individuals all else equal. The intuition is similar to that with respect to basic property crimes. Since each individual's utility function exhibits diminishing marginal utility in consumption (and therefore income), the greater the individual's legal income each period, the smaller is the utility lost from getting a relatively small amount of money taken from them in any given period, and the smaller is the utility gained by taking a relatively small amount of money from someone else. This intuition can be seen in Fig. 1.

These differing incentives across income types leads to Proposition 2.

Proposition 2. *In any neighborhood k, a greater fraction of low-income individuals will participate in interpersonal violent crimes (i.e. be thugs) than high-income individuals, or* $\pi^*_\ell(\lambda_k) > \pi^*_h(\lambda_k)$.

Proof. In Appendix A. □

At this point it is worth noting that the model suggests that there is likely be a strong correlation between those who choose to become thieves and those who choose to become thugs for two reasons. First, it is certainly reasonable to think that there is a strong correlation between ϵ^i_p and ϵ^i_v within individuals. In other words, those who incur relatively low disutility from committing burglaries also likely incur relatively low disutility from committing assaults and robberies. However, even in the absence of a correlation between the preference parameters ϵ^i_p and ϵ^i_v, the model still suggests that there will be a large overlap between those who commit property crimes and those who commit violent crimes due to the fact that both types of crime will be correlated with an individual's economic circumstances.

The next thing to examine is how the likelihood of becoming a thug depends on the poverty rate of one's neighbors (i.e. λ_k), which leads to Proposition 3.

Proposition 3. *Given there is sufficient variation in violent criminal propensity over the population (namely* $\sigma_v > \frac{\delta_t(\omega_\ell)+(1-p)c-\delta_a(\omega_\ell)}{\sqrt{2\pi}}$*), then for both high and low-income individuals, the fraction choosing to participate in interpersonal violent crimes is increasing in the fraction of their neighborhood that has low income, or* $\frac{\partial \pi^*_j(\lambda_k)}{\partial \lambda_k} > 0$ *for* $j = h, \ell$.

Proof. In Appendix A. □

Proposition 3 shows that, in this model, an individual with income level j is more likely to become a thug if he lives in a relatively poor neighborhood than in a relatively rich neighborhood, meaning there exist *neighborhood effects* with respect to violent crime. Intuitively, when an individual expects a relatively high fraction of his neighbors to be thugs (as he would in a high-poverty neighborhood), his own incentive to become a thug is primarily defensive, in the sense of being able to prevent other thugs from taking his property. Alternatively, when an individual expects very few of his neighbors to be thugs (as he would in a low poverty neighborhood), his own incentive to become a thug is primarily to offensive, in the sense of being able to successfully take property from others. Due to the diminishing marginal utility of consumption, we know $\delta_t(\omega_j) + (1-p)c > \delta_a(\omega_j)$ for $j = h, \ell$ and $p \in [0, 1]$, meaning the defensive incentive for becoming a thug in a poor neighborhood is greater than the offensive incentive in a richer neighborhood.

Pushing the model a little bit further, we can examine whether the strength of this neighborhood effect (i.e. $\frac{\partial \pi^*_j(\lambda_k)}{\partial \lambda_k}$) differs by the income level of the individual. This leads to Proposition 4.

Proposition 4. *If* $\pi^*_\ell(1) \leqslant 0.5$*, the neighborhood effect will be stronger for low-income individuals than high-income individuals, or* $\frac{\partial \pi^*_\ell(\lambda_k)}{\partial \lambda_k} > \frac{\partial \pi^*_h(\lambda_k)}{\partial \lambda_k}$.

Proof. In Appendix A. □

Intuitively, from Fig. 1 above, we know that regardless of income level, the opportunity cost associated with not being a thug when encountering a law-abider is smaller than the opportunity cost of not being a thug when encountering another thug. This difference in opportunity costs accounts for the greater likelihood of becoming a thug in low poverty neighborhoods (where the likelihood of encountering a thug is high) than in higher poverty neighborhoods (where the likelihood of encountering a thug is low). Moreover, the strict concavity of the u function implies that this difference in opportunity costs is greater for low-income individuals than high-income individuals (as can also be seen in Fig. 1). Thus, low-income individuals will be more influenced by the income characteristics of their neighbors than will higher-income individuals when it comes to committing violent crimes.[2]

We can now analyze how the income distribution within a neighborhood, as well as how income is distributed across neighborhoods within the overall community, affect the rate of interpersonal violent crime. To start this analysis, first recall that the equilibrium fraction of individuals in any particular neighborhood k choosing to become thugs is given by $\pi^*(\lambda_k) = \lambda_k \pi^*_\ell(\lambda_k) + (1-\lambda_k)\pi^*_h(\lambda_k)$. Taking the derivative of this equation and re-arranging gives

$$\frac{\partial \pi^*(\lambda_k)}{\partial \lambda_k} = \left(\pi^*_\ell(\lambda_k) - \pi^*_h(\lambda_k)\right) + \lambda_k \frac{\partial \pi^*_\ell(\lambda_k)}{\partial \lambda_k} + (1-\lambda_k)\frac{\partial \pi^*_h(\lambda_k)}{\partial \lambda_k}.$$

From Proposition 2 we know that the first expression in parentheses in the above equation is positive, and from Proposition 3 we know that the second and third expressions in the above expression are also positive. Therefore, if we assume the overall rate of interpersonal violent crime within a neighborhood at any given point time is a strictly increasing function of the fraction of the

[1] Technically, this result is only guaranteed when $\omega_h - b > \omega_\ell + b$. In other words, when the changes in wealth associated with mugging or being mugged are small compared to the overall income differences between high-income and low-income individuals.

[2] The sufficient condition for this result, namely $\pi^*_\ell(1) \leqslant 0.5$, essentially says that this result will always hold if a relatively large fraction of poor individuals incur sufficient disutility from choosing to engage in the thug life such that they will still choose not to become thugs even if all of their neighbors are poor (i.e. a large fraction of individuals have a low criminal propensity or high ϵ^i_v). It is worth noting that this is a sufficient condition for Proposition 4 to hold, but is not necessary. If we drop the assumption that ϵ^i is normally distributed, a sufficient condition for Proposition 4 is that the cumulative distribution of ϵ^i is simply weakly convex prior to $\pi^*_\ell(1)$.

residents in the neighborhood who are thugs at that time, then increasing the fraction of the neighborhood made up of low-income individuals will increase the overall rate of interpersonal violent crime in the neighborhood.

Furthermore, note that if we take the second derivative of $\pi^*(\lambda_k)$ and re-arrange, we obtain

$$\frac{\partial^2 \pi^*(\lambda_k)}{\partial \lambda_k^2} = 2\left(\frac{\partial \pi_\ell^*(\lambda_k)}{\partial \lambda_k} - \frac{\partial \pi_h^*(\lambda_k)}{\partial \lambda_k}\right) + \lambda_k \frac{\partial^2 \pi_\ell^*(\lambda_k)}{\partial \lambda_k^2} + (1-\lambda_k)\frac{\partial^2 \pi_h^*(\lambda_k)}{\partial \lambda_k^2}.$$

From Proposition 4, we know the expression in parentheses in the above equation is positive. Moreover, as long as we again assume $\pi_\ell^*(1) \leqslant 0.5$ and $\sigma_v > \frac{\delta_t(\omega_\ell)-\delta_a(\omega_\ell)}{\sqrt{2\pi}}$, both of the latter two terms in the above expression are also positive (see Appendix A for formal proof), implying $\frac{\partial^2 \pi^*(\lambda_k)}{\partial \lambda_k^2} > 0$. This leads to Proposition 5.

Proposition 5. *The rate of interpersonal violent crime in the community as a whole is increasing in the degree to which its neighborhoods are segregated by income.*

Proof. Given $\frac{\partial \pi^*(\lambda_k)}{\partial \lambda_k} > 0$ and $\frac{\partial^2 \pi^*(\lambda_k)}{\partial \lambda_k^2} > 0$, we know the fraction of a neighborhood that chooses to become thugs is an increasing strictly convex function of the fraction of the neighborhood that is poor. Therefore, for any given community-wide fraction poor λ, the interpersonal crime rate in the overall community is minimized when all neighborhoods have the same fraction of the poor. Alternatively, the rate of interpersonal crime in the community as a whole becomes greater the more its neighborhoods are segregated by income. □

The intuition for Proposition 5 comes from Proposition 4, which showed that neighborhood effect more strongly influences poor individuals than rich individuals. Therefore, the rate of interpersonal violent crime should necessarily be higher the more poor individuals are segregated from richer individuals all else equal. Note that just the opposite was true with respect to the basic property crime model developed above.

In summary, this model reveals the important role of poverty at both the individual and neighborhood level can play when it comes to crime rates. The model shows how even with very few assumptions, the strategic role of violence can interact with neighborhood poverty rates to create neighborhood effects on individual decisionmaking. Notably, a key feature of this model is that it implies that neighborhood effects with respect to violent crime are likely to be more pronounced and different in nature than any that might arise with respect to basic property crime. This is in contrast to many of the other models of neighborhood effects highlighted at the beginning of Section 2, where the neighborhood effects that arise in those models would not necessarily differ across crime types. While this is not the first model suggesting a link between segregation and crime (for example the O'Flaherty and Sethi (2007) paper discussed above, as well as sociological work such as Blau and Blau (1982)), it is unique in its focus on primarily income driven forces, and its differential predictions of the effect of segregation on different types of crime.

4. Empirically evaluating the model

While the model presented above was in many ways quite simplistic, abstracting from many other potentially important determinants of crime rates in cities, there is still value in attempting some empirical analysis of the model. In particular, I will focus on the key community-level implications inherent in Section 3.2 and Proposition 5 that are direct implications coming from this model. Specifically, all else equal, greater economic segregation in a metropolitan area ("MSAs" from here on) should lead to higher rates of violent interpersonal crimes but have negligible or even a negative impact on basic property crime rates. However, it is important to note that the preceding models of basic property and violent crime considered segregation to be exogenous to the community. Clearly, while the degree to which a city is segregated by poverty status may have an impact on different types of crime for the reasons specified in the model, it may also be true that greater criminal activity affects the level of segregation by poverty status in overall community. Therefore, given the model treated the level of segregation in a community as exogenous, I will attempt to focus on plausibly exogenous variation in segregation across cities using instrumental variable methods as will be discussed in more detail below.

4.1. FBI Uniform Crime Reports (UCR)

The FBI Uniform Crime Reporting Program is a nationwide program covering roughly 94% of the total US population, and 96% of the population living in MSAs (Federal Bureau of Investigation, 2000).[3] In this analysis, I use the 2000 FBI UCR data to look separately at the five most common *Index* crimes—robbery, aggravated assault, burglary, larceny-theft, and motor vehicle theft. I do not focus on the two other Index crimes—rape and murder—both because the motivation for these crimes is likely often distinct from the forces inherent in the model developed above, as well as the fact that because the number of these crimes are relatively small, especially in smaller cities, making the rates somewhat uninformative. In particular, in many smaller cities there are less than five of such crimes reported in a given year, meaning for example that one more murder in a given year will increase the murder rate in that city by 20% or more. However, for the interested reader, Appendix A, Table A1 shows the basic results for murder rates. A major distinction between these crime categories is that aggravated assault and robbery are defined to be violent crimes involving a direct confrontation with the victim, while burglary, larceny, and motor vehicle thefts are defined as non-violent property crimes that explicitly do not involve a direct confrontation with the victim (Federal Bureau of Investigation, 2000).[4] Therefore, I will refer to aggravated assault and robbery as "violent interpersonal crimes," and burglary, larceny, and motor vehicle thefts as "basic property crimes."

I also use the FBI UCR data from 1999 to measure the number of officers per 1000 residents in each city, to account for differences in police presence across cities. Moreover, since the FBI UCR crime data are reported at the county level, I determined crime rates and officer rates for each MSA by aggregating all relevant data for counties that fall within a particular MSA. Because most counties either fall in one MSA or fall in zero MSAs, this generally provided accurate MSA crime information. However, several New England counties are divided between two or more distinct MSAs. Since I could not determine which MSA to assign the reported crimes in these counties to, I had to exclude these New England MSAs that contained shared counties from the analysis.[5] Finally, I dropped all MSAs with fewer than 150,000 residents since their crime rates, especially for violent crime, fluctuate substantially from year to year even though there are relatively modest changes in the number of crimes.

[3] This data was made available through the National Archive for Criminal Justice Data (NACJD) and the Inter-University Consortium for Political and Social Research (ICPSR) study #3451.

[4] Car-jacking, or taking an individual's car by threat or force, is counted as robbery, not a motor vehicle theft.

[5] This criteria excluded the following MSAs: Bangor ME, Boston MA–NH, Burlington VT, Hartford CT, Lewiston–Auburn ME, Manchester NH, Pittsfield MA, Portland ME, Portsmouth–Rochester NH–ME, Springfield MA. The Miami FL MSA, the Bloomington–Normal IL MSA, and the Champaign–Urbana IL MSA were also dropped from the sample because the FBI UCR reports did not provide crime data for these MSAs in 2000.

Table 1
Descriptive statistics of data.

Variable	Mean	Std. dev.
Crime		
Basic property crimes (in 2000)		
Burglaries per 100 K residents	836	323
Larcenies per 100 K residents	2903	944
Motor vehicle thefts per 100 K residents	388	210
Interpersonal crimes (in 2000)		
Robberies per 100 K residents	133	85
Aggravated assaults per 100 K residents	321	165
Law enforcement officers per 1000 residents	244	150
MSA characteristics		
Total population	893,239	1,298,606
Percent poor	12.3	4.5
Percent urban	82.0	11.1
Percent immigrant	1.2	0.9
Percent black	11.0	9.9
Percent hispanic	11.5	15.8
percent of adults with college degree	23.9	6.8
Percent of households headed by single mother	7.5	1.6
Percent of households receiving housing assistance	2.1	1.0
Unemployment rate	5.9	1.9
Percent of days over 100°	11.1	8.0
Percent of days below 32°	22.2	13.4
Segregation		
Isolation index of segregation by poverty status	0.20	0.06
Isolation index of segregation by poverty status (excluding tracts with over 60% of population in college)	0.19	0.05
Isolation index of segregation by race	0.33	0.07
Number of observations	228	

4.2. MSA population characteristics

Data regarding MSA population characteristics come for the most part from the 2000 United States Census Summary File 3. I use these data to obtain measures of a variety of demographic characteristics for each MSA (see Table 1 for particular variables). I also use data from the Department of Housing and Urban Development's "A Picture of Subsidized Households – 1998" to determine the fraction of households in each MSA that receive housing assistance. Finally, to control for the potential effects of weather on criminal activity (see Jacob et al., 2007), I determined the average number of very hot days (i.e. temperature of 90° or higher) per 100 days for each state, as well as the average number of very cold days (i.e. temperature of 32° or lower) per 100 days for each state using data from the National Climatic Data Center.

Table 1 summarizes all of the above variables for the sample used in this analysis.

4.3. The correlation between economic segregation and crime

While there exist several plausible ways to measure the level of income segregation within a city, I primarily employ the *isolation index*.[6] This index attempts to measure the extent to which individuals of one group are only exposed to one another, rather than members of the other group, in their neighborhoods (Massey and Denton, 1988). In the context of segregation by poverty status, this index is essentially computed to be the fraction poor in the census tract occupied by the average poor individual in that MSA, and is given by the following formula:

$$\text{Poverty Isolation Index} = \sum_{i=1}^{N} \frac{\text{poor}_i}{\text{poor}_{\text{total}}} \frac{\text{poor}_i}{\text{persons}_i},$$

where i denotes census tract. The higher this index is, the greater the level of segregation.

Looking at this segregation statistic when it is computed using all census tracts in each MSA reveals a potentially problematic issue, namely that among the twenty MSAs with the highest Poverty Isolation Index are College Station TX, Gainsville FL, Athens GA, Talahassee FL, Lafayette IN, Madison WI, Provo-Orem UT, and Las Cruces NM; all moderate to small MSAs containing large universities. The concern this raises is that full-time students who do not live in dormitories will generally be counted as poor, since they earn little or no income while in school. Moreover, such students tend to live almost exclusively in census tracts surrounding their University, causing MSAs with relatively high college student populations to appear relatively segregated by poverty status, but not in the way we generally are attempting to capture. Therefore, I also computed the Poverty Isolation Index for each MSA excluding those census tracts containing over 60% students. This will be the preferred measure of poverty segregation, however, as I will also show, results do not differ substantively by using Poverty Isolation Indices computed using all census tracts.

We can begin by looking at the relationship between crime and poverty segregation by using simple OLS specifications, regressing the MSA crime rate, for each type of crime separately, on the Poverty Isolation Index for the MSA and a variety of other MSA characteristics. Table 2 shows separate specifications for each type of crime, where the dependant variable is the rate of that crime per 100,000 residents, standardized to have a mean of zero and standard deviation of one. I use these standardized rates in order to facilitate comparing magnitudes across crimes, as the overall rates per 100,000 residents differ dramatically across crimes (as can be seen in Table 1).

Looking at the first row of Table 2, we can see that conditional on the MSA level characteristics, the correlation between segregation by poverty status and crime rates is relatively weak across all crime categories, but there is some evidence that segregation by poverty status is positively correlated with motor vehicle theft, robbery, and aggravated assault.[7]

While these OLS results reveal some small differences in the correlation between poverty segregation and crime across different types of crime, these results are not necessarily very informative about the degree to which such economic segregation actually *affects* MSA-wide crime rates for these different types of crimes. In particular, as alluded to previously, the level of segregation in an MSA may be endogenous since people generally have substantial choice about where to live within a city and crime rates might affect this decision.

Such selection may bias the causal interpretation of the OLS results for several reasons. For example, as alluded to previously, O'Flaherty and Sethi (2007) (as well as O'Flaherty and Sethi (2010c) and Cullen and Levitt (1999)) argue that rising crime rates may lead to flight from central cities, especially by the wealthy (and therefore disproportionately white). This means that any positive relationship between crime and segregation may arise not because greater segregation increases crime, but

[6] All segregation measures used in this paper were computed using the Census Summary File 3 data discussed previously.

[7] I also constructed Racial Isolation Indices for each MSA (using all census tracts). The coefficients on the Racial Isolation Index in specifications otherwise analogous to those in Table 2 are insignificantly different from zero at any standard level of significance in the burglary, larceny, and motor vehicle theft specifications. However, the coefficients on the Racial Isolation Index are positive and significant at the 1% level in the robbery and aggravated assault specifications. Like the coefficients in Table 2, these coefficients were also relatively small in magnitude, with the coefficients indicating that a one standard deviation in the Racial Isolation Index is correlated with a 0.25 and 0.31 standard deviation increase in robbery and aggravated assault rates respectively—findings consistent with Shihadeh and Flynn (1996). The estimated coefficients on the other variables are almost identical to those in Table 2. Also, Table A1 in Appendix A shows the results for murder rates.

Table 2
OLS regression results.

Variable	Dependant variable				
	Standardized burglary rate	Standardized larceny rate	Standardized motor vehicle theft rate	Standardized robbery rate	Standardized aggravated assault rate
Std. Poverty Isolation Index	0.078	0.075	0.240	0.346	0.213
(excluding tracts with over 60% college)	(0.125)	(0.137)	(0.123)*	(0.106)***	(0.129)
Officers per 1000 residents 1999	0.109	0.055	0.014	0.107	0.098
	(0.037)***	(0.040)	(0.036)	(0.031)***	(0.038)***
Percent in poverty	0.052	0.063	−0.027	−0.059	−0.005
	(0.039)	(0.043)	(0.038)	(0.033)*	(0.040)
Log of population	−0.131	−0.252	0.373	0.310	0.162
	(0.074)*	(0.081)***	(0.073)***	(0.063)***	(0.076)**
Percent urban	0.017	0.018	0.027	0.018	0.003
	(0.007)**	(0.008)**	(0.007)***	(0.006)***	(0.008)
Percent immigrant	−0.072	−0.053	−0.096	−0.023	−0.023
	(0.064)	(0.070)	(0.063)	(0.054)	(0.066)
Percent black	0.009	−0.014	−0.001	0.039	−0.010
	(0.012)	(0.013)	(0.011)	(0.010)***	(0.012)
Percent hispanic	−0.015	−0.013	−0.015	0.001	−0.004
	(0.007)***	(0.007)*	(0.006)**	(0.006)	(0.007)
Percent with college degree	−0.025	0.005	−0.011	−0.018	−0.016
	(0.010)***	(0.011)	(0.010)	(0.008)**	(0.010)
Percent of HH with single mother	0.088	0.213	0.171	−0.007	0.110
	(0.064)	(0.070)***	(0.063)***	(0.054)	(0.066)*
Percent of HH subsidized	−0.120	−0.010	−0.105	0.036	−0.078
	(0.061)*	(0.067)	(0.060)*	(0.052)	(0.063)
Percent unemployed	−0.029	−0.124	0.066	0.012	−0.011
	(0.054)	(0.059)**	(0.053)	(0.046)	(0.056)
Percent of days above 90°	0.100	0.114	0.066	0.070	0.118
	(0.032)***	(0.034)***	(0.031)**	(0.027)***	(0.032)***
Sq. of percent of days above 90	−0.002	−0.002	−0.001	−0.002	−0.003
	(0.001)**	(0.001)*	(0.001)	(0.001)***	(0.001)***
Percent of days below 32°	0.025	0.059	0.019	0.029	−0.049
	(0.017)	(0.019)***	(0.017)	(0.014)**	(0.017)***
Sq. of percent of days below 32	−0.000	−0.001	−0.000	−0.000	0.001
	(0.000)	(0.000)**	(0.000)	(0.000)*	(0.000)**
N	228	228	228	228	228
R-square	0.40	0.38	0.51	0.63	0.38

Standard errors in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

because greater crime leads to greater economic and racial segregation. Therefore, the OLS results presented previously may be upwardly biased.

Alternatively, as violent crime increases in a city, for example as gangs become more prominent, individuals living in the neighborhoods where these gangs operate have a greater incentive to take on the expenses associated with moving. Indeed, escaping from gangs and crime was the primary reason participants in the MTO housing relocation program gave for signing up for the program (Kling et al., 2005). Given that these neighborhoods where violence and gang activity are greatest are often the poorest neighborhoods in a city, those emigrating from these neighborhoods will generally be poorer than the residents of the neighborhoods they move to. Therefore, it is also possible that as crime increases, a city becomes somewhat less economically segregated than it would be otherwise, meaning the OLS results discussed previously could also be downwardly biased.[8]

4.4. Controlling for the potential endogeneity of segregation

To overcome the potential simultaneity bias we must find some characteristics that vary across Metropolitan areas that affect current income segregation, but can be credibly excluded from having any direct relationship to current levels of criminal activity. Given the existence of such variables, we can then use them as instruments in Two-stage Least Squares (2SLS) approach.[9]

The first instrument for segregation by poverty status that I employ is the fraction of public housing assistance that was allocated in the form of apartments in government owned public housing structures as opposed to allocated via Section 8 housing vouchers or certificates (or other types of subsidies to non-government property owners). The data used to create this instrument once again comes from the HUD's "A Picture of Subsidized Households – 1998" described above. By design, public housing structures group poor individuals together to a greater extent than do housing vouchers which can generally be used anywhere in the city. Indeed, the HUD data shows that the census tracts surrounding public housing structures are almost 40% poor on average, compared with an average of around 20% poor for census tracts surrounding those units procured via vouchers or certificates.

[8] For more formal and detailed discussions of racial and economic segregation that are not related to crime, see Sethi and Somanathan (2004) and Bayer et al. (2004).

[9] Optimally, one might want to look at the relationship between *changes* in economic segregation over time and *changes* in crime rates. However, such a method would not allevaite the basic endogeneity concern on its own, and therefore would require time-varying instruments for economic segregation. As will be seen below, the instruments used here are not time varying.

Table 3
First stage of 2SLS regression results.

Control variables	Dependant variable	
	Std. Poverty Isolation Index	Std. Poverty Isolation Index (excluding census tracts with over 60% in college)
Instruments meeting exclusion restriction		
Percent of housing assistance via public housing	0.003 (0.002)**	0.004 (0.001)***
Percent of local rev. coming from state and federal government	−0.009 (0.003)***	−0.010 (0.003)***
Non-excluded		
Officers per 1000 residents 1999	−0.012 (0.000)	0.009 (0.000)
Percent in poverty	0.245 (0.016)***	0.213 (0.014)***
Log of population	0.037 (0.042)	0.122 (0.038)***
Percent urban	0.012 (0.004)***	0.010 (0.004)**
Percent immigrant	−0.077 (0.037)**	−0.077 (0.034)**
Percent black	0.023 (0.007)***	0.021 (0.006)***
Percent hispanic	−0.006 (0.004)	−0.000 (0.003)
Percent with college degree	0.027 (0.006)***	0.001 (0.005)
Percent of HH with single mother	−0.008 (0.037)	0.060 (0.034)*
Percent of households subsidized	−0.113 (0.035)***	−0.068 (0.032)**
Unemployment rate	−0.043 (0.031)	−0.031 (0.028)
Percent of days above 90°	−0.050 (0.018)***	−0.038 (0.016)**
Sq. of percent of days above 90	0.001 (0.001)**	0.001 (0.001)**
Percent of days below 32°	0.012 (0.010)	0.011 (0.009)
Sq. of percent of days below 32	−0.000 (0.000)	−0.000 (0.000)
Constant	−4.304 (0.605)***	−5.073 (0.552)***
N	228	228
R-square	0.82	0.86
F-statistic for excluded instruments	6.72**	10.69***

Standard errors in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Moreover, the stock of public housing in a particular city has generally existed for a considerable number of years prior to the year 1998 (the year in which the measures come from for this analysis). Indeed, data from "A Picture of Subsidized Housing in the 1970s" (also made available by HUD) confirms that the number of in-kind public housing units used to provide housing assistance throughout US cities in 1998 was essentially determined several decades ago. Specifically, over 87% of the public housing projects that existed in 1977 still existed and were in use in 1998. Moreover, very few public housing projects were built between the 1970s and 1998, with 62% of all public housing projects that existed and were in use in 1998 being constructed prior to 1977, and over 92% of those projects larger than 200 units being constructed prior to 1977. This is important because, as argued by Brueckner and Rosenthal (2009), a key force of economic integration of cities is the richer population following the cyclical updating of the housing stock throughout different neighborhoods in a metropolitan area. However, the presence of large public housing developments in a neighborhood impedes updating of the housing stock in that neighborhood, since the projects are not subject to standard market interventions, which also likely impact the surrounding housing structures as well. Therefore, cities that provided a relatively large fraction of their public housing through large public housing structures in the 1970s and 1980s likely faced greater barriers to the economic integration forces highlighted by Brueckner and Rosenthal (2009) than other cities.

The second variable I use as an instrument for segregation by poverty status was first used by Cutler and Glaeser (1997) as an instrument for racial segregation—namely the share of local government revenue in an MSA that comes from the state or federal government in 1962.[10] With more money coming from outside sources, there is less of an incentive for individuals within a city to segregate by income, since a smaller fraction of local public goods are funded through local taxes. Therefore, a greater fraction of local revenue coming from the state or federal government should lead to less economic segregation in an MSA.[11]

The first column of numbers in Table 3 shows the results of the first stage regression of the Poverty Isolation Index calculated using all census tracts on the two instruments meeting the exclusion restriction and the other MSA characteristics included in the original regressions from Table 2. The second column of numbers in Table 3 shows the analogous results that arise when using the Poverty Isolation Index calculated using only census tracts made up of less than 60% students. As can be seen, the two instruments discussed above are significantly related to Poverty Isolation Index (using either calculation method) in the predicted manner. However, as should be expected, both the magnitude of the estimated coefficients on the excluded instruments, as well as the *F*-statistic for the joint significance of the two instruments, are larger when using the Poverty Isolation Index calculated using only census tracts made up of less than 60% students.[12] Therefore, I will again focus on the results using this construction of the Poverty Isolation Index.

Table 4 shows the results from the 2SLS specifications. The first column in Table 4 reveals that, if anything, greater segregation by poverty status actually decreases rates of burglary. While not statistically significant at standard levels of significance (*p*-value 0.117), the coefficient is relatively large in magnitude, suggesting that a one standard deviation increase in the Poverty Isolation Index leads to a decrease in burglary rates by roughly two-thirds of a standard deviation.[13] Alternatively, the results shown in Table 4 with respect to larceny and motor vehicle theft reveal little effect of segregation by poverty status on these crimes.

Finally, the most notable results are shown in the last two columns of Table 4, which suggest that greater segregation by poverty status leads to much higher rates of the violent interpersonal crimes of robbery and aggravated assault, with these effects being statistically significant at the 10% level or higher. The point estimates indicate that a one standard deviation increase in the Poverty Isolation Index is associated with an increase in robbery rates by roughly two-thirds of a standard deviation and almost

[10] This data comes from the Census of Governments 1962, made available by the Inter-University Consortium for Political and Social Research (ICPSR) website.

[11] Note that while Cutler and Glaeser (1997) motivate this instrument identically to here, they use theirs to instrument for racial segregation, under the further motivation that income and race are strongly correlated in the US

[12] Indeed, the *F*-statistic on the joint significance of the instruments when using the Poverty Segregation Index calculated without the student heavy census tracts is arguably large enough to mitigate any concerns regarding weak instrument bias (Stock and Yogo, 2002).

[13] This translates to an almost 50% lower burglary rate (computed using the mean and standard deviation for burglaries from Table 1).

Table 4
2SLS regression results.

Variable	Dependant variable				
	Standardized burglary rate	Standardized larceny rate	Standardized motor vehicle theft rate	Standardized robbery rate	Standardized aggravated assault rate
Std. Poverty Isolation Index (excluding tracts with over 60% in college)	−0.672	−0.053	0.002	0.643	0.886
	(0.429)	(0.434)	(0.394)	(0.343)*	(0.433)**
Officers per 1000 residents in 1999	0.132	0.059	0.022	0.100	0.077
	(0.040)***	(0.041)	(0.037)	(0.032)***	(0.041)*
Percent in poverty	0.220	0.092	0.026	−0.126	−0.156
	(0.100)**	(0.101)	(0.092)	(0.080)	(0.101)
Log of population	−0.029	−0.235	0.405	0.270	0.071
	(0.095)	(0.096)**	(0.087)***	(0.076)***	(0.096)
Percent urban	0.027	0.020	0.030	0.015	−0.006
	(0.009)***	(0.009)**	(0.009)***	(0.007)**	(0.009)
Percent immigrant	−0.109	−0.059	−0.108	−0.008	0.010
	(0.070)	(0.070)	(0.064)*	(0.056)	(0.070)
Percent black	0.026	−0.011	0.004	0.032	−0.025
	(0.015)*	(0.015)	(0.014)	(0.012)***	(0.015)*
Percent hispanic	−0.016	−0.013	−0.016	0.001	−0.003
	(0.007)**	(0.007)*	(0.006)**	(0.005)	(0.007)
Percent with college degree	−0.030	0.004	−0.012	−0.016	−0.012
	(0.010)***	(0.011)	(0.010)	(0.008)**	(0.011)
Percent of HH with single mother	0.114	0.217	0.179	−0.017	0.087
	(0.068)*	(0.069)***	(0.063)***	(0.055)	(0.069)
Percent of HH subsidized	−0.175	−0.020	−0.122	0.058	-0.029
	(0.070)**	(0.071)	(0.065)*	(0.056)	(0.071)
Unemployment rate	−0.060	−0.129	0.056	0.024	0.017
	(0.059)	(0.059)**	(0.054)	(0.047)	(0.059)
Percent of days above 90°	0.063	0.107	0.054	0.084	0.151
	(0.038)*	(0.039)***	(0.035)	(0.031)***	(0.039)***
Sq. of percent of days above 90	−0.001	−0.002	−0.001	−0.003	−0.004
	(0.001)	(0.001)	(0.001)	(0.001)***	(0.001)***
Percent of days below 32°	0.035	0.060	0.022	0.025	−0.058
	(0.018)*	(0.019)***	(0.017)	(0.015)*	(0.019)***
Sq. of percent of days below 32	−0.001	−0.001	−0.000	−0.000	0.001
	(0.000)*	(0.000)**	(0.000)	(0.000)*	(0.000)***
Constant	−4.961	−1.917	−9.484	−4.093	0.979
	(2.549)*	(2.576)	(2.338)***	(2.036)**	(2.573)
N	228	228	228	228	228
F-statistic on excluded instruments	10.69***	10.69***	10.69***	10.69***	10.69***
p-value on Sargan statistics (i.e. overidentification test	0.23	0.39	0.13	0.81	0.20

Standard errors in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

nine-tenths of a standard deviation increase in the rate of aggravated assault.[14]

Table 5 shows that the above 2SLS results are robust to other measures of segregation. In particular, the top panel of Table 5 reveals the coefficients on different poverty segregation measures in otherwise analogous 2SLS specifications. The top row of numbers in Table 5 simply repeats the coefficients on the Poverty Segregation Index shown in the top row of Table 4. The remaining rows show the analogous coefficients when using several alternative construction methods for measures of segregation. Specification 1 uses the (standardized) Poverty Isolation Index calculated using *all* census tracts (including those made up of over 60% college students). Specification 2 uses the (standardized) Poverty Dissimilarity Index, which answers the question "what share of the poor population would need to change census tracts for the poor and non-poor to be evenly distributed within a city?" and is constructed using the formula $\frac{1}{2}\sum_{i=1}^{N}\left|\frac{\text{poor}_i}{\text{poor}_{total}} - \frac{\text{non-poor}_i}{\text{non-poor}_{total}}\right|$ for each MSA (Massey and Denton, 1988). Specification 3 uses the (standardized) Poverty Dissimilarity Index but constructed without using those census tracts made up of over 60% college students. Specification 4 once again uses the (standardized) Isolation Index computed using all census tracts, but adjusts for the overall fraction of the MSA that is poor (see Cutler and Glaeser, 1997).[15] Specification 5 again uses this "adjusted" Poverty Isolation Index, but computes it excluding those census tracts made up of over 60% college students. One concern regarding all of these poverty segregation measures is

[14] Given the mean and standard deviation for robbery rates per 100 thousand residents are 133 and 85 respectively, the above estimates suggest that a one standard deviation increase in poverty segregation leads to a roughly 40% higher robbery rate, all else equal. Similarly, given the mean and standard deviation for aggravated assault rates are 321 and 165 respectively, the above estimates suggest that a one standard deviation increase in poverty segregation leads to roughly 45% higher aggravated assault rate, all else equal. Again, see Appendix A, Table A1 for the 2SLS results corresponding to murder rates.

[15] This index is constructed to be the following $\frac{\sum_{i=1}^{N}\left(\frac{\text{poor}_i}{\text{poor}_{\text{total}}}\frac{\text{poor}_i}{\text{persons}_i}\right)-\left(\frac{\text{poor}_{\text{total}}}{\text{persons}_{\text{total}}}\right)}{\min\left(\frac{\text{poor}_{\text{total}}}{\text{persons}_{\ell}},1\right)-\left(\frac{\text{poor}_{\text{total}}}{\text{persons}_{\text{total}}}\right)}$, where persons_{ℓ} is the number of persons in the census tract with the lowest population with in the city and i once again denotes census tract. The first term in the top part of the above equation is the fraction poor in the census tract occupied by the average poor individual. From this, we can subtract the percentage poor in the city as a whole to eliminate the effect coming from the overall size of the poor population. This whole term is then normalized to be between zero and one, with one indicating the city is the most segregated it can possibly be.

that they treat the poor as being distinct from everyone else including the near poor, which obviously is not true. Therefore, Specification 6 again uses the Poverty Isolation Index calculated using all census tracts, but excludes those individuals in each census tract whose household earnings are above the poverty line but less than one and a half times the poverty line (the "near" poor), and Specification 7 uses the Poverty Isolation Index that both excludes those census tracts made up of over 60% students and excludes those individuals who are "near" poor.

As the upper panel of Table 5 shows, the coefficients on the segregation index in the burglary specifications are negative using all seven alternative indices and significantly so in six of them with magnitudes both somewhat smaller and larger than the coefficients that arise using the preferred measure (i.e. top row). The coefficients on the Poverty Segregation Indices in the larceny and motor vehicle theft specifications are negative using all seven alternative measures, but never even close to being statistically different from zero. However, the coefficients on the alternative Poverty Segregation Indices in the robbery and aggravated assault specifications are always positive and significantly so in 12 of the 14 specifications, with magnitudes again both somewhat smaller and somewhat larger than those shown in the top row of Table 5.

Finally, the lower panel of Table 5 shows the 2SLS coefficients on three different indices of *racial* segregation. As can be seen, the results generally mimic the results of the economic segregation indices—greater racial segregation seems to have no effect on larceny and motor vehicle theft, a negative impact on burglary (in fact significantly so), and a positive and significant impact on robbery and aggravated assault. To the extent that racial segregation is another measure of economic segregation, this evidence is consistent with the model. Interestingly, even thought the instruments are argued to be affecting income segregation, the *F*-statistic on the instruments suggests that they are even more strongly correlated with racial segregation. This could either be because the instruments do indeed have a more direct impact on racial segregation than economic segregation, or because race is actually a better measure of true economic circumstances than income from one year as measured by the Census.

In their paper on the consequences of ghettos, Cutler and Glaeser (1997) also use two further instruments for (racial) segregation beyond the fraction of local government revenue coming from the state or federal governments—namely the number of local municipal governments in each MSA, and the number of rivers flowing through each MSA.[16] I also estimated the 2SLS specifications using these variables as additional instruments for poverty segregation. However, as the top three rows of Table 6 show, using these additional instruments leads to no substantive differences in the estimated results, but does dramatically lower the *F*-statistic on the joint significance of the excluded instruments in the first stage regression. The fact that these further instruments do not seem to add much power to the analysis is the reason why they were not included in the "preferred" specifications shown in Table 6. Moreover, the last row of Table 6 shows the coefficients on the Poverty Isolation Index variable in the five 2SLS specifications using only the Cutler and Glaeser instruments (i.e. the fraction of local revenue received from state or federal sources, the number of municipal governments in each MSA, and the number of rivers in each MSA). These coefficients show that not using the "fraction of public housing given in-kind" instrument does alter the coefficients a good deal (especially those in the larceny and motor vehicle theft specifications), but the general conclusion remains that greater economic segregation appears to have differential effects on violent crimes versus basic property crimes.

4.5. Discussion of empirical results

Clearly, the validity of the 2SLS results presented above rest on the validity of the proposed instruments. On the most basic level, Table 3 (and the last column of Table 5) showed that both of the primary instruments argued to meet the exclusion restriction are indeed significantly correlated with segregation in the predicted manner. Therefore, one criteria for the validity of the instruments seems to be met. Moreover, given we have more excluded instruments than potentially endogenous variables, the model is overidentified, which means we can directly test whether it is inappropriate to exclude the instruments discussed above from being related to crime in 2000 other than through segregation (Wooldridge, 2002). The instruments pass this test. In particular, the *p*-value on the Sargan overidentification test statistics for the specifications in Table 4 range from 0.13 (motor vehicle theft) to 0.81 (robbery).

Another potential test of the validity of these instruments is to see if they have a significant relationship to other key metro area characteristics, such as the poverty rate or the poverty rate for blacks, after controlling for segregation by poverty status, as well as the other metro area characteristics. If they do, this suggests that these proposed instrumental variables may directly influence a variety of characteristics of a city in addition to segregation, which may then have their own direct affects on crime.[17] However, running similar first stage regressions to those shown in Table 3, but using "percent living in poverty" as the dependant variable and adding the Poverty Isolation Index to the right-hand side variables, the coefficients on the two excluded instruments are small in magnitude and statistically insignificant at any standard level of significance. Similarly, when I regress "percent of blacks living in poverty" on the instruments, any of the segregation indices, and the remaining right-hand side variables, the coefficients on the two instruments are again small and statistically insignificant. *In other words, other than through their relationship to segregation, the two instruments do not appear to be related to poverty rates as a whole or poverty rates for blacks.* Moreover, when I do the analogous exercise with "percent of households headed by single mother" as the dependant variable I again get coefficients that are small in magnitude on both instruments, but while small in magnitude, the coefficient on the public housing instrument is significantly negative at the 5% level. In other words, if anything, after controlling for all of the other metro area characteristics, a higher fraction of public housing given in-kind leads to lower rates of single parent headed households. Therefore, the above tests are consistent (though admittedly not conclusive) with the notion that instruments are not simply picking up the effects of some omitted variable that affects both current crimes rates and the composition of public housing subsidies and/or the historical sources of local revenue.

Even given the above discussion however, it is certainly plausible that something about one of the key instruments, namely fraction of public housing given in kind, is directly related to criminal activity. For example, there is a large literature on the relationship between public housing and crime related to the notion of "defensible space" (Jacobs, 1961; Newman, 1972). Advocates for defensible space in public housing argue that criminality in large public housing projects is often facilitated by unmaintained common areas that residents must pass through but have very little control over, such as hallways and common courtyards, and residents in such facilities are often separated by both physical barriers and distance from the wider community at large. While the empirical

[16] There has been considerable debate on how this variable should be property measured (see Rothstein (2007) and Hoxby (2007)). Given it is not the focus of this analysis, I simply decided to use the number of "long" rivers flowing through each MSA as coded by Jesse Rothstein. Thanks to Jesse Rothstein for providing me with this data.

[17] Thanks to Francisco Martorell for suggesting this.

Table 5
2SLS regression coefficients on alternative segregation indices.

Measure of segregation	Dependant variable					
	Standardized burglary rate	Standardized larceny rate	Standardized motor vehicle theft rate	Standardized robbery rate	Standardized aggravated assault rate	*F*-statistic on the instruments
Std. Poverty Isolation Index (excluding tracts with over 60% college)	−0.672 (0.429)	−0.053 (0.434)	0.002 (0.394)	0.643 (0.343)*	0.886 (0.433)**	10.69***
Alternative Poverty Segregation Indices						
1 – Std. Poverty Isolation Index (using all census tracts)	−0.840 (0.503)*	−0.185 (0.502)	−0.082 (0.455)	0.746 (0.419)*	0.936 (0.527)*	6.72**
2 – Std. poverty dissimilarity index (using all census tracts)	−0.541 (0.302)*	−0.216 (0.302)	−0.122 (0.271)	0.433 (0.241)*	0.461 (0.297)	8.18***
3 – Std. poverty dissimilarity index (excluding tracts with over 60% college)	−0.509 (0.287)*	−0.193 (0.286)	−0.107 (0.258)	0.413 (0.225)*	0.449 (0.279)	9.11***
4 – Std. "adjusted" Poverty Isolation Index (using all census tracts)	−0.501 (0.295)*	−0.121 (0.296)	−0.057 (0.268)	0.439 (0.246)*	0.542 (0.309)*	7.00***
5 – Std. "adjusted" Poverty Isolation Index (excluding tracts with over 60% college)	−0.356 (0.216)*	−0.054 (0.220)	−0.017 (0.200)	0.328 (0.173)*	0.432 (0.217)**	12.20***
6 – Std. Poverty Isolation Index (using all census tracts but excluding the "near" poor)	−0.917 (0.545)*	−0.244 (0.537)	−0.119 (0.485)	0.794 (0.452)*	0.960 (0.564)*	5.99**
7 – Std. Poverty Isolation Index (excluding census tracts with over 60% college and the "near" poor)	−0.758 (0.475)	−0.105 (0.473)	−0.030 (0.429)	0.703 (0.378)*	0.933 (0.476)**	8.99***
Racial Segregation Indices						
8 – std. Racial Isolation Index (using all census tracts)	−0.628 (0.368)*	−0.147 (0.373)	−0.067 (0.337)	0.554 (0.290)*	0.687 (0.342)**	12.93***
9 – Std. "adjusted" Racial Isolation Index (using all census tracts)	−0.468 (0.273)*	−0.111 (0.277)	−0.051 (0.251)	0.412 (0.215)*	0.511 (0.257)**	12.52***
10 – Std. racial dissimilarity index (using all census tracts)	−0.380 (0.209)*	−0.108 (0.218)	−0.055 (0.198)	0.325 (0.174)*	0.387 (0.214)*	15.24***

Standard errors in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Table 6
2SLS regression coefficients on segregation index using additional instruments.

	Dependant variable					
	Standardized burglary rate	Standardized larceny rate	Standardized motor vehicle theft rate	Standardized robbery rate	Standardized aggravated assault rate	*F*-statistic on the instruments
Instruments: (i) fraction public housing given 'in-kind', (ii) fraction of local public revenue received from state or federal						
Std. Poverty Isolation Index Excluding tracts with over 60% college	−0.672 (0.429)	−0.053 (0.434)	0.002 (0.394)	0.643 (0.343)*	0.886 (0.433)**	10.69***
Instruments: (i) fraction public housing given 'in-kind', (ii) fraction of local public revenue received from state or federal, (iii) number of municipalities in MSA						
Std. Poverty Isolation Index Excluding tracts with over 60% college	−0.667 (0.429)	−0.050 (0.434)	−0.011 (0.394)	0.639 (0.343)*	0.884 (0.433)**	7 .09***
Instruments: (i) fraction public housing given 'in-kind', (ii) fraction of local public revenue received from state or federal, (iii) number of municipalities in MSA, (iv) Number of 'long' rivers in MSA						
Std. Poverty Isolation Index Excluding tracts with over 60% college	−0.608 (0.429)	0.205 (0.424)	0.180 (0.387)	0.663 (0.346)*	0.991 (0.439)**	5 .11***
Instruments: (i) fraction of local public revenue received from state or federal, (ii) number of municipalities in MSA, (iii) number of rivers in MSA						
Std. Poverty Isolation Index Excluding tracts with over 60% college	−1.073 (0.638)*	−0.602 (0.605)	−0.410 (0.555)	0.671 (0.464)	0.423 (0.546)	3.64**

Standard errors in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

work on defensible space issues is varied and shows some mixed results (Repetto, 1974; Pyle, 1976; Duffala, 1976; Molumby, 1976; Mawby, 1977), this notion that lack of defensible space in public housing projects can directly impact crime does not necessarily invalidate this instrument. In particular, the instrument can still be argued to be an exogenous source of current segregation, as two of the key aspects of the theorized relationship between lack of defensible space and crime relate directly to the model of crime developed in the previous section. Namely, the presence of unmaintained common areas that residents must pass through increases the likelihood of encounters with thugs, which all else equal will increase relative payoff to becoming a thug. Similarly, the separation of the residents of large housing projects from the broader less poverty stricken community is exactly the type of segregation at work in the model. Hence, these issues with respect to defensible space closely mirror those in the model. All this being said, one must certainly consider the possibility that the results discussed in the previous subsection are influenced by some particular unmodeled relationship between criminality and the physical environment inherent in public housing structures.

In the end, though, perhaps the strongest evidence that mitigates concerns regarding the invalidity of the proposed instruments are the differential findings across crime types in the 2SLS specifications. In particular, if the instruments were strongly related to unmeasured variables affecting opportunities for the poor and/or black residents, such as social services availability and schooling, or there was a fundamental difference between living in public housing structures versus otherwise economically similar neighborhoods, we should expect the 2SLS results to be similar across crime types, or even more strongly positive for basic property crimes. Intuitively, if the instruments are directly related to some omitted variable measuring relative depravation or lack of opportunities for the poor and/or blacks, such relative deprivation should impact basic property crime behavior in similar ways to violent crime behavior. However, the 2SLS results discussed above contradict this, and instead reveal that these instruments appear to have a negative or negligible relationship with basic property crimes, but a strong positive relationship to violent crimes.

Finally, it is also interesting to consider the differences in the empirical results across the different basic property crime categories within the context of the model. For example, the empirical results suggested that greater economic segregation has an arguably negative impact on burglary rates. In the context of the model, this suggests that an individual's payoff to burglary depends on the economic characteristics of his neighbors. This would be true if burglars focused their crimes on the other residents of their neighborhoods. This seems plausible, as a person would generally only break and enter a residence or commercial establishment if he had knowledge of something valuable to steal. Clearly, such information would be better locally than more distantly. On the other hand, the empirical results suggested that greater segregation had no impact on larceny and motor vehicle theft. In the context of the model this would suggest an individual's payoff to these crimes does not depend on the economic characteristics of his neighbors. Given larceny and motor vehicle theft generally involve taking readily observable items, perpetrators of such crimes can easily travel to other neighborhoods to commit such crimes, suggesting it to be reasonable that the payoff to such crimes has little relationship to his own neighbor's characteristics.

5. Conclusion

The model developed in this paper showed that a very standard behavioral assumption, namely that individuals incur diminishing marginal utility of money, can have substantial implications when it comes to criminal participation. In particular, the model not only showed how such an assumption can cause poverty to affect an individual's likelihood of engaging in all types of crime, but also that "neighborhood effects" can actually arise under very minimal additional assumptions, particularly when it comes to violent crime.

Importantly, the model also showed that the diminishing marginal utility of money assumption will mean that this neighborhood effect with respect to violent crime will be stronger for poor individuals than non-poor individuals, or in other words, that violent criminal behavior of poor individuals may be more influenced by their neighborhood economic characteristics than is the violent criminal behavior of non-poor individuals. This in turn was shown to imply that while greater (exogenous) economic segregation might have no effect or potentially even a negative effect on the overall amount of basic property crime, it may be expected to lead to a higher level of violence than would occur if the poor were more evenly dispersed throughout the city.

While this implication was shown to be consistent with several empirical findings, certainly more evidence is necessary to definitively conclude that this model provides an important component regarding the connections between crime and poverty. However, if true, this theoretical model leads to a very important conclusion. Namely, that when it comes to violent crime, not only do an individual's own economic characteristics matter, but so do the economic characteristics of his neighbors. Therefore, while it is clear that policies dictating how public housing is allocated and how an urban area is developed will affect who is *victimized* by crime, such policies may also have a significant impact on who *commits* crime and the overall amount of crime that occurs. Moreover, the results of this paper suggest that demolition and decreased use of large housing projects over the last decade may be an important component in declining rates of violence over the last decade.

Acknowledgments

Thanks to Jenny Hunt, Lance Lochner, Nicolas Marceau, John Donohue, Rucker Johnson, Jacob Klerman, Paco Martorell, Steve Raphael, and Jeffrey Timberlake, as well as seminar participants at McMaster University, UQAM, York University, Rice University, University of Houston, Texas A&M, Claremont McKenna College, RAND, The Goldman School at UC-Berkeley, USC, and San Diego State University. I also want to thank Will Strange for his insightful and helpful editorial comments, as well as two anonymous referees for their helpful suggestions.

Appendix A. Proofs

A.1. Proof of Proposition 1

In the context of this model, given an overall thug participation rate in the neighborhood of π, optimal behavior will imply that the thug participation rate for individuals of income level $j \in \{h, \ell\}$ will equal $\pi_j = \Phi_v(\pi(\delta_t(\omega_j) + (1-p)c) + [1-\pi]\ \ \delta_a(\omega_j))$. Moreover, as discussed in the paper, a Nash Equilibrium for any given neighborhood k in this environment will be pair of thug participation rates for each income level, $\{\pi_\ell, \pi_h\}$, such that when each individual chooses optimally given these thug participation rates and neighborhood poverty rate λ_k, the resulting overall thug participation rate in neighborhood k actually equals $\pi = \lambda_k\pi_\ell + (1-\lambda_k)\pi_h$. Therefore, for a given neighborhood with poverty rate λ_k, an equilibrium will be characterized by a pair $\{\pi_\ell, \pi_h\}$ that jointly solve

$$\pi_\ell = \Phi_v([\lambda_k\pi_\ell + (1-\lambda_k)\pi_h](\delta_t(\omega_\ell) + (1-p)c) + [1 - [\lambda_k\pi_\ell + (1-\lambda_k)\pi_h]]\delta_a(\omega_\ell)),$$

$$\pi_h = \Phi_v([\lambda_k\pi_\ell + (1-\lambda_k)\pi_h](\delta_t(\omega_h) + (1-p)c) + [1 - [\lambda_k\pi_\ell + (1-\lambda_k)\pi_h]]\delta_a(\omega_h)).$$

Table A1
OLS and 2SLS regression results for murder.

Variable	OLS	2SLS
Std. Poverty Isolation Index	0.400	−0.149
(excluding tracts with over 60% in college)	(0.111)***	(0.372)
Officers per 1000 residents in 1999	0.035	0.052
	(0.033)	(0.035)
Percent in poverty	−0.070	0.053
	(0.035)**	(0.087)
Log of population	0.156	0.231
	(0.066)**	(0.082)***
Percent urban	0.008	0.015
	(0.007)	(0.008)*
Percent immigrant	−0.095	−0.122
	(0.057)*	(0.060)**
Percent black	0.052	0.064
	(0.010)***	(0.013)***
Percent hispanic	0.005	0.004
	(0.006)	(0.006)
Percent with college degree	−0.030	−0.033
	(0.009)***	(0.010)***
Percent of HH with single mother	0.001	0.020
	(0.057)	(0.059)
Percent of HH subsidized	0.014	−0.026
	(0.054)	(0.061)
Unemployment rate	−0.009	−0.031
	(0.048)	(0.051)
Percent of days above 90°.	0.090	0.063
	(0.028)***	(0.033)*
Sq. of percent of days above 90	−0.002	−0.002
	(0.001)	(0.001)*
Percent of days below 32°	0.023	0.030
	(0.015)	(0.016)*
Sq. of percent of days below 32	−0.000	−0.000
	(0.000)	(0.000)
Constant	−2.487	−5.45
	(1.091)**	(2.210)**
N	228	228
F-statistic on excluded instruments	*n/a*	10.69***
p-value on Sargan stat. (i.e. overidentification test)	*n/a*	0.16

Standard errors in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

With some re-arranging, and defining $\Delta\delta_j = [\delta_t(\omega_j) + (1-p)c - \delta_a(\omega_j)]$, an equilibrium will be characterized by a pair $\{\pi_\ell, \pi_h\}$ that jointly solve

$$\pi_\ell = \Phi_v(\Delta\delta_\ell[\lambda_k\pi_\ell + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_\ell)), \tag{5}$$
$$\pi_h = \Phi_v(\Delta\delta_h[\lambda_k\pi_\ell + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_h)). \tag{6}$$

To prove there exists such an equilibrium and it is unique for any given λ_k let us first use Eq. (5) to define the function

$$g(\pi_\ell|\pi_h) = \Phi_v(\Delta\delta_\ell[\lambda_k\pi_\ell + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_\ell)) - \pi_\ell. \tag{7}$$

Given this definition, consider the following lemma.

Lemma 1. *For any given $\lambda_k \in [0,1]$, there exists a unique one-to-one mapping of π_h into [0,1], denoted by $\pi_\ell(\pi_h)$, such that $g(\pi_\ell|\pi_h) = 0$ if and only if $\pi_\ell = \pi_\ell(\pi_h)$.*

Proof. First note that given Φ_v is the normal CDF, we know $\Phi_v(\Delta\delta_\ell[\lambda_k 0 + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_\ell)) - 0$ must be greater than zero for any $\pi_h \in [0,1]$ since $\delta_a(\omega_\ell) > 0$. Therefore, for any $\pi_h \in [0,1]$, it must be true that $g(0|\pi_h) > 0$. Similarly, again noting that Φ_v is the normal CDF, we know $\Phi_v(\Delta\delta_\ell[\lambda_k 1 + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_\ell)) - 1$ must be less than 0 since $\Phi_v(z) < 1$ for all $z \in \Re$, implying that for any $\pi_h \in [0,1]$ it must be true that $g(1|\pi_h) < 0$. Finally, note that $\frac{dg(\pi_\ell|\pi_h)}{d\pi_\ell} = \phi_v(X)\Delta\delta_\ell\lambda_k - 1$ (where $X = \Delta\delta_\ell[\lambda_k\pi_\ell + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_\ell)$). Recognizing that $\phi_v(\mu_v) = \frac{1}{\sigma_v\sqrt{2\pi}}$ and $\phi_v(\mu_v) > \phi_v(z)$ for all $z \in \Re$ (since ϕ_v being the normal pdf), and recalling the regularity assumption $\frac{\delta_t(\omega_\ell)+(1-p)c-\delta_a(\omega_\ell)}{\sqrt{2\pi}} < \sigma_v$, or equivalently $\frac{1}{\sigma_v\sqrt{2\pi}}\Delta\delta_\ell < 1$, we know $\phi_v(X)\Delta\delta_\ell\lambda_k < 1$, implying $\frac{dg(\pi_\ell|\pi_h)}{d\pi_\ell} < 0$, meaning $g(\pi_\ell|\pi_h)$ is strictly decreasing in π_ℓ. Therefore, since $g(\pi_\ell|\pi_h)$ is a continuous function in π_ℓ, with $g(0|\pi_h) > 0$ and $g(1|\pi_h) < 0$, by Balzano's theorem we know that for any π_h there exists a π_ℓ such that $g(\pi_\ell|\pi_h) = 0$, and moreover since $g(\pi_\ell|\pi_h)$ strictly decreasing in π_ℓ, the value for which $g(\pi_\ell|\pi_h) = 0$ must be unique for any given π_h. The unique value of π_ℓ such that $g(\pi_\ell|\pi_h) = 0$ for each $\pi_h \in [0,1]$ defines a function $\pi_\ell(\pi_h)$ which provides a one-to-one mapping of π_h into [0,1]. □

Intuitively, Lemma 1 shows that for any given $\pi_h \in [0,1]$ there exists a unique π_ℓ that solves Eq. (5). Next, the following Lemma describes a key property of the function defined by Lemma 1.

Lemma 2. $0 < \frac{d\pi_\ell(\pi_h)}{d\pi_h} < 1$.

Proof. First, recall from above that $\pi_\ell(\pi_h)$ was implicitly defined to be such that $g(\pi_\ell(\pi_h)|\pi_h) = 0$. Therefore, $\pi_\ell(\pi_h)$ must react to a small increase in π_h in such a way that $g(\pi_\ell(\pi_h)|\pi_h)$ still equals zero, meaning $\frac{dg(\pi_\ell(\pi_h)|\pi_h)}{d\pi_h} = 0$, or substituting $\pi_\ell(\pi_h)$ in for π_ℓ in Eq. (7), taking the derivative, and setting it equal to zero gives

$$\phi_v(X)\Delta\delta_\ell\left[\lambda_k\frac{d\pi_\ell(\pi_h)}{d\pi_h} + (1-\lambda_k)\right] - \frac{d\pi_\ell(\pi_h)}{d\pi_h} = 0,$$

where again $X = \Delta\delta_\ell[\lambda_k\pi_\ell + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_\ell)$). Re-arranging the above expression gives

$$\frac{d\pi_\ell(\pi_h)}{d\pi_h} = \phi_v(X)\Delta\delta_a\frac{1-\lambda_k}{1-\phi_v(X)\Delta\delta_\ell\lambda_k}.$$

It is easy to confirm that the above expression implies $\frac{d\pi_\ell(\pi_h)}{d\pi_h} > 0$. Moreover, the regularity condition $\frac{1}{\sigma_v\sqrt{2\pi}}\Delta\delta_\ell < 1$ will be sufficient to ensure that $\frac{d\pi_\ell(\pi_h)}{d\pi_h} < 1$. To see this is true, start with the expression $\frac{d\pi_\ell(\pi_h)}{d\pi_h} = \phi_v(X)\Delta\delta_\ell\frac{1-\lambda_k}{1-\phi_v(X)\Delta\delta_a\lambda_k} < 1$, which we can re-arrange to obtain $\phi_v(X)\Delta\delta_\ell(1-\lambda_k) < 1 - \phi_v(X)\Delta\delta_\ell\lambda_k$, which can be re-arranged again to be $\phi_v(X)\Delta\delta_\ell(1-\lambda_k+\lambda_k) < 1$, or $\phi_v(X)\Delta\delta_\ell < 1$. Using a similar argument from Lemma 1, we know $\phi_v(X)\Delta\delta_\ell < \phi_v(\mu_v)\Delta\delta_\ell = \frac{1}{\sigma_v\sqrt{2\pi}}\Delta\delta_\ell$ which is less than one by the regularity assumption, confirming $\frac{d\pi_\ell(\pi_h)}{d\pi_h} < 1$. □

Finally, using Eq. (6), let us define the function $h(\pi_h) = \Phi_v(\Delta\delta_h[\lambda_k\pi_\ell(\pi_h) + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_h)) - \pi_h$. Given this definition, and using the result in Lemma 2, we can state and prove Lemma 3 below.

Lemma 3. *For any given $\lambda_k \in [0,1]$ there exists a unique value π_h^* such that $h(\pi_h) = 0$ if and only if $\pi_h = \pi_h^*$.*

Proof. Again noting that Φ_v is the normal CDF, we know $\Phi_v(\Delta\delta_h[\lambda_k\pi_\ell(0) + (1-\lambda_k)0] + 2\delta_a(\omega_h)) - 0 > 0$, implying $h(0)$ must be greater than zero for any $\lambda_k \in [0,1]$. Similarly, given Φ_v is the normal CDF, we know $\Phi_v(z) < 1$ for all $z \in \Re$, meaning $\Phi_v(\Delta\delta_h[\lambda_k\pi_\ell(1) + (1-\lambda_k)1] + 2\delta_a(\omega_h)) - 1 < 0$, which directly implies $h(1)$ must be less than zero for any $\lambda_k \in [0,1]$. Finally, note that $\frac{\partial h(\pi_h)}{\partial\pi_h} = \phi_v(Z)\Delta\delta_h(\lambda_k\frac{d\pi_\ell(\pi_h)}{d\pi_h} - (1-\lambda_k)) - 1$ (where $Z = \Delta\delta_h[\lambda_k\pi_\ell + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_h)$). Recalling that the regularity assumption implies $\phi_v(Z)\Delta\delta_\ell < 1$, which in turn implies $\phi_v(Z)\Delta\delta_h < 1$ (since $\Delta\delta_h < \Delta\delta_\ell$), and recalling from Lemma 2 that $0 < \frac{d\pi_\ell(\pi_h)}{d\pi_h} < 1$, we then know it must be true that $\frac{\partial h(\pi_h)}{\partial\pi_h} < 0$. Therefore, since $h(\pi_h)$ is a continuous function in π_h, with $h(0) > 0$ and $h(1) < 0$, by Balzano's theorem we know that there exists a π_h such that $h(\pi_h) = 0$. Moreover, since $h(\pi_h)$ is strictly decreasing, there must exists a unique value π_h^* such that $h(\pi_h) = 0$ if and only if $\pi_h = \pi_h^*$. □

Using Lemma 3, we can then define the unique value of π_h^* such that $h(\pi_h^*)=0$ for a given $\lambda_k \in [0,1]$ as $\pi_h^*(\lambda_k)$. If we then define $\pi_\ell^*(\lambda_k)=\pi_\ell(\pi_h^*(\lambda_k))$, then from Lemma 1 we know $g(\pi_\ell^*(\lambda_k) \mid \pi_h^*(\lambda_k))=0$. Given the definitions for the functions $g(\cdot)$and $h(\cdot)$ given above, we then know that $\{\pi_\ell^*(\lambda_k),\pi_h^*(\lambda_k)\}$ must the unique pair that jointly solve

$$\pi_\ell = \Phi_v(\Delta\delta_\ell[\lambda_k\pi_\ell + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_\ell)),$$
$$\pi_h = \Phi_v(\Delta\delta_h[\lambda_k\pi_\ell + (1-\lambda_k)\pi_h] + 2\delta_a(\omega_h)).$$

A.2. Proof of Proposition 2

Say $\pi_h^*(\lambda_k) \geqslant \pi_\ell^*(\lambda_k)$. From Proposition 1 (and the fact that the cdf Φ_v must be an increasing function), we then know that this would imply

$$\begin{aligned}&[\lambda_k\pi_\ell^*(\lambda_k) + (1-\lambda_k)\pi_h^*(\lambda_k)](\delta_t(\omega_h) + (1-p)c)\\ &\quad + [1-[\lambda_k\pi_\ell^*(\lambda_k) + (1-\lambda_k)\pi_h^*(\lambda_k)]]\delta_a(\omega_h)\\ &\geqslant [\lambda_k\pi_\ell^*(\lambda_k) + (1-\lambda_k)\pi_h^*(\lambda_k)](\delta_t(\omega_\ell) + (1-p)c)\\ &\quad + [1-[\lambda_k\pi_\ell^*(\lambda_k) + (1-\lambda_k)\pi_h^*(\lambda_k)]]\delta_a(\omega_\ell).\end{aligned}$$

Noting $\pi^*(\lambda_k) = \lambda_k\pi_\ell^*(\lambda_k) + (1-\lambda_k)\pi_h^*(\lambda_k)$, we can re-arrange the above expression to get

$$\pi^*(\lambda_k)[\delta_t(\omega_h) - \delta_t(\omega_\ell)] + (1-\pi^*(\lambda_k))[\delta_a(\omega_h) - \delta_a(\omega_\ell)] \geqslant 0.$$

Noting that $\delta_t(\omega_\ell) > \delta_a(\omega_\ell) > \delta_t(\omega_h) > \delta_a(\omega_h)$ (as can be seen in Fig. 1), we can see that the above equation cannot hold. Therefore, it cannot be true that $\pi_h^*(\lambda_k) \geqslant \pi_h^*(\lambda_k)$, meaning $\pi_h^*(\lambda_k) < \pi_h^*(\lambda_k)$, confirming Proposition 1.

A.3. Proof of Proposition 3

From Proposition 1 we know that in equilibrium the following equation must hold for $j = h, \ell$,

$$\begin{aligned}\pi_j^*(\lambda_k) = \Phi_v(&[\lambda_k\pi_\ell^*(\lambda_k) + (1-\lambda_k)\pi_h^*(\lambda_k)](\delta_t(\omega_j) + (1-p)c)\\ &+ [1-\lambda_k\pi_\ell^*(\lambda_k) - (1-\lambda_k)\pi_h^*(\lambda_k)]\delta_a(\omega_j)).\end{aligned}$$

Taking the derivative of the above equation with respect to λ_k and re-arranging we get

$$\begin{aligned}\frac{\partial\pi_j^*(\lambda_k)}{\partial\lambda_k} = \phi_v\left(x_j^*\right)&[\delta_t(\omega_j) - \delta_a(\omega_j)\\ &+ (1-p)c]\left[\gamma + \lambda_k\frac{\partial\pi_\ell^*(\lambda_k)}{\partial\lambda_k} + (1-\lambda_k)\frac{\partial\pi_h^*(\lambda_k)}{\partial\lambda_k}\right], \qquad (8)\end{aligned}$$

where ϕ_v is the pdf of the normal distribution evaluated at $x_j^* = [\lambda_k\pi_\ell^*(\lambda_k) + (1-\lambda_k)\pi_h^*(\lambda_k)](\delta_t(\omega_j) + (1-p)c) + [1-\lambda_k\pi_\ell^*(\lambda_k) - (1-\lambda_k)\pi_h^*(\lambda_k)]\delta_a(\omega_j)$ and $\gamma = \pi_\ell^*(\lambda_k) - \pi_h^*(\lambda_k)$ (where γ is known to be strictly positive by Proposition 2). Next, note that the above equation implicitly defines two equations of the following form:

$$A = \phi_v(x_\ell^*)\Delta\delta_\ell[\gamma + \lambda_k A + (1-\lambda_k)B], \qquad (9)$$
$$B = \phi_v(x_h^*)\Delta\delta_h[\gamma + \lambda_k A + (1-\lambda_k)B], \qquad (10)$$

where $A = \frac{d\pi_\ell^*(\lambda_k)}{d\lambda_k}$, $B = \frac{d\pi_h^*(\lambda_k)}{d\lambda_k}$, and again $\Delta\delta_j = [\delta_t(\omega_j) + (1-p)c - \delta_a(\omega_j)]$ (where $\Delta\delta_j$ can easily be confirmed to be strictly positive). Solving Eq. (9) for A then substituting into Eq. (10) and re-arranging we get

$$\frac{B}{\phi(x_h^*(\lambda_k))\Delta\delta_h} = \gamma' + \frac{\lambda\phi(x_\ell^*)\Delta\delta_\ell(1-\lambda_k)}{1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k}B + (1-\lambda_k)B, \qquad (11)$$

where $\gamma' = \frac{\gamma}{1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k}$. Now recall the assumed regularity condition that $\frac{\delta_t(\omega_\ell)+(1-p)c-\delta_a(\omega_\ell)}{\sqrt{2\pi}} < \sigma_v$. This implies $\frac{1}{\sigma_v\sqrt{2\pi}}\Delta\delta_\ell < 1$, or equivalently (again noting that $\phi(\mu_v) = \frac{1}{\sigma_v\sqrt{2\pi}}$ given ε_v^i is normally distributed) that

$$1-\phi(\mu_v)\Delta\delta_\ell > 0. \qquad (12)$$

Furthermore, given that ε_v^i is normally distributed we know $\phi(\mu_\varepsilon) \geqslant \phi(x)$ for all $x \in \Re$. Finally, recalling that $0 < \lambda_k < 1$, we know that it must be true that $1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k > 0$, and therefore $\gamma' > 0$.

We can now simplify Eq. (11) to get

$$B = \frac{\phi(x_h^*)\Delta\delta_h[1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k]}{1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k - (1-\lambda_k)\phi(x_h^*)\Delta\delta_h}\gamma'. \qquad (13)$$

Now note that the above equation implies $B > 0$ as long as (i) $1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k > 0$, and (ii) $1-\phi(x_\ell^*)\Delta\delta_\ell\lambda - (1-\lambda)\phi(x_h^*)\Delta\delta_h > 0$. Condition (i) was shown to hold above. To prove that condition (ii) will hold true, note again that $\Delta\delta_\ell > \Delta\delta_h$ (as can be confirmed in Fig. 1) and recalling $0 < \lambda_k < 1$, we know that $1-\lambda_k\phi(x_\ell^*)\Delta\delta_\ell - (1-\lambda_k)\phi(x_h^*)\Delta\delta_h > 1-\phi(x_\ell^*)\Delta\delta_\ell$. Noting again that the regularity assumption implies $1-\phi(x)\Delta\delta_\ell > 0$ for all x (as shown above), we know that condition (ii) must also hold. Therefore, $B > 0$, which confirms $\frac{\partial\pi_h^*(\lambda_k)}{\partial\lambda_k} > 0$. An analogous argument can be made to show $\frac{\partial\pi_\ell^*(\lambda_k)}{\partial\lambda_k} > 0$ by solving Eq. (10) for B and substituting into Eq. (9).

A.4. Proof of Proposition 4

First, recall Eq. (8) above

$$\begin{aligned}\frac{\partial\pi_j^*(\lambda_k)}{\partial\lambda_k} = \phi\left(x_j^*\right)&[\delta_t(\omega_j) - \delta_a(\omega_j) + (1\\ &-p)c]\left[\gamma + \lambda_k\frac{\partial\pi_\ell^*(\lambda_k)}{\partial\lambda_k} + (1-\lambda_k)\frac{\partial\pi_h^*(\lambda_k)}{\partial\lambda_k}\right].\end{aligned}$$

From the above expression we can see that $\frac{\partial\pi_\ell^*(\lambda_k)}{\partial\lambda_k} > \frac{\partial\pi_h^*(\lambda_k)}{\partial\lambda_k}$ if (i) $\delta_t(\omega_\ell) - \delta_a(\omega_\ell) > \delta_t(\omega_h) - \delta_a(\omega_h)$ and (ii) $\phi(x_\ell^*) \geqslant \phi(x_h^*)$. Given our assumption that u in an increasing strictly concave function, we know condition (i) will always be true (see Fig. 1). Regarding condition (ii), given our assumption that Φ_v is the cdf of the normal distribution and therefore convex for all x such that $\Phi(x) \leqslant 0.5$, we know $\phi_v(x_1) \geqslant \phi_v(x_2)$ for all $x_2 \leqslant x_1 \leqslant 0.5$. Therefore, recalling that

$$\pi_j^*(\lambda_k) = \Phi_v\left(x_j^*\right)$$

and $\pi_\ell^*(\lambda_k) > \pi_h^*(\lambda_k)$, we know that $\phi_v(x_\ell^*) \geqslant \phi_v(x_h^*)$ as long as $\pi_\ell^*(\lambda_k) < 0.5$. Given $\pi_\ell^*(\lambda_k) < \pi_\ell^*(1)$ (as implied by Proposition 3), this confirms that if $\pi_\ell^*(1) < 0.5$ then $\frac{\partial\pi_\ell^*(\lambda_k)}{\partial\lambda_k} > \frac{\partial\pi_h^*(\lambda_k)}{\partial\lambda_k}$ for all $\lambda_k \leqslant 1$.

A.5. Proof that $\frac{\partial^2\pi_j^*(\lambda_k)}{\partial\lambda_k^2} > 0$ for $j = h,\ell$

From the proof of Proposition 2 we know

$$\frac{\partial\pi_j^*(\lambda_k)}{\partial\lambda_k} = \phi_v(x_j^*)\Delta\delta_j\left[\gamma + \lambda_k\frac{\partial\pi_\ell^*(\lambda_k)}{\partial\lambda_k} + (1-\lambda_k)\frac{\partial\pi_h^*(\lambda_k)}{\partial\lambda_k}\right],$$

where again $x_j^* = [\lambda_k\pi_\ell^*(\lambda_k) + (1-\lambda_k)\pi_h^*(\lambda_k)](\delta_t(\omega_j) + (1-p)c) + [1-\lambda_k\pi_\ell^*(\lambda_k) - (1-\lambda_k)\pi_h^*(\lambda_k)]\delta_a(\omega_j)$, $\Delta\delta_j = [\delta_t(\omega_j) - \delta_a(\omega_j) + (1-p)c]$, *$\delta_j = [\delta_t(\omega_j) - \delta_a(\omega_j) + (1-p)c]$, and $\gamma = \pi_\ell^*(\lambda_k) - \pi_h^*(\lambda_k)$. Taking the derivative of the above expression we get*

$$\begin{aligned}\frac{\partial^2\pi_j^*(\lambda_k)}{\partial\lambda_k^2} = \phi_v'(x_j^*)\Delta\delta_j^2&\left[\gamma + \lambda_k\frac{\partial\pi_\ell^*(\lambda_k)}{\partial\lambda_k} + (1-\lambda_k)\frac{\partial\pi_h^*(\lambda_k)}{\partial\lambda_k}\right]^2\\ &+ \phi_v(x_j^*(\lambda_k))\Delta\delta_j\left[2\left(\frac{\partial\pi_\ell^*(\lambda_k)}{\partial\lambda_k} - \frac{\partial\pi_h^*(\lambda_k)}{\partial\lambda_k}\right) + \lambda_k\frac{\partial^2\pi_\ell^*(\lambda_k)}{\partial\lambda_k^2}\right.\\ &\left.+(1-\lambda_k)\frac{\partial^2\pi_h^*(\lambda_k)}{\partial\lambda_k^2}\right]\end{aligned}$$

Simplifying the above expression we get

$$\frac{\partial^2 \pi_j^*(\lambda_k)}{\partial \lambda_k^2} = \kappa_j + \phi_v\left(x_j^*\right)\Delta\delta_j\left[\eta_j + \lambda_k \frac{\partial^2 \pi_\ell^*(\lambda_k)}{\partial \lambda_k^2} + (1-\lambda_k)\frac{\partial^2 \pi_h^*(\lambda_k)}{\partial \lambda_k^2}\right],$$

where $\kappa_j = \phi'_v(x_j^*)\Delta\delta_j^2[\gamma + \lambda_k \frac{\partial \pi_\ell^*(\lambda_k)}{\partial \lambda_k} + (1-\lambda_k)\frac{\partial \pi_h^*(\lambda_k)}{\partial \lambda_k}]^2$ and $\eta_j = 2(\frac{\partial \pi_\ell^*(\lambda_k)}{\partial \lambda_k} - \frac{\partial \pi_h^*(\lambda_k)}{\partial \lambda_k})$. From Proposition 2 we know $\kappa_j > 0$ as long as $\pi_\ell^*(1) \leqslant 0.5$ (which will ensure $\phi'_v(x_j^*) > 0$ as discussed above in the proof to Proposition 3), and from Proposition 3 we know $\eta_j > 0$. We can simplify the above equation one more time to become

$$\frac{\partial^2 \pi_j^*(\lambda_k)}{\partial \lambda_k^2} = \phi_v(x_j^*)\Delta\delta_j\left[\gamma_j + \lambda_k \frac{\partial^2 \pi_\ell^*(\lambda_k)}{\partial \lambda_k^2} + (1-\lambda_k)\frac{\partial^2 \pi_h^*(\lambda_k)}{\partial \lambda_k^2}\right], \qquad (14)$$

where now $\gamma_j = \frac{\kappa_j}{\phi_v(x_j^*)\Delta\delta_j} + \eta_j$. Note again that since $\kappa_j > 0$ and $\eta_j > 0$ (as shown above), it must also be true that $\gamma_j > 0$.

As in the proof of Proposition 2, note that Eq. (14) implicitly defines two equations

$$A = \phi_v\left(x_\ell^*\right)\Delta\delta_\ell[\gamma_\ell + \lambda_k A + (1-\lambda_k)B], \qquad (15)$$

$$B = \phi_v\left(x_h^*\right)\Delta\delta_h[\gamma_h + \lambda_k A + (1-\lambda_k)B], \qquad (16)$$

where $A = \frac{\partial^2 \pi_\ell^*(\lambda_k)}{\partial \lambda_k^2}$, $B = \frac{\partial^2 \pi_h^*(\lambda_k)}{\partial \lambda_k^2}$. Solving Eq. (15) for A then substituting into Eq. (16) and re-arranging we get

$$\frac{B}{\phi(x_h^*)\Delta\delta_h} = \gamma'_h + \frac{\lambda\phi\left(x_\ell^*\right)\Delta\delta_\ell(1-\lambda_k)}{1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k}B + (1-\lambda_k)B, \qquad (17)$$

where $\gamma'_h = \frac{\gamma_h(1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k) + \gamma_\ell\phi(x_\ell^*)\Delta\delta_\ell\lambda_k}{1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k}$. Note that $\gamma'_h > 0$ as long as $1-\phi(x_\ell^*)\Delta\delta_\ell\lambda_k > 0$. As shown in the proof to Proposition 2, this will always hold given the assumption that ε_v^i is normally distributed and the regularity condition that $\frac{\Delta\delta_\ell}{\sqrt{2\pi}} < \sigma_v$. Given this, note that Eq. (17) is of the same form as Eq. (13) in the proof to Proposition 2. Therefore, the proof that $B > 0$ follows identically from that in the analogous proof to Proposition 2, meaning $\frac{\partial^2 \pi_h^*(\lambda_k)}{\partial \lambda_k^2} > 0$. As in the proof to Proposition 2, the proof that $A > 0$, and therefore that $\frac{\partial^2 \pi_\ell^*(\lambda_k)}{\partial \lambda_k^2} > 0$, follows in a similar manner.

References

Anderson, E., 1999. Code of the Street: Decency, Violence and the Moral Life of the Inner City. W.W. Norton and Company, New York.
Bayer, P., McMillan, R., Rueben, K., 2004. What drives racial segregation? New evidence using census microdata. Journal of Urban Economics 56 (3), 535–574.
Blau, J.R., Blau, P.M., 1982. The cost of inequality: metropolitan structure and violent crime. American Sociological Review 47, 114–129.
Brueckner, J., Rosenthal, S., 2009. Gentrification and neighborhood housing cycles: will America's future downtowns be rich? Review of Economics and Statistics 91 (4), 725–743.
Brock, W., Durlauf, S.N., 2001. Discrete choice with social interactions. Review of Economic Studies 68 (2), 235–260.
Calvo-Armengol, A., Zenou, Y., 2004. Social networks and crime decisions: the role of social structure in facilitating delinquent behavior. International Economic Review 45 (3), 939–958.
Calvo-Armengol, A., Verdier, T., Zenou, Y., 2007. Strong ties and weak ties in employment and crime. Journal of Public Economics 91, 203–233.
Cullen, J.B., Levitt, S.D., 1999. Crime, urban flight, and the consequences for cities. Review of Economics and Statistics 81 (2), 159–169.
Cutler, D., Glaeser, E., 1997. Are ghettos good or bad? Quarterly Journal of Economics 112 (3), 827–872.
Duffala, T., 1976. Convenience stores: armed robbery, and the physical environmental features. American Behavioral Scientist 20, 227–247.
Federal Bureau of Investigation, 2000. Crime in the United States 2000, Uniform Crime Reports. US Department of Justice, Washington, DC.
Ferrer, R., 2010. Breaking the law when others do: a model of law enforcement with neighborhood externalities. European Economic Review 54 (2), 163–180.
Glaeser, E., Sacerdote, B., Scheinkman, J., 1996. Crime and social interactions. Quarterly Journal of Economics 111, 507–548.
Heavner, L., Lochner, L., 2002. Social Networks and the Aggregation of Individual Decision. Working Paper, University of Western Ontario, Department of Economics.
Helsley, R.W., Strange, W.C., 2005. Mixed markets and crime. Journal of Public Economics 89, 1251–1275.
Hoxby, C., 2007. Does competition among public schools benefit students and taxpayers? Reply. American Economic Review 97 (5), 2026–2037.
Jacob, B., Lefgren, L., Moretti, E., 2007. Dynamics of criminal behavior: evidence from weather shocks. Journal of Human Resources 42 (3), 489–527.
Jacobs, J., 1961. The Death and Life of Great American Cities. Random House, New York.
Jankowski, M.S., 1991. Islands in the Street: Gangs and American Urban Society. University of California Press, Berkeley.
Kling, J., Ludwig, J., Katz, L., 2005. Neighborhood effects on crime for female and male youth: evidence from a randomized housing voucher experiment. Quarterly Journal of Economics 120 (1).
Kotlowitz, A., 1991. There Are No Children Here. Random House Inc., New York.
Krivo, L., Peterson, R.D., 1996. Extremely disadvantaged neighborhoods and urban crime. Social Forces 75 (2), 619–650.
Landesman, P., 2007. Nine miles and spreading. LA Weekly, December 14–20, pp. 44–51.
Massey, D.S., 1995. Getting away with murder: segregation and violent crime in urban America. University of Pennsylvania Law Review 143 (5), 1203–1232.
Mawby, R.I., 1977. Defensible space: a theoretical and empirical appraisal. Urban Studies 14, 169–180.
Massey, D.S., Denton, N., 1988. Suburbanization and segregation in US metropolitan areas. American Journal of Sociology 94 (3), 592–626.
Messner, S.F., Tardiff, K., 1986. Economic inequality and levels of homicide: an analysis of urban neighborhoods. Criminology 24, 297–317.
Molumby, T., 1976. Patterns of crime in a university housing project. American Behavioral Scientist 20, 247–261.
Newman, O., 1972. Architectural Design for Crime Prevention. US Department of Justice, Washington, DC.
O'Flaherty, B., Sethi, R., 2007. Crime and segregation. Journal of Economic Behavior and Organization 64, 391–405.
O'Flaherty, B., Sethi, R., 2010. Peaceable kingdoms and war zones: preemption, ballistics and murder in newark. In: DiTella, R., Edwards, S., Schargrodsky, E. (Eds.), The Economics of Crime: Lessons for and from Latin America. University of Chicago Press, Chicago.
O'Flaherty, B., Sethi, R., 2010. Homicide in black and white. Journal of Urban Economics 68 (3), 215–230.
O'Flaherty, B., Sethi, R., 2010c. Racial geography of street vice. Journal of Urban Economics. 67 (3), 270–286.
Patacchini, E., Zenou, Y., 2008. The strength of weak ties in crime. European Economic Review 52, 209–236.
Patterson, E.B., 1991. Poverty, income inequality, and community crime rates. Criminology 29, 755–776.
Pyle, G.F., 1976. Spatial and temporal aspects of crime in Cleveland, Ohio. American Behavioral Scientist 20, 175–199.
Repetto, T.A., 1974. Residential Crime. Ballinger, Cambridge, MA.
Rothstein, J., 2007. Does competition among public schools benefit students and taxpayers? A comment on Hoxby (2000). American Economic Review 97 (5), 2026–2037.
Savitz, L.D., Rosen, L., Lalli, M., 1980. Delinquency and gang membership as related to victimization. Victimology 5 (2–4), 152–160.
Sethi, R., Somanathan, R., 2004. Inequality and segregation. Journal of Political Economy 112 (6), 1296–1321.
Shihadeh, E., Flynn, N., 1996. Segregation and crime: the effect of black isolation on the rates of black urban violence. Social Forces 74 (4), 1325–1352.
Silverman, D., 2004. Street crime and street culture. International Economic Review 45 (3), 761–786.
Spergel, I.A., 1990. Youth gangs: continuity and change. Crime and Justice 12, 171–275.
Stock, J., Yogo, M., 2002. Testing for Weak Instruments in Linear IV Regression. NBER Technical Working Paper 284.
Verdier, T., Zenou, Y., 2004. Racial beliefs, location, and the causes of crime. International Economic Review 45 (3), 731–760.
Wilson, W.J., 1987. The Truly Disadvantaged. University of Chicago Press, Chicago.
Wooldridge, J.W., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge MA.