DISCUSSION PAPER SERIES

No. 10707

PROPERTY CRIME AND PRIVATE PROTECTION ALLOCATION WITHIN CITIES: THEORY AND EVIDENCE

Bruno Decreuse, Steeve Mongrain
and Tanguy van Ypersele

PUBLIC ECONOMICS

Centre for Economic Policy Research

ISSN 0265-8003

PROPERTY CRIME AND PRIVATE PROTECTION ALLOCATION WITHIN CITIES: THEORY AND EVIDENCE

Bruno Decreuse, Steeve Mongrain and Tanguy van Ypersele

Discussion Paper No. 10707
July 2015
Submitted 30 June 2015

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

PROPERTY CRIME AND PRIVATE PROTECTION ALLOCATION WITHIN CITIES: THEORY AND EVIDENCE[†]

Abstract

We model the allocation of property crime and private protection within cities. We provide a theory where city-specific criminals choose a neighborhood and whether they pay a search cost to compare potential victims, whereas households invest in self-protection. The model features strategic complementarity between criminals' search efforts and households' protection investments. As criminals' return to search increases with neighborhood wealth, households in rich neighborhoods are more likely to enter a rat race to ever greater protection that drives criminals towards poorer areas. The mechanisms of our model are tested with the Canadian General Social Survey. Household protection increases with household and neighborhood incomes, neighborhood protection, and neighborhood victimization.

JEL Classification: K14 and K42
Keywords: economics of crime, private protection, search frictions and social multiplier

Bruno Decreuse decreuse@univmed.fr
Aix-Marseille University

Steeve Mongrain mongrain@sfu.ca
Simon Fraser University

Tanguy van Ypersele tanguy.vy@univ-amu.fr
Aix-Marseille University and CEPR

[†] We wish to thank Jean Marie Baland, Paul Beaudry, Pierre-Philippe Combes, Patrick Francois, Cecilia Garcia-Penalosa, Krishna Pendakur, Marc Sangnier, the RDC staff at SFU and UBC, seminar participants in Aix-Marseille University, Université de Cergy-Pontoise, University of Namur, University of Paris-Ouest, University of Alberta, University of Rennes, CES-Ifo, and participants at ENSAI days in Rennes and at the SaM conference in Edinburgh. We are grateful to Statistics Canada for making the data available. Our data also include information copied with permission from Canada Post Corporation. Steeve Mongrain thanks the Sciences and Humanities Research Council of Canada for financial support. Bruno Decreuse and Tanguy van Ypersele thank the Institut Universitaire de France.

1 Introduction

This paper provides a theory of property crime and protection investment allocation within cities. Since Becker (1968), most of the research addressing the determinants and consequences of crime focuses on the supply of criminal activities. By contrast, less has been done to understand the demand side of crime. However, many individuals respond to the threat of crime to their household by buying locks or alarms or other forms of protection (see Shavell, 1991 and Hotte and van Ypersele, 2008, for early analyses). Although there are no studies documenting the effectiveness of such tools on crime reduction, they can be considered to affect the location of criminal activities.

We are motivated by three Canadian stylized facts: (i) there is no correlation between household income and break-ins, (ii) rich neighborhoods are less victimized than poor ones, (iii) rich households as well as rich neighborhoods invest more in protection. Figure 1 shows the yearly mean victimization rate is about 3.4% and is roughly the same for households below and above the median household income. It also shows that neighborhoods below the median income experience victimization rates 35% higher than neighborhoods above the median income. Thus, rich people are more victimized than their neighbors. Figure 2 shows that the percentage of households equipped with an alarm is about 30-35% larger among households above the median income than among households below the median income.

Beyond Canada, the correlation between household income and exposure to property crime is very volatile throughout the world, whereas rich households always invest more in private protection. In his survey, Freeman (1996) reports that "The [US] rate of victimization for violent crimes (which range from robbery to assault to rape) is inversely related to household income, while the rate of victimization for property crimes rises only modestly with income". Figure 3 uses data from the International Crime Victim Survey for a selection of 15 OECD countries / provinces. It displays two ratios: the ratio of the yearly victimization probability for more-than-median income to less-than-median income households, and a similar ratio of household shares who have an alarm. The former ratio is below one in half of the cases, whereas the latter ratio is always above one.

Section 2 describes a model where a fixed number of criminals allocate across two neighborhoods, whereas households invest in protection. Once in a neighborhood, criminals randomly pick one of the households. They can also make a search effort to draw an additional potential victim. They then select the theft opportunity with the highest payoff. Protection heterogeneity provides an incentive to search. In turn, criminal effort motivates protection investment. Thus there is strategic complementarity between criminals' search efforts and households' protection investments. Strategic complementarity originates a social multiplier in individual protection investment decisions: beyond any price interaction, my neighbor's choice of installing an alarm increases my own need for such a protection device.

In the absence of protection investment, the rich neighborhood is more attractive to criminals. However, its residents also invest more in protection, which tends to repel criminals. Indeed, criminals in the rich neighborhood are more willing to spend resources to compare potential victims than in the poor neighborhood. Thus rich households expect to be frequently compared to their neighbors. This affects both the mean (people invest more on average) and the dispersion of protection investments (there are still some people who choose not to invest, whereas others invest more). That the rich neighborhood invests more in protection implies low returns to crime, and therefore criminals may be attracted to the poor neighborhood. Meanwhile, rich households are more victimized than poorer ones in each neighborhood. A parameterization of our model broadly reproduces the patterns displayed by Figures 1 and 2.

Section 3 quantifies the roles played by the different ingredients of our theory. We use three waves of the Canadian General Social Survey. The Canadian territory is divided into Census Metropolitan Areas (limiting criminals' mobility) and Forward Sortation Areas (the neighborhoods). We do not study individual victimization, whose measure is too noisy at neighborhood level. Instead, we focus on individual protection and estimate the magnitude of the social multiplier. In theory, individual investment depends on household income, FSA income, FSA income inequality, FSA mean protection investments, and FSA victimization. The statistical model is a variant of the linear-in-mean model of social interactions. Its estimation involves a specific IV strategy. Namely, we follow Manski (1993) and instrument FSA mean protection investments by FSA mean characteristics. As for FSA victimization, we use our theoretical model to select the following instruments: CMA victimization (the supply of criminals at city level), and mean characteristics of alternative FSAs belonging to the same CMA (because criminals compare the different neighborhoods).

The data feature the kind of strategic complementarity between household investments that our model predicts: households invest more in protection when their neighbors are well protected. The coefficient associated with mean protection investment is about 0.3. Thus the social multiplier is about $1/0.7 = 1.4$. Richer households and households in richer and more unequal neighborhoods also invest more in protection. Finally, more victimized neighborhoods provide incentives to invest in protection. Accounting for the social multiplier, a one-standard deviation increase in household income translates into a 4.5 percentage point increase in the probability of having an alarm, about 30% of its standard deviation. Increasing neighborhood victimization by one standard deviation raises the former probability by 2 percentage points.

We now discuss our assumptions in light of the literature. First, note that we standardily model private protection (see Shavell, 1991, for an early discussion, and Hotte and van Ypersele, 2008, for a normative analysis). Private protection has three effects: crime diversion, which makes one's belongings less attractive to criminals, theft reduction, which decreases what is stolen in the case of victimization, and deterrence, which implies that the entry into criminal activity drops. In our framework, the total amount of crime is fixed at city level. Thus deterrence occurs at neighborhood level.

Second, we model imperfect information with a search block in each neighborhood. Imperfect information is essential to our theory. If criminals could sample as many households as they want at no cost, then the incentive to invest in protection would not differ with neighborhood wealth. The exact formulation of the model is consensual in the search community: we heavily borrow from Burdett and Judd (1983) who focus on the good market. In this model, customers choose the number of buy offers they sample from the distribution of offer prices. They typically sample two prices, and there is price dispersion as a result (see also Acemoglu and Shimer, 2000, Albrecht et al, and Galenianos and Kircher, 2009, for labor market applications, and Moen, 1999, for an application to education). Beyond our subject, the main innovation of the paper consists in considering two different market places, e.g. neighborhoods, that are interconnected through the return to property crime. We also innovate by introducing within-neighborhood heterogeneity. Both assumptions are fundamental to answer the general puzzle exposed above.

Third, protection costs and gains are proportional to property value, which rules out cases where rich households find it easier to invest in protection. We in no way suggest that such cases do not exist. However, opposite effects may also occur: for instance, rich households may not have much time to allocate to self-protection activities. Our theory refers to a more general principle linking neighborhood income to the protection incentive: the criminals' return to search increases with neighborhood income. In the same vein, this link is not due to cooperative behavior at neighborhood level. Households do not take into account that protection investments deter criminals from prospecting the neighborhood. The non-cooperative protection game played by neighbors leads to an escalation in protection investments, which drives criminals out of the neighborhood.

Our model is close in spirit to social interaction models of criminal activities featuring a social multiplier (see Glaeser et al, 1996, Zenou, 2003, Calvo-Armengol et al, 2007, and Patacchini and Zenou, 2008). In these models, individuals are more prone to engage in criminal activities when their peers also adopt criminal behavior. We highlight a novel mechanism of social influence that can lead to multiple equilibria. Namely, there is strategic complementarity among households (the demand side), and also between households and criminals (the supply side). When protection investments are more dispersed, criminals have a stronger incentive to search, which further raises incentive to protect and implies protection heterogeneity.

Finally, there already exist search models of crime, but they do not focus on protection. A first strand of models study the joint determination of crime, unemployment, and inequality (see, e.g., Burdett et al, 2003, 2004; Huang et al, 2004; Engelhardt et al, 2008). Galenianos et al (2012) provide a search-theoretic model of the retail market for illegal drugs. The model predicts a non-degenerate distribution of drug purity offered in equilibrium, which they test with US data.

2 Property crime and protection: theory

2.1 The model

Model assumptions.—The model is static. There are two neighborhoods indexed by $j \in \{A, B\}$. Each neighborhood is composed of K_j residents. Residents may differ across neighborhoods, but they are homogenous within each neighborhood. All residents in neighborhood j have the same wealth endowment $V_j > 0$. We assume $V_A > V_B$ and refer to neighborhood A as the rich one, and to neighborhood B as the poor one.

There are also C criminals. They have limited mobility across neighborhoods. Each criminal must choose a neighborhood and then commits one crime in the neighborhood.

The number of criminals in neighborhood j is C_j. Thus the crime rate in neighborhood j is $c_j = C_j/K_j$. By definition, we have $C_A + C_B = C$.

Residents independently choose the level of self-protection $\theta \geq 0$. This comes at the cost $\gamma\theta V$, $\gamma > 0$. Protection reduces the loss incurred during a theft. This loss is $(\alpha - \theta)V$, with $\alpha \in (0, 1)$ and is proportional to wealth. However, this proportion is decreased by θ. The loss $(\alpha - \theta)V$ is also the criminal's gain.[1]

We denote by H_j the equilibrium distribution of protection efforts in neighborhood j. We also denote by $\Theta \subset \mathbb{R}_+$ the support of this distribution. We assume that the crime rate is bounded as follows: $C/\min\{K_1, K_2\} < \gamma/2$. This restriction ensures that the highest protection level $\bar{\theta}$ is lower than α.

Wealth and protection are imperfectly observable. Criminals who plan a robbery in neighborhood j are presented with one theft opportunity at random. They then observe the protection level θ. They may also benefit from a second opportunity. This comes at the cost s. Thus the margin of decision is whether to pay the cost and compare two households or not.

We denote by C_{j1} the expected number of criminals in neighborhood j who have a single theft option (*single-option criminals*), and by C_{j2} the expected number of criminals who have two options (*double-option criminals*). Similarly, c_{j1} and c_{j2} denote the corresponding criminal-to-household ratios.

Model agenda.—The timing is as follows.

Stage 1. Households choose their protection level θ.

Stage 2. Criminals choose a neighborhood $j \in \{A, B\}$ and whether to pay the search cost or not.

Stage 3. Criminals with two theft opportunities select the most interesting one. If indifferent, criminals select one of the two options with probability one half.

We will solve the model in two steps. In the first step, we characterize the equilibrium

[1] We may distinguish the household's loss from the criminal's gain. Households may be insured, and consequently they suffer smaller losses, but they may also endure larger losses because of the psychological cost associated with crime. The model can easily be extended to cases where the loss and the gain differ by a factor of proportion.

at given criminals' location choice (equilibrium with immobile criminals). In the second step, we study the allocation of criminals across the two neighborhoods (equilibrium with mobile criminals). To simplify notations, we neglect the neighborhood index j until we discuss criminals' location choices.

Agents' payoffs.—Let Ω denote the expected payoff for a given criminal. When criminals play mixed strategies, we have:

$$\Omega = \alpha V - (1-q)\mathbb{E}(\theta)V - q\mathbb{E}(\min\{\theta,\theta'\})V - sq, \tag{1}$$

where $q \in [0,1]$ is the probability of making the search effort. A theft opportunity is a random draw in the distribution of households' protection investments. When criminals have a single theft opportunity, their expected payoff thus depends on the unconditional mean $\mathbb{E}(\theta)$ of protection investments. When they have two opportunities, they choose the least protected household. Their expected payoff now depends on the mean of the minimum protection level.

The return to search is the difference in expected loss due to protection:

$$\Gamma = [\mathbb{E}(\theta) - \mathbb{E}(\min\{\theta,\theta'\})]\, V. \tag{2}$$

The return to search is proportional to wealth. Having two theft options is advantageous when protection levels are heterogenous. When the distribution of protection investments collapses to a single mass point, then $\Gamma = 0$ and criminals have no reason to search.

We now define the expected payoff W of a household with protection θ. Let η_1 and η_2 denote the expected number of visits by single-option and double-option criminals. We have

$$W = V - (\eta_1 + \eta_2)(\alpha - \theta)V - \gamma\theta V. \tag{3}$$

Wealth is reduced by theft and by the protection cost.

Appendix B computes the numbers of visits by single-option and double-option criminals. It leads to

$$\eta_1 = c_1 = c(1-q), \tag{4}$$

$$\eta_2(\theta) = 2c_2\Big[1 - H(\theta)\Big] = 2cq\Big[1 - H(\theta)\Big]. \tag{5}$$

Single-option criminals randomly sample within the household set. Thus η_1 is equal to the ratio of such criminals to households. As for the expected number of visits by double-option criminals, when a given household is compared to another potential victim, the household with the lowest protection level is robbed whereas the other stays safe. By assumption all households sample their protection investment in the distribution H. Thus the proportion of households with a protection level above θ is simply $1 - H(\theta)$.

2.2 Equilibrium with immobile criminals

We focus on protection investment and search effort at a given number of criminals C and corresponding crime rate $c = C/K$. We start with the equilibrium definition. We then examine existence and uniqueness. We finally turn to comparative statics.

Definition 1 An *equilibrium with immobile criminals* (EIC) is a search effort q^* and a cdf H such that

(i) $\theta \in \Theta$ if and only if (iff) $\theta \in \arg\max_{\theta' \geq 0} W(\theta', q^*)$,

(ii) $q^* \in \arg\max_{q \in [0,1]} \Omega(q, H(., q^*))$.

In equilibrium, protection efforts maximize households' well-being, whereas the search effort maximizes criminals' payoffs. If there is a unique value of θ that maximizes individual well-being, then the distribution H is degenerate. Otherwise, the exact distribution results from the equality of payoffs $W(\theta, .)$ over the equilibrium support of the distribution.

Proposition 1 (Equilibrium distribution of protection investment) *Let* $x(q) = \frac{2cq}{\gamma - c(1-q)}$. *In an EIC, the equilibrium distribution of protection investment is such that*

(i) the cdf H and the pdf h are defined by

$$H(\theta) = \frac{\theta}{\alpha - \theta} \Big/ \frac{\bar{\theta}}{\alpha - \bar{\theta}}, \tag{6}$$

$$h(\theta) = \frac{\alpha}{(\alpha - \theta)^2} \Big/ \frac{\bar{\theta}}{\alpha - \bar{\theta}}, \tag{7}$$

for all $\theta \in \Theta = \left[0, \bar{\theta}\right]$, with $\bar{\theta} = \alpha x(q^*)$;

(ii) the upper bound $\bar{\theta}$ of the support is strictly increasing in the search probability q and crime rate c;

(iii) the density is strictly decreasing in q and c, but the effect decreases with θ, i.e. $h_q(.) < 0$ and $h_c(.) < 0$, whereas $h_{q\theta}(.) > 0$ and $h_{c\theta}(.) > 0$.

There is a continuous equilibrium distribution of protection investment. This result closely mimics Burdett and Judd (1983). In their model devoted to the goods market, prices are imperefctly observed. Some of the buyers sample two firms, and there is an equilibrium distribution of prices as a result. Here similar households invest differently because protection is a positional good. If the distribution were not continuous, investing a little more would achieve a mass gain. The household would only pay slightly more, but they would not be chosen when a criminal compared them with another household. Zero protection always belongs to the support of the equilibrium distribution. On the contrary, let us suppose that the lower bound is strictly positive. Investing less in protection would cost less and would not increase the probability of being visited. This contradicts the fact that the lower bound is strictly positive.

The density function $h(\theta) \equiv H'(\theta)$ is strictly increasing in θ. Along the equilibrium distribution, the marginal return to investment must equal the marginal cost. Formally, we have

$$[c_1 + 2c_2(1 - H(\theta))]\, V + 2c_2 h(\theta)(\alpha - \theta)V = \gamma V. \tag{8}$$

The marginal cost stands in the right-hand side. It is constant and equal to γV. The marginal return stands in the left-hand side. It has two components. The first component is due to the loss reduction effect. A marginal increase in protection reduces the impact of theft by the quantity V. The second component is due to the marginal decrease in probability of being robbed. This gain is proportional to the property loss $(\alpha - \theta)V$. When θ is large, there is not much to lose. This small gain must be compensated by a large increase in occurrence probability. Thus the density h must be increasing in θ.

The equilibrium distribution of protection investment depends on the crime rate c and on the search probability q. Both increase the expected number of visits by double-option criminals. Thus households anticipate being compared with their neighbors more frequently and invest more as a result. The first implication is that the maximum effort $\overline{\theta}$ increases, so that the support of the distribution expands. The second implication is that households assign more weight to higher investments. Thus the distribution becomes more concentrated around the highest values of investment ($h_{q\theta}(.) > 0$ and $h_{c\theta}(.) > 0$).

To compute the return to search, we need to define the distribution of $\min\{\theta, \theta'\}$. Its density is equal to $2h(\theta)[1 - H(\theta)]$. Thus,

$$\Gamma = V \int_0^{\overline{\theta}} \{h(\theta)\theta - 2h(\theta)[1 - H(\theta)]\theta\} d\theta. \tag{9}$$

The computation gives

$$\Gamma \equiv \Gamma(x(q)) = \alpha V \frac{x(q)-1}{x(q)^2}[(2-x(q))\ln(1-x(q)) + 2x(q)]. \tag{10}$$

The equilibrium resolution reduces to finding $q^* \in \arg\max_{q\in[0,1]}\{q\Gamma(x(q^*)) - sq\}$.

Proposition 2 (Existence of EIC) *Let E_{IC} be the set of EIC with typical element q^*. Let also $x_{\max} \in (0,1)$ be uniquely defined by $\Gamma'(x_{\max}) = 0$. Then,*

(i) If $x(1) < x_{\max}$, then

$$E_{IM} = \begin{cases} \{0\} \text{ if } \Gamma(x(1)) < s, \\ \{0, \underline{q}, 1\} \text{ if } \Gamma(x(1)) > s, \end{cases}$$

(ii) If $x(1) > x_{\max}$, then

$$E_{IM} = \begin{cases} \{0, \underline{q}, 1\} \text{ if } \Gamma(x(1)) > s, \\ \{0, \underline{q}, \bar{q}\} \text{ if } \Gamma(x(1)) < s, \end{cases}$$

with $\Gamma(x(\bar{q})) = \Gamma(x(\underline{q})) = s$.

Figure 4 depicts the equilibrium. The variable x, as defined in Proposition 2, lies in the vertical axis. The return to search Γ and the search cost s lie on the left horizontal axis. The search probability q lies on the right horizontal axis. The return to search is non-monotonic in x. It is equal to 0 when criminals do not search, then increases until it reaches $x_{\max}$, and finally decreases to 0 when x is equal to one. The return to search depends on the difference between the mean protection investment and the mean of the min of two investments picked randomly. Thus the return to search increases with protection investment dispersion. Proposition 1, part (iii), shows that when criminals search more, this affects such dispersion in two ways: the max of the distribution increases, but the distribution becomes more concentrated at its top. The former effect dominates at low levels of search efforts, whereas it may be dominated at larger ones. The effects cancel each other when x reaches $x_{\max}$. Thus, the return to search may decrease with search effort when $x(1) > x_{\max}$.

As the return to search is 0 when criminals do not search, there will always be the pure-strategy equilibrium $q^* = 0$. There is also a mixed-strategy equilibrium with $q^* \in (0,1)$ when the return to search is equal to the expected search cost, i.e. $\Gamma(x(q^*)) = s$. Finally, the pure-strategy equilibrium $q^* = 1$ obtains when $\Gamma(x(1)) \geq s$. In Figure 4, parameters γ, c, s and V have been chosen so that $\Gamma(x(1)) > s$. Thus there are three equilibria: the two pure-strategy equilibria 0 and 1, and a mixed-strategy equilibrium $\underline{q} \in (0, 1)$.

Pure-strategy equilibria are stable in the usual meaning. Thus, the equilibrium without comparison is robust to individual deviations, and the equilibrium where all criminals compare is also robust to such deviations when it exists. When $1 \in E_{IC}$, the mixed-strategy equilibrium $\underline{q} \in (0, 1)$ is unstable. When $\underline{q}$ and $\bar{q}$ belong to E_{IC}, $0 < \underline{q} < \bar{q} < 1$, the mixed-strategy equilibrium $\underline{q} \in (0, 1)$ is unstable, whereas $\bar{q} \in (0, 1)$ is stable.

Let $x(q) \equiv x(q, c)$ to highlight the dependence with respect to c. In the remaining, we make two additional assumptions. First, we only focus on stable equilibria. Thus we omit unstable equilibria from the equilibrium set E_{IC}. Second, we assume that $x(1, C/K) < x_{\max}$. This ensures that the return to search is below its maximum, i.e. $\Gamma(x(.)) < \Gamma(x_{\max})$. Thus a marginal increase in x coincides with a marginal increase in the return to search. This assumption is very natural in a context where property crime remains a rare event, so that c is small. Hence $x(1, c)$ is also small, whereas $x_{\max} > 0.7$.

Properties of EIC.—We examine how equilibrium outcomes vary with the crime rate c, and with wealth V. We denote $\Gamma(x, V)$ to highlight the dependence vis-à-vis V.

Proposition 3 (The effects of c and V) *Let $x(1, C/K) < x_{\max}$, $c_0 \in (0, C/K)$, and $\Gamma(x(1, C/K), V_0) > s$. The following properties hold:*

(i) (Effects of c) For $V = V_0$, $E_{IM} = \begin{cases} \{0\} \text{ if } c < \gamma\bar{x}/2 \\ \{0, 1\} \text{ else} \end{cases}$

(ii) (Effects of V) For $c = c_0$, $E_{IM} = \begin{cases} \{0\} \text{ if } \Gamma(x(1, c_0), V) < s \\ \{0, 1\} \text{ else} \end{cases}$

with $\Gamma(\bar{x}, V_0) = s$.

Figure 4 also depicts the effects of c and V. Both wealth and the crime rate favor the occurrence of search equilibria, i.e. with $q^* = 1$. Both c and V increase the return to search, but through different channels. The expected number of times households are compared to their neighbors increases with the crime rate. This affects the equilibrium distribution of investment: its support widens and there is more weight at its top. In so far as $x(1)$ remains lower than $x_{\max}$, the former effect dominates the latter. Thus the return to search increases. As for wealth, the gain that criminals expect from saving on protection damage is proportional to the property value V. Thus the return to search increases as well.

We now briefly discuss the parametric assumptions. That $x(1, C/K) < x_{\max}$ has already been discussed. If it does not hold, we may have additional types of stable EIC with $q^* = \bar{q} \in (0,1)$. Then, an increase in c may sometimes imply that the set E_{IC} goes from $\{0,1\}$ to $\{0,\bar{q}\}$. Conversely, an increase in V may sometimes imply that E_{IC} goes from $\{0,\bar{q}\}$ to $\{0,1\}$. The next assumption involves fixing c to let V vary. The final assumption consists in fixing V to let c vary. Wealth V must be sufficiently large, so that there exists $c < C/K$ such that $1 \in E_{IM}$. If not, then $E_{IC} = \{0\}$ for all c.

2.3 Equilibrium with mobile criminals

Criminals' location decisions.—Criminals choose a neighborhood based on their expected payoff. They do so on the basis of their expectations about the crime rate, the search effort made by the other criminals, and households' protection investments.

Definition 2 An *equilibrium with mobile criminals* (EMC) is a pair of search efforts (q_A^*, q_B^*), a collection of cdf $\{H_A, H_B\}$ and an allocation of criminals across neighborhoods (C_A^*, C_B^*) such that, for $j = A, B$,

(i) $\theta \in \Theta_j$ iff $\theta \in \arg\max_{\theta' \geq 0} W_j(\theta', q_j^*)$;

(ii) $q_j^* \in \arg\max_{q \in [0,1]} \Omega_j(q, H(., q_j^*, C_j^*/K_j))$;

(iii) $C_i^* > 0$ iff $i \in \arg\max_{j \in \{A,B\}} \Omega_j(q_j^*, H(., q_j^*, C_j^*/K_j))$;

(iv) $C_A^* + C_B^* = C$.

An EMC is composed of two EIC (parts (i) and (ii)) with the additional requirement that the number of criminals by location is endogenous. Thus part (iii) requires that criminals enter neighborhoods with the highest payoffs. Finally, part (iv) states that the total supply of criminals is fixed.

In neighborhood j, the search effort for a given number of criminals is $q^*(C_j/K_j)$. Thus, the neighborhood-j specific return to crime is

$$\Omega_j = \alpha V_j - \mathbb{E}(\theta \mid q^*(C_j/K_j), C_j/K_j) V_j + q^*(C_j/K_j)(\Gamma(x(q^*(C_j/K_j), C_j/K_j), V_j) - s). \quad (11)$$

Proposition 4 (Existence of EMC) *Assume $x(1, C/K) < x_{\max}$, and let E_{MC} be the set of EMC with typical element $\{(q_A^*, C_A^*), (q_B^*, C_B^*)\}$. The following properties hold:*

(i) There always exists an EMC, i.e. $E_{MC} \neq \emptyset$;

(ii) If $\{(1, C_A^*), (1, C_B^*)\} \in E_{MC}$, then $C_A^*/K_A > C_B^*/K_B$;

(iii) If $\{(q_A^*, C_A^*), (q_B^*, C_B^*)\} \in E_{MC}$ with $C_A^*/K_A < C_B^*/K_B$, then $q_A^* = 1$ and $q_B^* = 0$.

There always exists an EMC where all criminals enter the rich neighborhood, residents do not protect themselves, and criminals do not search. Indeed, Proposition 2 shows that there always exists an EIC with no search and no protection. All criminals enter the rich neighborhood in this case, because our model does not feature any other source of congestion. In line with the phenomenon of multiple victimization, a household can be broken into as many times as they are visited by criminals. In the particular case where the neighborhoods are equally rich, the allocation of criminals across neighborhood is indeterminate.

There may exist other types of EMC where criminals enter in both neighborhoods and search in at least one of them. The reason why all criminals do not enter the same neighborhood is because location choices convey a congestion externality. Congestion operates through protection investments. As explained below Proposition 1, the distribution of protection investments widens with the crime rate. This reduces the return to crime and discourages further entry.

Parts (ii) and (iii) describe two key properties of such equilibria. Part (ii) states that if criminals make the search effort in both locations, then the crime rate is larger in the rich neighborhood. Part (iii) completes the scenario: the poor crime rate exceeds the rich one only if criminals make the search effort in the rich neighborhood, whereas they do not compare potential victims in the poor one.

Crime, wealth and protection.—Figure 5 illustrates the case where the poor neighborhood B features a higher crime rate than the rich neighborhood A. When few criminals operate in a neighborhood, Proposition 3 states that in the unique EIC criminals do not search and households do not invest in protection. With no search, crime has constant returns. This explains the horizontal part of the return to crime Ω_j in each neighborhood, with $\Omega_A > \Omega_B$. When the number of criminals is sufficiently large, Proposition 3 also states that another EIC occurs where criminals search. Then, crime has decreasing returns, as more criminals stimulate protection investments.

Figure 5 displays two EMC. The no-search equilibrium where all criminals locate in neighborhood A, and an equilibrium where criminals are more numerous in the rich neighborhood and pay the search cost there, whereas criminals do not compare potential victims in the poor neighborhood.

2.4 Within-neighborhood heterogeneity

Our model can predict that the rich neighborhood invests so much in protection that the crime rate is higher in the poor neighborhood than in the rich one. In this sub-section, we focus on within-neighborhood heterogeneity.

Household heterogeneity.—There are a fraction μ of poor households and the remaining fraction $1-\mu$ of rich households. The former have wealth V^P, whereas the latter have V^R, with $V^R = (1+\delta)V^P$, $\delta > 0$. Wealth heterogeneity has two implications: households' behavior is likely to depend on wealth, and criminals' search is likely to be affected as well.

The expected number of visits by double-option criminals now depends on type, and so $\eta_2 \equiv \eta_2^i$. Consider a type-i household of protection θ compared with a household of similar type with protection θ'. The former is chosen by the criminal if and only if $\theta < \theta'$. This event occurs with probability $\Pr(\theta' > \theta \mid i) = 1 - H^i(\theta)$. Now, when compared to a household of a different type, they are chosen if and only if $V^i(\alpha - \theta) > V^{-i}(\alpha - \theta')$. This event occurs with probability $\Pr(\theta' > \theta V^i/V^{-i} - \alpha(V^i - V^{-i})/V^{-i} \mid -i) = 1 - H^{-i}[\theta V^i/V^{-i} - \alpha(V^i - V^{-i})/V^{-i}]$. It follows that

$$\eta_2^R(\theta) = 2c_2\left[1 - (1-\mu)H^R(\theta) - \mu H^P\left(\theta(1+\delta) - \alpha\delta\right)\right], \tag{12}$$

$$\eta_2^P(\theta) = 2c_2\left[1 - (1-\mu)H^R\left(\theta/(1+\delta) + \alpha\delta/(1+\delta)\right) - \mu H^P(\theta)\right]. \tag{13}$$

Distribution of protection investment.—As in the homogenous case, households play mixed strategies. Two cases may arise (described in Appendix D). In the *separating* case, we always have $\theta < \alpha\delta/(1+\delta)$ for all $\theta \in \Theta^R$. Thus the rich are always preferred to the poor regardless of their protection investment. The distribution of protection investment is contingent on type. It has the same qualitative properties as described in Proposition 1.

In the *pooling* case, poor agents who do not protect much are preferred to rich households who invest a lot in protection. The distribution of protection investment accounts for such cases where rich and poor compete in repelling criminals.

The key properties of the distribution of protection investment are as follows. First, in both the separating and the pooling cases, the support of the distribution widens with the crime rate and with the search effort. This property is similar to the homogenous case, meaning that criminals' search and criminals' entry increase the dispersion of protection investment. Second, in both cases, rich households are more victimized. With mixed strategies, some of the rich may be less victimized than some of the poor. However, the probability of being visited by a criminal is larger on average for the rich than for the poor.

Criminals' search effort.—Criminals' expected payoff is now

$$\Omega = \mathbb{E}(\alpha V - \theta V) + q\Gamma - sq, \tag{14}$$

where the return to search is

$$\Gamma = \mathbb{E}[\max\{\alpha V - \theta V, \alpha V' - \theta' V'\}] - \mathbb{E}(\alpha V - \theta V). \tag{15}$$

In the separating case, we have

$$\begin{aligned}\Gamma = &(1-\mu)^2 \left[\mathbb{E}(\theta^R) - \mathbb{E}(\min\{\theta^R, \theta^{R\prime}\})\right] V^R + \mu^2 \left[\mathbb{E}(\theta^P) - \mathbb{E}(\min\{\theta^P, \theta^{P\prime}\})\right] V^P \\ &+\mu(1-\mu)[(\alpha - \mathbb{E}(\theta^R))V^R - (\alpha - \mathbb{E}(\theta^P))V^P], \end{aligned} \tag{16}$$

where $\theta^i \in \Theta^i$, $i = R, P$.

The return to search has three components. With probability $(1-\mu)^2$, criminals sample two rich households and exploit (endogenous) dispersion in protection investment. Similarly, with probability μ^2 criminals sample two poor households and, here again, they benefit from choosing the least protected one. The third component is due to (exogenous) wealth dispersion. With probability $2\mu(1-\mu)$ criminals pick a rich and a poor household. In a separating equilibria they choose the rich one. In half of such cases, they would have picked a rich household anyway. The corresponding gain is zero. In the remaining half, they gain the expected difference $(\alpha - \mathbb{E}(\theta_R))V^R -$ $(\alpha - \mathbb{E}(\theta_P))V^P$.

In the pooling case, we have

$$\Gamma = (1-\mu)^2 \left[\mathbb{E}(\theta^R) - \mathbb{E}(\min\{\theta^R, \theta^{R\prime}\})\right] V^R + \mu^2 \left[\mathbb{E}(\theta^P) - \mathbb{E}(\min\{\theta^P, \theta^{P\prime}\})\right] V^P$$
$$+2\mu(1-\mu)[(\mathbb{E}(\max\{(\alpha-\theta^R)V^R, (\alpha-\theta^P)V^P\}) - (1-\mu)(\alpha - \mathbb{E}(\theta^R))V^R - \mu(\alpha - \mathbb{E}(\theta^P))V^P]. \tag{17}$$

The first line is similar to (16). The second line differs because, in a pooling equilibrium, criminals do not always prefer rich households. They only prefer them on average: sampling among rich households is more attractive, but some draws can be very unlucky as the household is so well protected that a poor household with little protection would have been better.

In both cases, consideration of wealth heterogeneity does not change the mechanism of strategic complementarity highlighted in the homogenous case. When criminals search more, the mean and dispersion of protection investments increase. This further raises the return to search. Wealth inequality also increases the return to search. More unequal neighborhoods should also be characterized by more protected residents, and more dispersed protection investments.

Numerical example.—We conclude this sub-section by a parameterization that broadly replicates the main features of the Canadian income-victimization nexus as depicted by Figure 1. Parameters are set as follows: $\alpha = 10\%$, $V = 1.0$, $s = 8.3\%$, $\gamma = 15\%$, $\delta = 35\%$, $K_A = K_B = 1.0$, $C = 6.6\%$, $\mu_A = 15\%$ and $\mu_B = 85\%$. Rich and poor are equally numbered. The rich are 35% richer than the poor; neighborhood A is mostly composed of rich households, whereas neighborhood B contains a majority of poor households. This parameterization admits an EMC with a pooling equilibrium in the protection subgame played by households in the rich neighborhood. This EMC is completely described in the Web Appendix. Figure 6 shows resulting mean victimization by household and neighborhood income. The mean crime rate is $c = 3.3\%$. In neighborhood A, the crime rate is $c_A^* = 2.9\%$ and the search effort is $q_A^* = 1$. Criminals search and both poor and rich households invest in protection. The rich invest more in protection, and those who invest the most are less attractive than the poor households who do not invest. In this neighborhood, the mean victimization probability of the rich households is 3.2%, whereas the mean victimization probability of the poor is 0.7%. In neighborhood B, $c_B^* = 3.8\%$ and $q_B^* = 0$. Criminals do not search and households do not invest in protection. The rich and the poor are equally likely to be victimized. Overall, victimization is 34% larger in the poor neighborhood than in the rich one, but the rich households are as victimized as the poor.

3 Property crime and protection: Canadian evidence

We test the microeconomic mechanisms of our model on Canadian data. More particularly, we estimate the social multiplier and disentangle the roles played by individual and neighborhood income in the protection game played by Canadian households. We linearize the structural equation defining individual investment in protection, and we estimate it with a specific IV procedure.

3.1 Data

We use cross-sectional individual data from the victimization survey of the Canadian GSS. This survey is conducted every five years, or cycles. We consider years 1999, 2004, and 2009, which correspond to cycles 13, 18, and 23. We supplement the GSS with the Census to compute the mean neighborhood income and neighborhood poor proportion. Years do not exactly match. We consider the following Census years: 1996, 2001, and 2006.

The GSS is representative at CMA level. Each individual is located by a six-digit postal code. The first three digits define Forward Sortation Areas (FSAs), our neighborhoods. FSA is not a common geographic unit in the Census and so we use a table mapping Dissemination Areas in the Census with FSAs. FSAs are exceptionnally shared by several CMAs. When this happens, the FSA is divided into several sub-neighborhoods belonging to different CMAs.

The survey samples dwellings. The respondent may be anyone belonging to the household. We only keep the following types of households: singles or couples with or without children. This excludes less traditional households with additional relatives and non relatives at home. Such households lead to never-ending questions as to the identity of wage-earners and the respondent's statute in the household. We also exclude observations where the respondent is a child. Household income is not well reported in such cases, and time use is missing for the parents.

We exclude neighborhoods with less than 10 observations in a given wave, and CMAs with a single FSA. This leaves us with about 48 CMAs, 442 FSAs, and 18 observations/FSA by wave. These numbers do not change much across waves. Over the three waves, there are 681 different FSAs, of which 249 are represented once, 236 twice, and 202 three times.

We use two variables of protection. In the survey, respondents declare whether they have ever installed an alarm and if they have ever installed locks or bars. The variable PR proxies the magnitude of household investment. It is equal to 0 when neither bars / locks nor an alarm have ever been installed. It is equal to 1 when only bars / locks have been installed, to 2 when only an alarm, and to 3 when both have been installed. The quantitative values taken by this variable are somewhat arbitrary. Thus the other variable $PA = 1$ when there is an alarm and $PA = 0$ otherwise. For robustness purposes, we also consider the dummy variable $PA2$, which takes the value one when the household has installed an alarm or bars / locks, and $PA3$, which takes the value one when the household has an alarm, bars / locks and the respondent holds a gun. Tables A1 and A2 in Appendix **??** show that all these variables similarly vary across household and neighborhood income.

The main variable of victimization BE takes the value one when the household experienced a successful or attempted break-in over the past 12 months before the interview date. Here again, we have alternative measures. The dummy $BE2 = 1$ when a break-in occurred, or something was destroyed in the property like a window. The dummy $BE3 = 1$ when $BE2 = 1$ or something was stolen in the property but outside the house. Tables A1 and A2 in Appendix **??** show that the stylized facts highlighted in the Introduction also hold when BE is replaced by $BE2$ and $BE3$. Mean neighborhood protection $\overline{PA}_{jt}$ and mean victimization $\overline{BE}_{jt}$ are computed by (weighted) aggregation of household observations.

The household characteristics are as follows. A set of dummy variables describes the household type: couple, single woman, single man. Then come the number of children, the household age (defined as the mean age in five-year classes of the leading members of the household) and education (defined as the highest level of education of the leading members). We include a dummy equal to 1 when the leading members of the household are all occupied full time (because they follow an educational program or because they work). We also consider individual lifetime victimization. This includes the events presented above (break-in, property damage, theft outside property), but also events like assault, sexual or moral harassment, rape, etc. We exclude the events presented above and that occurred during the previous 12 months. We finally include the following dwelling characteristics: three dummy variables describing the dwelling type (detached house, semi-detached and apartment), the age of the household in the dwelling, and a dummy equal to one if the household owns the dwelling.

Household income is declared in 10 classes. We attribute the class mean to each income class. We then divide by the contemporaneous CMA average and compute the log of this ratio. Thus household income is specific to each CMA. This is in line with the model where criminals are mobile within CMAs but not across them.

As for neighborhood characteristics, we compute the mean of household characteristics. We also use two variables from the Census: the log of the ratio of neighborhood income to CMA income, and the low-income proportion. These variables are predicted by the theoretical model.

As for $\bar{y}_{-j}$, μ_{-j} and $\mathbb{E}_{-j}(PA)$, we have several possibilities. The GSS does not sample individuals from all possible neighborhoods of a given CMA. We limit our proxies to neighborhoods actually contained in the survey. For each neighborhood, we consider the variables of the most populated neighborhood after the one of interest. Alternatively, we average the different variables over the remaining neighborhoods belonging to the same CMA.

Table 1 provides descriptive statistics on the final sample, whereas Table 2 displays the matrix of correlation coefficients between variables aggregated at FSA level.

3.2 Empirical model

A linear model.—We start from the model version developed in Section xx where people differ in wealth within each neighborhood. Having $PA_i = 1$ means that the individual invested in protection above some threshold, say $\theta > 0$. The model predicts

$$\Pr[PA_i = 1 \mid y_i] = \Pr[\theta_i \geq \theta] = 1 - H(\theta | (C/K)_{j(i)}, q_{j(i)}, \mu_{j(i)}), \tag{18}$$

where $j(i)$ denotes the neighborhood where i resides, H is the cdf of protection investment and μ_j is the proportion of poor. A linear version gives:

$$\Pr[PA_i = 1 | y_i] = a_0 + a_1 y_i + a_2 (C/K)_{j(i)} + a_3 \mu_{j(i)} + q_{j(i)} + u_i, \tag{19}$$

Individual income affects the probability of having an alarm, because the distribution of protection investment differs across income groups. This effect is positive, even though we can find cases where rich people invest less than poor ones.

In our data the crime-to-resident ratio $(C/K)_j$ is observed, whereas the criminals' effort q_j is not. Thus we must find an alternative observed variable to the unobserved q. In an EIC, the criminals' effort depends on the distribution of protection investment, on neighborhood mean wealth $\bar{y}_j$, and on μ_j. We assume that the functional form linking q to its determinants is linear:

$$q_j = b_0 + b_1\bar{y}_j + b_2\mu_j + b_3\mathbb{E}_j(PA). \tag{20}$$

Replacing q_j by its value, we obtain

$$\Pr[PA_i = 1|y_i] = c_0 + c_1 y_i + c_2(C/K)_{j(i)} + c_3\bar{y}_{j(i)} + c_4\mu_{j(i)} + c_5\mathbb{E}_{j(i)}(PA) + u_i. \tag{21}$$

The new set of regressors features the proportion of households who have an alarm $\mathbb{E}_{j(i)}(PA)$. This term reflects the rat race to protection induced by criminals' effort to find the most attractive house. Parameter c_5 measures the magnitude of neighbors' social influence through their protection investment. The new set of regressors also includes the proportion of poor μ_j. This variable was already among the regressors. Thus $c_4 = a_3 + b_2$ and a_3 and b_2 cannot be identified. This is not a problem for the theory, but we must bear it in mind when interpreting empirical estimates of parameter c_4.

In an EMC, there is a positive number of criminals in neighborhood j if and only if the return to crime in neighborhood j is larger than or equal to the overall return to crime in the rest of the city $-j$:

$$\Omega_j^* \equiv \Omega(\mathbb{E}_j(PA))y_j \geq \Omega(\mathbb{E}_{-j}(PA))y_{-j} \equiv \Omega_{-j}^*. \tag{22}$$

Imposing the equilibrium condition whereby $\mathbb{E}_j(PA)$ is a function of q_j, C/K, $\overline{y}_j$, and, possibly, μ_j, we obtain

$$\Omega_j^* \equiv \Omega(q_j, C_j/K_j, \overline{y}_j, \mu_j) \geq \Omega(q_{-j}, (C - C_j)/K_{-j}, \overline{y}_{-j}, \mu_{-j}) \equiv \Omega_{-j}^*. \tag{23}$$

Using the fact that q_j increases with $\bar{y}_j$, μ_j and $\mathbb{E}_j(PA)$, C_j/K_j is a function of $\mathbb{E}_j(PA)$, $\bar{y}_j$, μ_j, $\mathbb{E}_{-j}(PA)$, $\bar{y}_{-j}$, and μ_{-j}. We assume a linear functional form:

$$(C/K)_j = d_0 + d_1\bar{y}_j + d_2\bar{y}_{-j} + d_3\mu_j + d_4\mu_{-j} + d_5\mathbb{E}_j(PA) + d_6\mathbb{E}_{-j}(PA) + v_j. \tag{24}$$

Empirical model.—We adapt equations (21) and (24) to our data and we explain how we estimate the resulting model. There are three important aspects that are missing from the theoretical model. First, the model features two neighborhoods and a unique geographic area over which criminals are mobile. In practice, the Canadian territory is divided into a partition of Census Metropolitan Areas (CMAs), themselves divided into numerous neighborhoods. Second, even if households are observed once, they are observed at different dates corresponding to different survey waves. We index survey waves by t. Third, households not only differ in income but also with respect to other characteristics that may affect victimization and the incentive to install an alarm. Let x_i denote the vector of household characteristics except income.

For survey wave t, the neighborhood-j specific victimization rate $\overline{BE}_{jt}$ proxies C/K_j, whereas the household share who have an alarm $\overline{PA}_{jt}$ proxies $\mathbb{E}_j(PA)$. The empirical model is

$$PA_i = c_0 + c_1 y_i + c_{1n} x_i + c_2 \overline{BE}_{j(i)t(i)} + c_3 \bar{y}_{j(i)t(i)} + c_4 \mu_{j(i)t(i)} + c_5 \overline{PA}_{j(i)t(i)} + u_i, \quad (25)$$

$$\begin{aligned}\overline{BE}_{jt} = {} & d_0 + d_1 \bar{y}_{jt} + d_{1n} \bar{x}_{jt} + d_2 \bar{y}_{-jt} + d_{2n} \bar{x}_{-jt} + d_3 \mu_{jt} + d_4 \mu_{-jt} \\ & + d_5 \overline{PA}_{jt} + d_6 \overline{PA}_{-jt} + v_{jt}. \end{aligned} \quad (26)$$

Estimation techniques.—The model has two particularities. On the one hand, mean protection belongs to the set of regressors in equation (25). On the other hand, the neighborhood victimization probability is endogenous and responds to neighborhood characteristics differences. Such characteristics include neighborhood mean protection as well as mean protection in alternative neighborhoods of the same city. These two particularities motivate a specific IV strategy.

Equation (25) is a linear-in-means model of social interactions. Taking the wave-t neighborhood-j mean of equation (25) and then solving in $\overline{PA}_{jt}$, we obtain

$$\overline{PA}_{jt} = \frac{c_0 + c_1 \bar{\bar{y}}_{jt} + c_{1n} \bar{\bar{x}}_{jt} + c_2 \overline{BE}_{jt} + c_3 \bar{y}_{jt} + c_4 \mu_{jt}}{1 - c_5}, \quad (27)$$

where a variable with two upper bars is the wave-t neighborhood-j variable mean. Thus, we can use group-specific means $\bar{\bar{x}}_{jt}$ and $\bar{\bar{y}}_{jt}$ to instrument $\overline{PA}_{j(i)t(i)}$ in equation (25).

In an EIC, the regressor $\overline{BE}_{jt}$ is exogenous and thus we can limit the set of instruments to $\bar{\bar{x}}_{jt}$ and $\bar{\bar{y}}_{jt}$. In an EMC, the crime rate is endogenous in each neighborhood. Consideration of (26) shows that $\overline{BE}_{jt}$ is affected by $\overline{PA}_{jt}$ and $\overline{PA}_{-jt}$. Neglecting such dependence leads to biasing the coefficient of $\overline{BE}_{jt}$ downward. We try to detect the positive effects that criminals have on the incentive to invest in protection. However, simultaneously, protection investments deter the criminals from entering the neighborhood.

The logic of equation (26) provides a set of instruments to $\overline{BE}_{jt}$. Criminals compare different neighborhoods before they choose one. Thus we can use the mean characteristics of alternative neighborhoods to instrument $\overline{BE}_{jt}$. More precisely, we will use $\bar{y}_{-jt}$, $\bar{x}_{-jt}$, μ_{-jt}, $\overline{\overline{x}}_{-jt}$ and $\overline{\overline{y}}_{-jt}$. We also use the CMA-specific victimization rate.

3.3 Results

Main results.—Table 3 displays the main results. It consists of six columns. The dependent variable is PR in columns a to c, and PA in columns d to f. Columns a and d display OLS estimates. In columns b and e, the only instruments are the FSA mean characteristics of the household respondents. In theory, this corresponds to the EIC case. In columns c and f, instruments also include the CMA mean victimization proportion and the mean characteristics of the respondents in the most populated neighborhood after the one of interest. This corresponds to the EMC case. Note that we only display the parameters of interest. The full estimates as well as the results of the first-step regressions can be found in the online Appendix.

OLS estimates in columns a and d feature a very strong impact of mean protection on household protection. The estimated coefficient is slightly below one. By contrast, mean victimization has a very small quantitative impact and is not significant anyway. As we progressively enrich the set of instruments, mean protection declines in impact, whereas mean victimization gains in both impact and significance. These effects were expected. Mean protection declines as we project it onto the space of FSA mean characteristics, thereby lowering the influence of individual protection. Mean victimization matters more as we purge it from the influence of protection. This eliminates the causal negative impact of neighborhood protection on criminals' incentive to enter the neighborhood.

Columns c and f display the estimates of the model whose estimation strategy was exposed in the previous section. From a qualitative perspective, they validate the key predictions of the theory. Not only do mean protection and mean victimization matter, but also household and FSA mean household incomes, as well as the FSA low-income proportion. Thus social interactions go through average protection and average income. In our model, there is strategic complementarity between households' protection investments, and this implies that individual protection increases with average protection. Quantitatively, the parameter associated to mean protection is about 0.3. Thus the social multiplier is about $1/(1-0.3) \approx 1.4$. Criminals are also more likely to search and compare potential victims when the neighborhood is richer, that is when the return to search is sufficiently high. Household investment also increases with household income. In our model this arises because richer households of a given neighborhood sample their investment from a distribution with a longer support and higher mass in the right part.

Columns c and f reveal a strong impact of income variables. Through aggregation of individual investments, the mean protection is impacted twice by neighborhood income: because of social interactions, and because of the direct effect of individual income on protection. Summing coefficients of the two income variables, we obtain about 0.2 in the PR case and 0.1 in the PA case. Now consider an increase in FSA household income by 30 percentage points, its standard error in the GSS. Accounting for the social multiplier, this increases FSA mean protection investments by $0.2 \times 30/0.7 \approx 8.5$ percentage points in the PR case and by $0.1 \times 30/0.7 \approx 4.5$ percentage points in the PA case. These figures broadly amount to 30% of the standard deviation of both variables.

Changes in mean victimization also have sizeable effects on protection investments. An increase by one-standard deviation in mean victimization causes an $1.8 \times 0.05/0.7 \approx 12.5$ percentage point increase in mean protection, and a $0.6 \times 0.05/0.7 \approx 4.5$ percentage point increase in the mean number of households who have installed an alarm.

The instrument set is statistically valid. The p-value of the Hansen test of over-identifying restrictions is largely above usual rejection thresholds. Thus we do not reject the null hypothesis of instrument exogeneity. Instruments are also sufficiently strong to explain a significant part of the variance of the two endogenous variables. The Angrist-Pischke F statistics are very large, and the Cragg-Donald statistics are well above the typical thresholds.

Robustness.—Table 4 presents different changes in the dependent variables. In columns a and c, the mean protection investment does not include the household of interest. In columns b and d, the estimation is only run over the set of nonvictimized households. This reduces the influence of the FSA mean victimization probability, but it is still positive and significant. In columns e and f, PA is modified to account for alarm and bars / locks simultaneously (PA2), and to account for alarms, bar / locks, gun and guard dog (PA3). Table 5 reproduces the basic estimates with different variables of mean FSA victimization. Victimization also includes damages to property (BE2) in columns a and c, and theft outside the property (BE3) in columns b and d. Estimates are very stable across the different specifications.

Missing variables.—We cannot exclude that a missing personal characteristic is jointly correlated with individual protection and individual income. We control for a large number of household characteristics. However, the personal time discount rate for instance is unobserved. This variable positively affects the propensity to invest in many different types of activities. This may include self-protection and human capital. Thus its absence creates a spurious correlation between income and protection. The only way to deal with such variables would be to include household fixed effects, but this strategy is not feasible because each household is observed once.

The case of missing variables at FSA level, the correlated effects, is somewhat different. Such effects are not worrying if they are implicitly included in the set of controls. Consider for instance the mapping of police forces over the different FSAs of a given CMA. It is likely that such a mapping will affect the willingness of the residents to self-protect. Thus we may think that excluding it from the regression may cause an artificial correlation between individual and FSA mean protection. However, the reason why police forces would matter is because of their supposed impact on the local crime rate. Here is the point: we control for FSA mean victimization. Of course, having police data by neighborhood would allow us to increase the set of instruments for first-step regressions.

Still, there may be local amenities shaping protection incentives. One may consider the presence of an alarm shop for instance. Despite fact that the causal effect of the shop could be disputed, it can be argued that such a shop facilitates alarm purchases. Neglecting such a variable positively biases the coefficient associated with FSA mean protection in the regression. The way to eliminate this bias consists in controlling for FSA fixed effects. Unfortunately, we only have three waves of data, and FSA covariates do not vary a lot over time.

Table 6 displays the estimates with such FSA fixed effects. As expected, most of the time-varying FSA characteristics become nonsignificant, except average protection. The parameter associated with mean protection has the same magnitude and significance as in columns c and f of Table 3. This result holds whether the computation of mean protection includes the individual observation or not. Thus there is strategic complementarity between FSA neighbors, and the impact of the mean cannot be attributed to some time-invariant correlated effect.

Residents' sorting.—In our paper, criminals are mobile whereas residents are not. This asymmetry corresponds to the realistic case where moving involves much higher costs than choosing a neighborhood to prospect potential victims. However, sorting may bias the different coefficients associated with income and FSA variables. This arises when sorting is made on the basis of characteristics that are not among the explicative variables. We now illustrate this claim with three examples.

'Rich households prefer rich neighborhoods'. Income sorting complicates the identification of household income from FSA mean income. The problem is not hard to solve as long as sorting is imperfect and we can observe rich (poor) households in poor (rich) neighborhoods. In any case, there is no other problem than identification of the different income variables.

'Rich households prefer safe neighborhoods'. This example is much less obvious than the previous one. On the one hand, property crime is not attached to a location when criminals are perfectly mobile. On the other hand, neighborhood income inequality must also be taken into account. The crime rate may be low in a given poor neighborhood, but a rich household residing there would be much more victimized than the others. However, let us assume that this type of sorting occurs. This does not cause any statistical problem because FSA mean victimization is also among the regressors.

'Rich households prefer well-protected neighborhoods'. Here again this must be discussed. Our model can predict that protected neighborhoods are less victimized than others. However, households in such neighborhoods are forced to self-protect because criminals search intensively. Thus residing there means paying the investment cost, and this may exceed the gain due to reduced victimization. Here again, that households make residential choices on the basis of neighborhood mean protection is not a problem because we include FSA mean protection among the regressors.

We cannot exclude that sorting is made on unobservables that affect household protection. However, the different scenarii we have described so far prove that it is not easy to come up with a story that does not involve any of the different explicative variables.

4 Conclusion

We develop a model of property crime and private protection allocation within cities. Criminals are mobile over different neighborhoods of a city. Once in a neighborhood, they choose whether to make a search effort to compare potential victims. If not they are randomly assigned to one of the households of the neighborhood. In turn, households choose how much they invest in private protection. Such investments reduce criminals' gains and divert their attention to less protected neighbors. Households in a rich neighborhood may enter in a rat race to protection characterized by strong incentives to self protect and to search, and a weak incentive to enter the neighborhood. When sufficiently fierce, the rat race leads criminals to prefer poorer and less protected neighborhoods.

We test this theory with the Canadian GSS. We divide the Canadian territory into a set of Census Metropolitan Areas (our cities) and Forward Sortation Areas (our neighborhoods). We regress different measures of household protection on household characteristics, including income, and FSA characteristics, such as mean income, mean protection, and mean victimization. We use 2SLS to account for endogeneity. The set of instruments is deduced from the theoretical model. We identify strong effects of social influence that transit through mean income and mean protection. Household protection is also positively impacted by household income and mean victimization.

5 References

Acemoglu, D. and R. Shimer (2000), "Wage and Technology Dispersion", *Review of Economic Studies* **67**, 587–607.

Albrecht, J., P. Gautier and S. Vroman (2006), "Equilibrium Directed Search with Multiple Applications", *Review of Economic Studies* **73**, 869–891.

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

Burdett, K. and K.L. Judd (1983), "Equilibrium price dispersion", *Econometrica* **51**, 955-970.

Burdett, K., R. Lagos and R. Wright (2003), "Crime, Inequality, and Unemployment", *American Economic Review* **93**, 1764–1777.

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

Calvó-Armengol, A., T. Verdier and Y. Zenou (2007), "Strong and Weak Ties in Employment and Crime", *Journal of Public Economics* **91**, 203–233.

Engelhardt, B., G. Rocheteau and P. Rupert (2008), "Crime and the Labor Market: A Search Model with Optimal Contracts", *Journal of Public Economics* **92**, 1876–1891.

Freeman, R.B. (1999), "The Economics of Crime", Handbook of Labor Economics. In Ashenfelter, O. and D. Card (ed.), Handbook of Labor Economics, vol. 3, ch. 52, 3529–3571, Elsevier.

Galenianos, M. and P. Kircher (2009), "Directed Search with Multiple Job Applications", *Journal of Economic Theory* **144**, 445–471.

Galenianos, M., R. Pacula and N. Persico (2012), "A Search-Theoretic Model of the Retail Market for Illicit Drugs", *Review of Economic Studies* **79**, 1239–1269.

Glaeser, E.L., B. Sacerdote and J.A. Scheinkman (1996), "Crime and Social Interactions", *Quarterly Journal of Economics* **111**, 507–48.

Hotte L. and T. Van Ypersele (2008), "Individual Protection against Property Crime: Decomposing the Effects of Protection Observability", *Canadian Journal of Economics*

41, 537–563.

Huang, C.-C., D. Laing and P. Wang (2004), "Crime and Poverty: A Search-Theoretic Approach", *International Economic Review* **45**, 909–938.

Manski, C.F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies* **60**, 531–542.

Moen, E. (1999), "Education, Ranking, and Competition for Jobs", *Journal of Labor Economics* **17**, 694–723.

Patacchini, E. and Y. Zenou (2008), "The Strength of Weak Ties in Crime", *European Economic Review* **52**, 209–236.

Shavell, S. (1991), "Individual Precautions to Prevent Theft: Private versus Socially Optimal Behavior", *International Review of Law and Economics* **11**, 123–132.

Zenou, Y. (2003), "The Spatial Aspects of Crime", *Journal of the European Economic Association* **1**, 459–467.

APPENDIX

A International Crime and Victim Survey

Figure 3 is built from the International Crime and Victim Survey. This survey is conducted in a large number of countries. It randomly samples households and the questions are addressed to one particular member of each household. Thus the sampling unit is the household. We consider OECD countries / states surveyed in 1996 and / or 2000.

We use two sets of questions. The first set is about property crime:

> "Over the past five years, did anyone actually get into your house or flat without permission and steal or try to steal something? I am not including here thefts from garages, sheds or lock-ups."

and

> "When did this happen? Was this... (1) this year; (2) last year; (3) before then; (4) don't know/can't remember"

We divide the population into two groups: households above median income and households below it. In each sub-group, we compute the proportion of persons / households who respond "yes" to the first question, and "(1) this year" to the second one. We finally compute the ratio of these proportions.

The second set consists of one question about private protection:

> "In order to help us understand why some homes are more at risk of crime than others, could I ask you a few questions about the security of your houses? Is your house protected by the following (multiple answers are allowed): 1) a burglar alarm, 2) special door locks, 3) special window or door grilles, 4) a dog that would detect a burglar, 5) a high fence, 6) a caretaker of security guard, 7) a formal neighborhood watch scheme, 8) friendly arrangements with neighbors to watch each others houses, 9) not protected by any of these, 10) respondent refuses to answer"

In each income group, we compute the share of households who declare they are protected by an alarm. We then compute the ratio of the two shares.

B Expected number of visits

Hereafter, we refer to theft options as matches. A criminal is matched with zero, one, or two households. All households draw their protection level from the continuously differentiable cdf H. The number of expected visits by single-option criminals is equal to the ratio of such criminals to households. Thus $\eta^1 = c_1$, irrespective of household's type. We now study the expected number of visits by double-option criminals η_2.

Consider a household whose protection level is θ. With probability

$$\Pi(t) = \binom{C_2}{t}\left(\frac{2}{K}\right)^t\left(1-\frac{2}{K}\right)^{C_2-t}, \tag{28}$$

the victim is matched with t double-option criminals. When the victim is compared to another household, s/he is chosen by the criminal if and only if the victim's protection level θ is lower than the other household's protection level θ'.

Let us write $(\theta_1, ...\theta_t)$ the t order of the sub-sample of households who are matched with criminals who are also matched with the victim, i.e. $\theta_1 < \theta_2 ... < \theta_t$. The victim is visited by $t-g$ criminals matched with another household if and only if $\theta \in [\theta_g, \theta_{g+1}]$. The joint density distribution (x, y) of $(g, g+1)$ order statistics is given by

$$\frac{t!}{(g-1)!(t-g-1)!} H(x)^{g-1}\left(1-H(y)\right)^{t-g-1} h(x)h(y). \tag{29}$$

Therefore the probability that $\theta \in [\theta_g, \theta_{g+1}]$ is given by

$$\begin{aligned} p(g) &= \int_\theta^{\bar{\theta}}\int_0^\theta \frac{t!}{(g-1)!(t-g-1)!} H(x)^{g-1}\left(1-H(y)\right)^{t-g-1} h(x)h(y)\ dx\ dy \\ &= \frac{t!}{g!(t-g)!} H(\theta)^g (1-H(\theta))^{t-g}. \end{aligned} \tag{30}$$

as $\int H(x)^{g-1}h(x)dx = H(x)^g/g$ and $\int (1-H(y))^{t-g-1}h(y)dy = (1-H(y))^{t-g}/(g-t)$.

Thus the expected number of visits by double-option criminals is

$$\begin{aligned} \eta_2 &= \sum_{t=0}^{C_2}\Pi(t)\sum_{g=0}^{t}(t-g)p(g) \\ &= 2c_2\left[1-H(\theta)\right]. \end{aligned} \tag{31}$$

C Proofs

Proof of Proposition 1. Part (i). We have

$$W(\theta,.) = \{1 - [c(1-q) + 2cq(1-H(\theta))](\alpha - \theta) - \gamma\theta\}V. \tag{32}$$

There cannot be a degenerate distribution of protection investment. On the contrary, suppose that the distribution is degenerate, and so the equilibrium investment is θ_0. This implies that $W(\theta,.) < W(\theta_0,.)$ for all $\theta \neq \theta_0$. Let $\theta = \theta_0 + \varepsilon > \theta_0$. We have

$$W(\theta) - W(\theta_0) = -(\gamma - c(1-q))\varepsilon V + cq(\alpha - \theta_0)V, \tag{33}$$

which is positive for ε sufficiently small. It follows that the distribution H cannot be degenerate.

In a mixed-strategy equilibrium, households must be indifferent between the different protection investments. The cdf $H(\theta)$ is obtained by solving $W^i(\theta,.) = W^i(0,.)$. This gives equation (6). At the upper bound of the support, $H(\bar{\theta}) = 1$. This gives $\bar{\theta}$.

Part (ii). We have $\bar{\theta} = \frac{2cq\alpha}{\gamma - c(1-q)}$. It is strictly increasing in c. As for q, $\partial\bar{\theta}/\partial q$ has the sign of $\gamma - c > \gamma/2 - c > 0$ by assumption.

Part (iii). Both derivatives are easy to compute.

Proof of Proposition 2. The function $\Gamma : [0,1] \to \mathbb{R}$ has the following properties: $\Gamma(0) = \Gamma(1) = 0$, $\Gamma(x) > 0$ for all $x \in (0,1)$, and $\Gamma'(x) \gtreqless 0$ iff $x \lesseqgtr x_{\max}$.

We have $1 \in E_{IM}$ if $\Gamma(x(1)) \geq s$, $\underline{q} \in E_{IM}$, $\underline{q} \in (0,1)$, if $\Gamma(x(\underline{q})) = s$, and $0 \in E_{IM}$ if $\Gamma(x(0)) < s$.

Part (i). Figure 4 represents the return to search Γ as a function of x. We already know that x strictly increases with q from 0 to $x(1) < 1$. When $x(1) < x_{\max}$, the return to search strictly increases with q. The result follows: if $\Gamma(x(1)) < s$, then $\Gamma(x(q)) < s$ for all q and $E_{IM} = \{0\}$. If $\Gamma(x(1)) > s$, there is a unique $\underline{q} \in (0,1)$ such that $\Gamma(x(\underline{q})) = s$, and we have $E_{IM} = \{0, \underline{q}, 1\}$.

Part (ii). When $x(1) > x_{\max}$, the reasoning is very similar. However, there is something new because we may have $\Gamma(x_{\max}) > s$ and $\Gamma(x(1)) < s$. In this case, there are $(\underline{q}, \bar{q})$, $0 < \underline{q} < \bar{q} < 1$, such that $\Gamma(x(\underline{q})) = \Gamma(x(\bar{q})) = s$.

Proof of Proposition 3. By assumption, $x(1,c) < x_{\max}$ for all $c \leq C/K$. Proposition 1 implies that

$$E_{IM} = \begin{cases} \{0\} \text{ if } \Gamma(x(1,c),V) < s, \\ \{0,1\} \text{ if } \Gamma(x(1,c),V) > s. \end{cases}$$

Part (i). Given the definition of V_0, and given that Γ strictly increases with x, there is a unique $\bar{x} < x(1,C/K)$ such that $\Gamma(\bar{x},V) = s$. When $c < \gamma\bar{x}/2$, $x(1,c) < \bar{x}$ and $\Gamma(x(q,c),V) < s$ for all $q \in [0,1]$. Thus $E_{IM} = \{0\}$. When $c > \gamma\bar{x}/2$, we have $\Gamma(x(1,c),V) > s$ and thus $E_{IM} = \{0,1\}$.

Part (ii). When $\Gamma(x(1,c_0),V) < s$, we have $\Gamma(x(q,c_0),V) < s$ for all q and so $E_{IM} = \{0\}$. When $\Gamma(x(1,c_0),V) > s$, $E_{IM} = \{0,1\}$.

Proof of Proposition 4. Following Proposition 2 and the focus on stable EIM, $x(1,C/K) < x_{\max}$ implies that $q_j^* \in \{0,1\}$, $j = A,B$.

Part (i). $\{(0,C),(0,0)\} \in E_{MC}$. We already know that $0 \in E_{IM}$ and so requirements (i) and (ii) of Definition 2 are satisfied. Moreover, we have $\Omega_A^* = \alpha V_A > \Omega_B^* = \alpha V_B$ and condition (iii) is satisfied. Finally, condition (iv) trivially holds.

Part (ii). Suppose on the contrary that $C_A^*/K_A \leq C_B^*/K_B$. Then $\Omega_A^* > \Omega_B^*$, which violates condition (iii) of Definition 2.

Part (iii). Suppose on the contrary that $q_A^* \leq q_B^*$. Then $\Omega_A^* > \Omega_B^*$, which, again, violates condition (iii) of Definition 2.

D Distribution of protection investment with household heterogeneity

We refer to the separating case as a separating equilibrium of the protection subgame played by the residents of a neighborhood. Similarly, we refer to the pooling case as a pooling equilibrium of this game.

Proposition A1 (Separating and pooling equilibria) *Let $q > 0$ be given and $\bar{\delta} = \frac{x(q)}{1-x(q)}(1-\mu)$. The following properties hold:*

(i) If $\delta > \bar{\delta}$, there is a separating equilibrium such that

$$H^i(\theta) = \frac{\theta}{\alpha - \theta}\frac{\alpha - \bar{\theta}^i}{\bar{\theta}^i}, \tag{34}$$

with $\bar{\theta}^R = \alpha\frac{x(q)}{1-\mu x(q)}$ *and* $\bar{\theta}^P = \mu\alpha x(q)$*;*

(ii) If $\delta \leq \bar{\delta}$*, there is a pooling equilibrium such that*

$$H^R(\theta) = \begin{cases} \frac{\theta}{\alpha-\theta}\frac{1-x(q)}{(1-\mu)x(q)} \equiv H_\ell^R(\theta) \text{ if } \theta \in \left[0, \frac{\alpha\delta}{1+\delta}\right) \\ \frac{\theta}{\alpha-\theta}\frac{1-x(q)}{(1-\mu)x(q)}\frac{\alpha\delta\mu(1-\mu x(q))+\theta(1-\mu)(1-x(q)(1+\delta)\mu}{1+\mu\delta-x(q)(1+\delta)\mu} \equiv H_u^R(\theta) \text{ if } \theta \in \left[\frac{\alpha\delta}{1+\delta}, \bar{\theta}^R\right] \end{cases}, \tag{35}$$

$$H^P(\theta) = \begin{cases} \frac{\theta}{\alpha-\theta}\frac{1-x(q)}{(1-\mu)x(q)}\frac{(1+\delta)(1-x(q)\mu)}{1+\mu\delta-x(q)(1+\delta)\mu} \equiv H_\ell^P(\theta) \text{ if } \theta \in \left[0, \bar{\theta}^R(1+\delta)-\alpha\delta\right) \\ \frac{\alpha(1+\delta)}{\mu x(q)}\frac{1-x(q)}{\alpha-\theta} + \frac{\mu x(q)-1}{\mu x(q)} \equiv H_u^P(\theta) \text{ if } \theta \in \left[\bar{\theta}^R(1+\delta)-\alpha\delta, \bar{\theta}^P\right] \end{cases} \tag{36}$$

with $\bar{\theta}^R = \alpha x(q) - \alpha(1-x(q))\delta\frac{2-\mu(1+x(q))}{(1-\mu)(1-\mu x(q))}$ *and* $\bar{\theta}^P = \alpha(x(q) - (1-\delta x(q))$*.*

Proof. Part (i). Suppose that there is a separating equilibrium, i.e. $V^R(\alpha - \bar{\theta}^R) > \alpha V^P$, which is equivalent to $\bar{\theta}^R < \alpha\delta/(1+\delta)$. Then, we follow Proposition 1 and we obtain (34). We finally check that this is indeed a separating equilibrium when $\delta > \bar{\delta}$.

Part (ii). Suppose that there is a pooling equilibrium with $\bar{\theta}^P > \bar{\theta}^R(1+\delta) - \alpha\delta$. This means that $\bar{\theta}^R$ is such that $V^R(\alpha - \bar{\theta}^R) < \alpha V^P$, which is equivalent to $\bar{\theta}^R > \alpha\delta/(1+\delta)$. Expected household payoffs are as follows:

$$W^R(\theta) = \begin{cases} \left[1 - c_1(\alpha-\theta) - 2c_2(\alpha-\theta)\left[1-(1-\mu)H_\ell^R(\theta)\right] - \gamma\theta\right] V^R \text{ if } \theta \in \left[0, \frac{\alpha\delta}{1+\delta}\right) \\ \left[1 - c_1(\alpha-\theta) - 2c_2(\alpha-\theta)\left[1-(1-\mu)H_u^R(\theta) - \mu H_\ell^P((1+\delta)\theta-\alpha\delta)\right] - \gamma\theta\right] V^R \text{ if } \theta \in \left[\frac{\alpha\delta}{1+\delta}, \bar{\theta}^R\right] \end{cases} \tag{37}$$

$$W^P(\theta) = \begin{cases} \left[1 - c_1(\alpha-\theta) - 2c_2(\alpha-\theta)\left[1-(1-\mu)H_u^R\left(\frac{\theta+\alpha\delta}{1+\delta}\right) - \mu H_l^P(\theta)\right] - \gamma\theta\right] V^P \text{ if } \theta \in \left[0, \bar{\theta}^R(1+\delta)-\alpha\delta\right) \\ \left[1 - c_1(\alpha-\theta) - 2c_2(\alpha-\theta)\mu\left[1-H_u^P(\theta)\right] - \gamma\theta\right] V^P \text{ if } \theta \in \left[\bar{\theta}^R(1+\delta)-\alpha\delta, \bar{\theta}^P\right] \end{cases} \tag{38}$$

The functions H^R and H^P are continuous, monotonically increasing, with $H^R(0) = 0$, $H^P(0) = 0$, $H^R(\bar{\theta}^R) = 1$ and $H^P(\bar{\theta}^P) = 1$. Thus they are two cdf. Moreover, W^R and W^P are constant over their respective support. Finally, $\bar{\theta}^R > \alpha\delta/(1+\delta)$ iff $\delta \leq \bar{\delta}$. ■

As claimed in the text, $\bar{\theta}^i$ is strictly increasing in c and q in both types of equilibrium. Moreover, the victimization probability is always larger for the rich than for the poor. This statement is obvious in the separating case where the rich are always preferred to the poor. In the pooling case, suppose that a criminal is randomly matched with a poor and a rich household. The probability that the rich household is selected is

$$H_u^R\left(\frac{\alpha\delta}{1+\delta}\right) + \int_{\frac{\alpha\delta}{1+\delta}}^{\bar{\theta}^R}\int_{(1+\delta)\bar{\theta}^R-\alpha\delta}^{\bar{\theta}^P} h^P(\theta^P)h^R(\theta^R)d\theta^R d\theta^P = \frac{1}{2} + \delta\frac{\alpha-\bar{\theta}^P}{\bar{\theta}^R} > 1/2. \tag{39}$$

E Figures and Tables

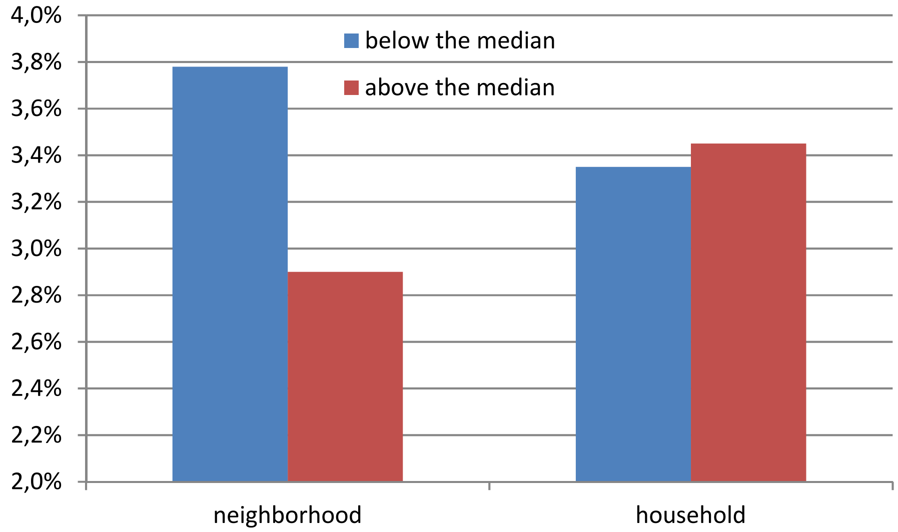

Fig.1: Property crime by household and neighborhood income in Canada. The figure displays the household share that experienced a break-in over the last 12 months. See Section 3.2 for more details. Source: Canadian GSS 1999, 2004 and 2009

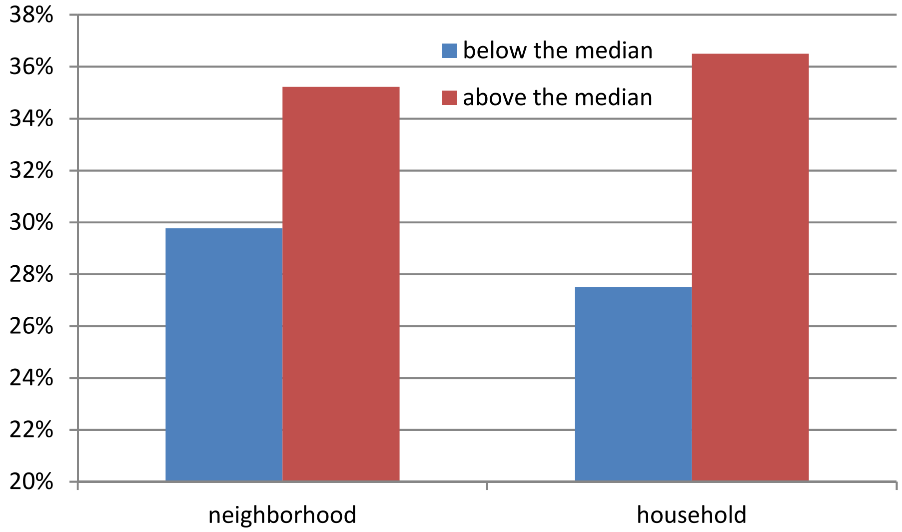

Fig.2: Household share with an alarm by household and neighborhood income. See Section 3.2 for more details. Source: Canadian GSS 1999, 2004, 2009.

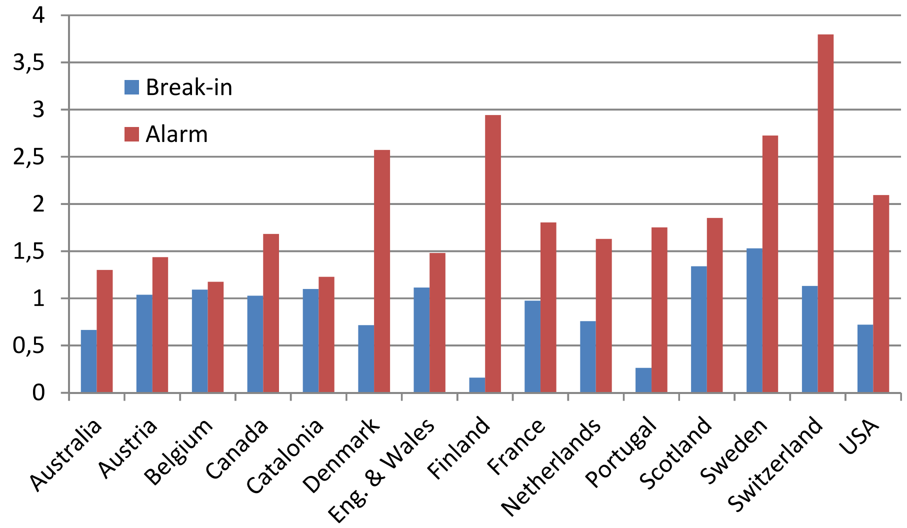

Fig.3: Relative victimization and protection for a selection of countries / states. Ratio of above-the-median-income to below-the-median-income household shares. The blue bar is the victimization probability ratio, whereas the red bar is the alarm probability ratio. Data are presented with greater detail in Appendix A. Source: International Crime and Victim Survey, 1996-2000.

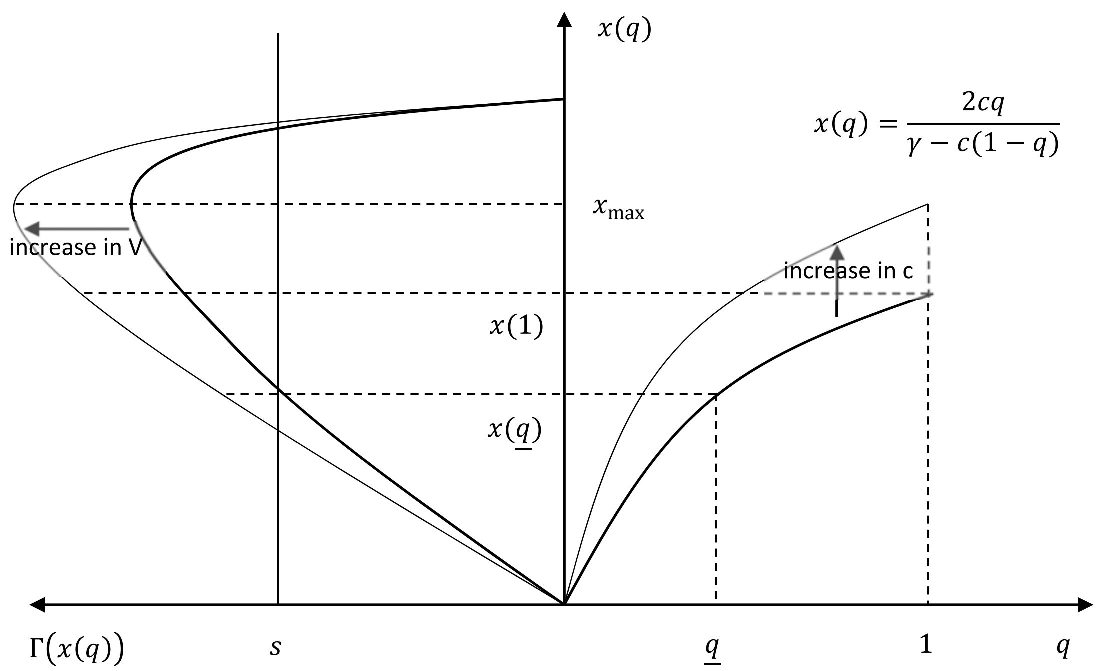

Fig.4: Existence and properties of EIC. The Figure depicts case (i) of Proposition 2. There are three equilibria, 0, $\underline{q}$ and 1. The figure also illustrates the impacts of c and V as detailed in Proposition 3.

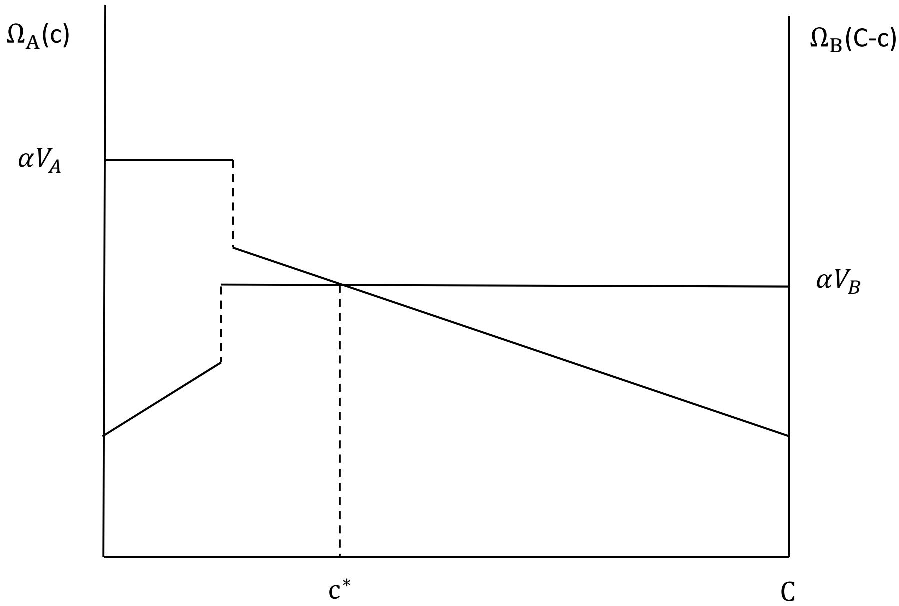

Fig.5: Existence of EMC. The figure displays the equilibrium $(q_A^*, C_A^*) = (1, c^*)$ and $(q_B^*, C_B^*) = (0, C-c^*)$.

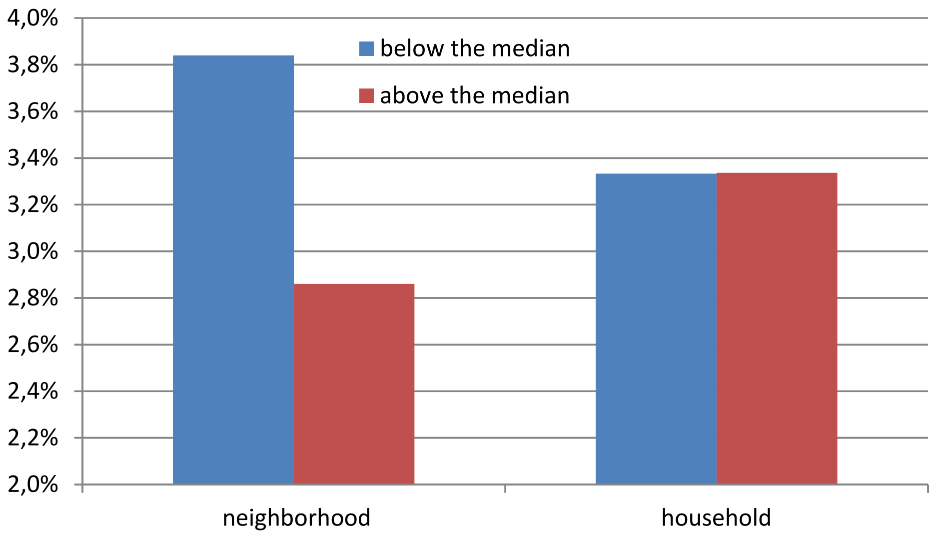

Fig.6: Equilibrium victimization by household and neighborhood income. The figure depicts the predicted victimization probability in the EMC with intra-neighborhood heterogeneity. The parameters are presented in Section 2.4.

Table 1: Descriptive statistics

Household characteristics	Mean	std.dev.
PR	1.0890	1.5311
BE	0.0375	0.2426
PA	0.3534	0.6219
income	0.0069	0.8755
couple	0.6208	0.6314
single woman	0.2309	0.5421
single man	0.1483	0.4677
education	5.3736	3.7001
age	6.9936	4.2504
occupation	0.6034	0.6369
children	0.7028	1.3190
past victimization	0.5144	0.6533
detached	0.5536	0.6464
semi-detached	0.1697	0.4969
apartment	0.2767	0.5885
owned	0.6741	0.6133
years in the dwelling	4.0180	1.6132
PA2	0.5426	0.6520
PA3	0.5681	0.6473
BE2	0.0740	0.3356
BE3	0.1128	0.4094

FSA characteristics	Mean	std.dev
BE	0.0378	0.0500
BE2	0.0746	0.0705
BE3	0.1111	0.0872
PR	1.0685	0.3808
PA	0.3473	0.1540
PA2	0.5303	0.1544
PA3	0.5567	0.1563
BE (CMA)	0.0370	0.0143
income	-0.0020	0.2701
income (census)	0.0153	0.1914
low income share	0.1702	0.0777

The sample size is about 24,000, but for household income, which is available for about 19,300 households. For FSA characteristics, we compute in each FSA the weighted mean of household characteristics. We then report the average and the standard deviation of such means. Source: GSS 1999, 2004, 2009

Table 2: Matrix of correlation coefficients between aggregate variables

	BE	PR	PA	income (census)	low income share	income (GSS)
BE	1					
PR	0.1274***	1				
PA	0.0789**	0.9414***	1			
income (census)	-0.1394***	0.1629***	0.1914***	1		
low income share	0.2186***	-0.1848***	-0.2439***	-0.6574***	1	
income (GSS)	-0.1153***	0.2652***	0.2992***	0.5478***	-0.5409***	1

All variables are weighted averages at FSA level. * 10%, ** 5%, *** 1%. Source: GSS 1999, 2004, 2009; Census 1996, 2001, 2006

Table 3: Basic results, second-stage regressions

	a	b	c	d	e	f
	OLS	IV	IV	OLS	IV	IV
dependent variable	PR	PR	PR	PA	PA	PA
hsd income	0.1188***	0.1144***	0.1151***	0.0472***	0.0457***	0.0460***
	0.016	0.016	0.016	0.006	0.007	0.007
FSA income	0.031	0.1230*	0.1073*	0.0152	0.0527**	0.0457*
	0.062	0.063	0.063	0.025	0.026	0.026
FSA low income share	1.5418***	1.1889***	1.0415***	0.5965***	0.4226***	0.3844***
	0.164	0.17	0.186	0.066	0.069	0.075
FSA BE	0.0616	0.6814***	1.8280***	0.0181	0.1939**	0.5680***
	0.185	0.196	0.531	0.074	0.077	0.21
FSA PR	0.8142***	0.2366***	0.2808***			
	0.025	0.055	0.053			
FSA PA				0.8213***	0.2717***	0.3330***
				0.025	0.055	0.052
hsd characteristics	yes	yes	yes	yes	yes	yes
year dummies	yes	yes	yes	yes	yes	yes
A-P Fstat [FSA PR or PA]		553.49	286.62		539.2	282.33
A-P Fstat [FSA BE]			131.95			136.78
P-Value Hansen Test		0.861	0.327		0.992	0.272
Cragg-Donald Statistics		739.017	156.201		729.07	163.436
R2	0.189	0.16	0.162	0.182	0.156	0.16
N	18525	18525	18525	18530	18530	18530

* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors below coefficients. Column a: OLS estimation; b: 2SLS where only FSA PR is instrumented; c: 2SLS with all instruments; d: OLS estimation; e: 2SLS where only FSA PA is instrumented; f: 2SLS with all instruments.

Table 4: Changes in the dependent variable

	a	b	c	d	e	f
	IV	IV	IV	IV	IV	IV
dependent variable	PR	PR	PA	PA	PA2	PA3
hsd income	0.1170***	0.1086***	0.0470***	0.0423***	0.0401***	0.0403***
	0.016	0.016	0.007	0.007	0.007	0.007
FSA income (census)	0.1091*	0.1085*	0.0465*	0.0468*	0.0314	0.0206
	0.065	0.064	0.026	0.026	0.028	0.028
FSA low income share	1.0566***	1.0488***	0.3911***	0.3792***	0.3800***	0.3188***
	0.191	0.187	0.077	0.076	0.082	0.084
FSA BE	1.8662***	1.2470**	0.5808***	0.4105*	0.6976***	0.7096***
	0.541	0.567	0.215	0.225	0.229	0.226
FSA PR-i	0.2649***					
	0.052					
FSA PR		0.2847***				
		0.054				
FSA PA-i			0.3159***			
			0.052			
FSA PA				0.3256***		
				0.053		
FSA PA2					0.2136***	
					0.065	
FSA PA3						0.2083***
						0.068
hsd characteristics	yes	yes	yes	yes	yes	yes
year dummies	yes	yes	yes	yes	yes	yes
A-P Fstat [FSA PR or PA]	289.95	277.61	284.59	272.36	187.37	201.4
A-P Fstat [FSA BE]	132.83	122.28	137.31	125.86	131.71	148.91
P-Value Hansen Test	0.333	0.181	0.278	0.253	0.318	0.227
Cragg-Donald Statistics	157.452	146.396	164.143	152.079	145.323	165.373
R2	0.13	0.164	0.122	0.159	0.127	0.134
N	18525	17781	18530	17786	18532	18531

* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors below coefficients. Columns a and c: FSA PR and FSA PA computed witout observation i; b and d: households with BE = 0.

Table 5: Changes in the victimization variable BE

	a	b	c	d
	IV	IV	IV	IV
dependent variable	PR	PR	PA	PA
hsd income	0.1160***	0.1163***	0.0462***	0.0463***
	0.016	0.016	0.007	0.007
FSA income	0.1378**	0.1525**	0.0540**	0.0564**
	0.064	0.064	0.026	0.026
FSA low income share	1.1622***	1.1555***	0.4305***	0.4343***
	0.174	0.172	0.07	0.069
FSA BE2	1.1070***		0.2767**	
	0.311		0.121	
FSA BE3		0.9784***		0.2246**
		0.229		0.088
FSA PR	0.2603***	0.2526***		
	0.056	0.057		
FSA PA			0.3228***	0.3290***
			0.054	0.053
hsd characteristics	yes	yes	yes	yes
year dummies	yes	yes	yes	yes
A-P Fstat [FSA PR or PA]	259.95	287.19	265.63	276.32
A-P Fstat [FSA BE]	207.78	282.95	222.53	284.35
P-Value Hansen Test	0.177	0.678	0.177	0.197
Cragg-Donald Statistics	210.722	275.812	234.211	292.434
R2	0.162	0.161	0.16	0.161
N	18525	18525	18530	18530

* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors below coefficients.
Columns a and c: BE2 instead of BE; b and d: BE3 instead of BE.

.

Table 6: FSA fixed effects

	e	f	g	h
	IV	IV	IV	IV
dependent variable	PR	PA	PR	PA
hsd income	0.1190***	0.0470***	0.1214***	0.0484***
	0.016	0.007	0.017	0.007
FSA income	-0.1986	-0.1015	-0.2013	-0.1035
	0.263	0.106	0.27	0.11
FSA low income share	-0.6379	-0.1996	-0.6565	-0.2042
	0.542	0.215	0.214	0.087
FSA BE	0.9543	0.2437	0.9918	0.2505
	0.842	0.334	0.862	0.346
FSA PR	0.3246***			
	0.113			
FSA PA		0.4363***		
		0.108		
FSA PR-i			0.2996***	
			0.112	
FSA PA-i				0.4151***
				0.109
hsd characteristics	yes	yes	yes	yes
year dummies	yes	yes	yes	yes
CMA fixed effects	no	no	no	no
FSA fixed effects	yes	yes	yes	yes
A-P Fstat [FSA PR or PA]	190.47	202.48	78.35	90.73
A-P Fstat [FSA BE]	65.56	64.27	95.98	97.95
P-Value Hansen Test	0.373	0.534	0.636	0.84
Cragg-Donald Statistics	84.638	82.303	124.129	131.355
R2	0.162	0.159	0.16	0.143
N	18525	18530	18525	18530

* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors below coefficients. Columns a to d: CMA fixed effects. Columns e to h: FSA fixed effects. Columns

Table A1: BE, BE2, BE3, PR, PA, PA2, PA3 by mean neighborhood income

	1st quartile	2nd quartile	3rd quartile	4th quartile
BE	0.0378	0.0379	0.0287	0.0292
std.dev.	0.0017	0.0018	0.0016	0.0016
PR	0.8534	0.9350	1.0110	1.0898
std.dev.	0.0100	0.0108	0.0111	0.0116
PA	0.2836	0.3118	0.3348	0.3697
std.dev.	0.0041	0.0044	0.0045	0.0047
BE2	0.0721	0.0740	0.0675	0.0639
std.dev.	0.0023	0.0025	0.0024	0.0024
BE3	0.1095	0.1148	0.1031	0.0960
std.dev.	0.0029	0.0030	0.0029	0.0029
PA2	0.4340	0.4671	0.5024	0.5284
std.dev.	0.0045	0.0047	0.0048	0.0049
PA3	0.4727	0.5090	0.5362	0.5580
std.dev.	0.0045	0.0047	0.0048	0.0049

Mean neighborhood income is in quartiles. std.dev. is the standard deviation of the mean. Source: GSS 1999, 2004, 2009 and Census 1996, 2001, 2006.

Table A2: BE, BE2, BE3, PR, PA, PA2, PA3 by household income

	1st quartile	2nd quartile	3rd quartile	4th quartile
BE	0.0320	0.0349	0.0374	0.0316
std.dev.	0.0016	0.0021	0.0021	0.0013
PR	0.7751	0.9428	1.0776	1.0598
std.dev.	0.0099	0.0130	0.0125	0.0093
PA	0.2426	0.3077	0.3637	0.3664
std.dev.	0.0040	0.0053	0.0051	0.0037
BE2	0.0623	0.0700	0.0775	0.0696
std.dev.	0.0022	0.0029	0.0029	0.0020
BE3	0.0987	0.1085	0.1186	0.1028
std.dev.	0.0028	0.0035	0.0035	0.0024
PA2	0.4133	0.4779	0.5290	0.5071
std.dev.	0.0046	0.0057	0.0053	0.0039
PA3	0.4497	0.5160	0.5675	0.5406
std.dev.	0.0046	0.0057	0.0053	0.0039

Household income is in quartiles. std.dev. is the standard deviation of the mean. Source: GSS 1999, 2004, 2009