Agglomeration, City Size and Crime

Carl Gaigné* and Yves Zenou†

April 5, 2013

Abstract

This paper analyzes the relationship between crime and agglomeration where the land, labor, product, and crime markets are endogenously determined. We show that in bigger cities there is *relatively* more crime, a standard stylized fact of most cities in the world. We also show that, in the short run when individuals are not mobile, a reduction in commuting costs (or a better access to jobs) decreases crime while, in the long run with free mobility, the effect is ambiguous. Finally, we show that the most efficient way of reducing total crime is to use *both* a transportation and a crime policy that decreases commuting costs and increases policy resources.

Key words: New economic geography, crime, agglomeration, policies.
JEL Classification: K42, R1.

*INRA, UMR1302 SMART, Rennes (France) and Laval University, Québec (Canada) France. Email: carl.gaigne@rennes.inra.fr.

†Stockholm University, IFN and GAINS. Email: yves.zenou@ne.su.se.

1 Introduction

It is well documented that there is more crime in big than in small cities (Glaeser and Sacerdote, 1999; Kahn, 2010). For example, the rate of violent crime in cities with more than 250,000 population is 346 per 100,000 inhabitants whereas in cities with less than 10,000 inhabitants, the rate of violent crime is just 176 per 100,000 (Glaeser, 1998). Similar figures can be found for property crimes or other less violent crimes.

The aim of this paper is to propose a model that captures some of these stylized facts and to analyze policies aiming at reducing crime. Our model delivers a full analytical solution that captures in a simple way how interactions between the land, product, crime and labor market yield agglomeration and criminal activity. Our model takes into account the following fundamental aspects of urban development: larger cities are associated with (*i*) higher nominal wages (Baum-Snow and Pavan, 2011); (*ii*) more varieties (Handbury and Weinstein, 2011); (*iii*) higher housing and commuting costs (Fujita and Thisse, 2013); (*iv*) higher crime rate (Glaeser and Sacerdote, 1999).

To be more precise, we develop an urban model where city size and the type of activities (crime and job) are endogenous within a full-fledged general equilibrium model. The individuals are freely mobile between and within the cities. We consider four different markets in each city: land, labor, product, and crime. The land market is assumed to be competitive and land is allocated to the highest bidders in each city. Land is owned by absentee landlords. The labor market is also competitive and wage are determined by free entry. Monopolistic competition prevails in the product market, which implies that each firm has a monopoly power on her variety. Finally, the crime market is competitive and the mass of criminals is determined by a cost-benefit analysis for each person.

In order to disentangle the various effects at work, we distinguish between what we call a *short-run equilibrium*, in which individuals are supposed to be spatially immobile and a *long-run equilibrium* when they are spatially mobile. In the short run, we find that a decrease in commuting costs will reduce crime because they reduce urban costs experienced by workers. In the long-run, when agents are perfectly mobile, this is no longer true. Indeed, if commuting costs are high, then the spatial equilibrium is such that there is full dispersion while, if there are small, then agglomeration prevails. In that case, a decrease in commuting costs, which favors agglomeration, has an ambiguous effect on crime. Indeed, in bigger cities, people earn higher wages but they experience higher urban costs and obtain higher proceeds from crime. As a result, from a social optimal viewpoint, it is better to have two cities of equal size (dispersion) than a city with 70 percent of people and another one with 30 percent. Finally, we consider a policy where each local government finance police resources by taxing workers. We show that this policy is efficient in reducing crime only if it decreases commuting costs and increases police resources in each city.

The paper is organized as follows. In the next section, we relate our model to the literature on crime and cities. Section 3 describes our framework. In Section 4, we study the decision to commit a crime and the share of criminals in each city with respect to their population size (short-run equilibrium) while, in Section 5, we determine the inter-city distribution of households (long-run equilibrium) when the share of criminals in each city adjusts to a change in its population size. In Section 6, we study the changes on criminal activity triggered by lower commuting costs by distinguished between the short-run and long-run effects. In Section 7, we examine a policy aiming at reducing criminal activity where local governments levy a tax on workers in order to increase policy resources. Finally, Section 8 concludes.

2 Related literature

To our knowledge, three types of theoretical models have integrated space and location in crime behavior. First, *social interaction* models that state that individual behavior not only depends on the individual incentives but also on the behavior of the peers and the neighbors are a natural way of explaining the concentration of crime by area. An individual is more likely to commit crime if his or her peers commit than if they do not commit crime (Glaeser et al., 1996; Calvó-Armengol and Zenou, 2004; Ballester et al., 2006, 2010; Calvó-Armengol et al., 2007; Patacchini and Zenou, 2012). This explanation is backed up by several empirical studies that show that indeed neighbors matter in explaining crime behaviors. Case and Katz (1991), using the 1989 NBER survey of young living in low-income, inner-city Boston neighborhoods, found that residence in a neighborhood in which many other youths are involved in crime is associated with an increase in an individual's probability of committing crime. Exploiting a natural experience (i.e. the Moving to Opportunity experiment that has assigned a total of 614 families living in high-poverty Baltimore neighborhoods into richer neighborhoods), Ludwig et al. (2001) and Kling et al. (2005) find that this policy reduces juvenile arrests for violent offences by 30 to 50 percent, relative to a control group. This also suggests very strong social interactions in crime behaviors. Patacchini and Zenou (2012) find that peer effects in crime are strong, especially for petty crimes. Bayer et al. (2009) consider the influence that juvenile offenders serving time in the same correctional facility have on each other's subsequent criminal behavior. They also find strong evidence of learning effects in criminal activities since exposure to peers with a history of committing a particular crime increases the probability that an individual who has already committed the same type of crime recidivates that crime.

Second, Freeman et al. (1996) provide a theoretical model that explains why criminals are concentrated in some areas of the city (ghettos) and why they tend to commit crimes in their own local areas and not in rich neighborhoods. Their explanation is based on the fact that, when criminals are numerous in an area, the probability to be caught is low so that criminals create a positive externality for each other. In this context, *criminals concentrate their effort in (poor) neighborhoods where the probability to be caught is small.*[1] This explanation has also strong empirical support (see e.g. O'Sullivan, 2000).

Finally, Verdier and Zenou (2004) show that prejudices and *distance to jobs* (legal activities) can explain crime activities, especially among blacks. If everybody believes that blacks are more criminal than whites -even if there is no basis for this- then blacks are offered lower wages and, as a result, locate further away from jobs. Distant residence increases even more the black-white wage gap because of more tiredness and higher commuting costs. Blacks have thus a lower opportunity cost of committing crime and become indeed more criminal than whites. Using 206 census tracts in city of Atlanta and Dekalb county and a state-of-the-art job accessibility measure, Ihlanfeldt (2002) demonstrates that modest improvements in the job accessibility of male youth, in particular blacks, cause marked reductions in crime, especially within category of drug-abuse violations. He found an elasticity of 0.361, which implies that 20 additional jobs will decrease the neighborhood's density of drug crime by 3.61%.[2]

Our contribution is different since we focus on the impact of inter-city mobility, city size and agglomeration effects on criminal behavior. We believe this is the first model that integrates crime and agglomeration economics in a unified framework by modeling the labor, crime, land and product market. In particular, our model is able to reproduce different stylized facts observed in real-world cities by showing that there exist big cities with both a high number of workers and criminals. We also show that the per-capita crime increases with the city size. In addition, our framework allows us to study the effect of different policies aiming at reducing crime.

[1] See also Deutsch et al. (1987).

[2] For a more detailed survey on the spatial aspects of crime, see Zenou (2003).

3 The model

Consider an economy with two cities, labeled $r = 1, 2$ and with a mass of population $N = 1$. Each individual exclusively is either a workers or a criminal. The spatial allocation of individuals and the choice of the type (worker or criminal) are endogenous.

We consider four different markets for each city: land, labor, product, and crime. The land market is assumed to be competitive and land is allocated to the highest bidders in each city. Land is owned by absentee landlords. The labor market is also competitive and wage are determined by free entry. Monopolistic competition prevails in the product market, which implies that each firm has a monopoly power on her variety. Finally, the crime market is competitive and the mass of criminals is determined by a cost-benefit analysis.

Cities Each city is formally described by a one-dimensional space. It can accommodate firms, criminals and workers. Whenever a city is formed, it has a Central Business District (CBD) located at $x = 0$ where the city r-firms are established.[3] In other words, any firm that wants to establish herself in a city r has to be located in the CBD. Workers and criminals can choose in which city they want to reside and will live between the CBD and the city fringe. Without loss of generality, we focus on the right-hand side of the city, the left-hand side being perfectly symmetric. Distances and locations are expressed by the same variable x measured from the CBD. Figure 1 describes a city r.

$$[Insert\ Figure\ 1\ here]$$

Each individual consumes a residential plot of fixed size (normalized to 1), regardless of her location. Denoting by λ_r the population residing in city r (with $\lambda_1 + \lambda_2 = N = 1$), the right-hand side of this city is then given by $\lambda_r/2$. The mass of workers in city r is denoted by L_r while that of criminals is C_r. As a result, $L_r + C_r = \lambda_r$ (see Figure 1).

Preferences and budget constraints Individuals are heterogeneous in their incentives to commit crime. They have different aversion to crime, denoted by c, so that higher c means more aversion towards crime. We assume that c is uniformly distributed on the interval $[0, 1]$.[4]

[3] See the survey by Duranton and Puga (2004) for the reasons explaining the existence of a CBD.

[4] See Verdier and Zenou (2004, 2012) and Conley and Wang (2006) for models with heterogenous disutilities of crime.

The individuals consume two types of goods: a *homogenous* good and *non-tradeable* goods (where trade costs are prohibitive), which are horizontally differentiated by varieties. One can think of a bundle of *services* locally produced like, for example, restaurants, retail shops, theaters, etc. (Glaeser et al, 2001). Preferences are the same across individuals and, for $v \in [0, n]$, the utility of a consumer in city r is given by:

$$U_r(q_0; q(v)) = \alpha \int_0^{n_r} q_r(v)\mathrm{d}v - \frac{\gamma}{2n_r}\left(\int_0^{n_r} q_r(v)\mathrm{d}v\right)^2 - \frac{(\beta-\gamma)}{2}\int_0^{n_r}[q_r(v)]^2\mathrm{d}v + q_0 \quad (1)$$

where $q_r(v)$ is the quantity of variety v of services and q_0 the quantity of the homogenous good, which is taken as the *numéraire*. All parameters α, β and γ are positive; $\gamma > 0$ measures the substitutability between varieties, whereas $\beta - \gamma > 0$ expresses the intensity for the love for variety. Equation (1), which has been extensively used in the economic geography literature (see, e.g. Ottaviano et al., 2002; Tabuchi and Thisse, 2002; Melitz and Ottaviano, 2008; Gaigné and Thisse, 2013), is a standard quasi-linear utility function with a quadratic sub-utility, symmetric in all varieties. Although this modeling strategy gives our framework a fairly strong partial equilibrium flavor, it does not remove the interaction between product, labor, land and crime markets, thus allowing us to develop a full-fledged model of agglomeration formation, independently of the relative size of the service sector. Note that the utility (1) degenerates into a utility function that is quadratic in total consumption $\int_0^{n_r} q_r(v)\mathrm{d}v$ when $\beta = \gamma$.

Each worker commutes to the CBD and pays a unit commuting cost per unit of distance of $t > 0$, so that a worker located at $x > 0$ bears a commuting cost equals to tx. The budget constraint of a *worker* residing at x in city r is given by

$$\int_0^{n_r} p_r(v)q_r(v)\mathrm{d}v + q_0 + R_r^w(x) + tx = I_r - \xi C_r + \overline{q}_0 \quad (2)$$

where $p_r(v)$ is the price of the service good for variety v, $R_r^w(x)$ is the land rent paid by workers (superscript w) located at x and I_r is the income of a worker. The homogenous good is available as an endowment denoted by $\overline{q}_0$; it can be shipped costlessly between the two cities. In this formulation, C_r is the mass of criminals in city r while ξ is a lump-sum amount stolen by each criminal. In other words, we assume that there are negative externalities of having criminals in a city r so that the higher is the number of criminals, the higher are these negative externalities. Observe that ξ is neither indexed by r nor by x meaning that the technology of criminals is the same in the two cities and within each city. On average, the stolen amount per worker increases with the mass of criminals in the city. In this formulation, each worker is "visited" by C_r criminals who each takes ξ.

By the law of large numbers, this means that, on average, a worker meets C_r criminals. This also implies that each criminal "attacks" L_r workers and takes ξ from each worker so that her average proceeds from crime is ξL_r.

Within each city, a worker chooses her location so as to maximize her utility (1) under the budget constraint (2).

The budget constraint of a *criminal* residing at x in city r is given by:

$$\int_0^{n_r} p_r(v)q_r(v)\mathrm{d}v + q_0 + R_r^c(x) = \xi L_r + \overline{q}_0 \tag{3}$$

where L_r is the mass of workers in city r and $R_r^c(x)$ is the land rent paid by criminals (superscript c) located at x . Observe that individuals are here specialized so that workers only work and do not commit crime while criminals only commit crime. As mentioned above, from equation (3), we see that the proceeds from crime ξL_r are increasing in the number of workers L_r in the city. For simplicity and to be consistent with (2), each criminal is assumed to steal a fraction ξ from these workers.

Technology and market structure Each variety of services v is supplied by a single firm and any firm supplies a single differentiated service under monopolistic competition. Labor is the only production factor. The fixed requirement of labor needed to produce variety v is denoted by $\phi > 0$, while the corresponding marginal requirement is set equal to zero for simplification. Note that a lower value of ϕ means a higher labor productivity. Hence, the profit made by a service firm v established in city r is given by:

$$\pi_r(v) = p_r(v)q_r(v)(L_r + C_r) - \phi I_r \tag{4}$$

where p_r is the price quoted by a service firm located in r and I_r the wage a service firm pays to its workers. Consistent with (2) and (3), this formulation (4) means that both criminals and workers consume all goods.

Services market, equilibrium prices and consumer's surplus The maximization of utility (1) under the budget constraint (2) or (3) leads to the demand for a service v given by:

$$q_r(v) = \frac{\alpha}{\beta} - \frac{p_r(v)}{(\beta-\gamma)} + \frac{\gamma}{\beta(\beta-\gamma)}\frac{P_r}{n_r} \tag{5}$$

where the price index $P_r = \int_0^{n_r} p_r(v)\mathrm{d}v$ is defined over the range of services produced in city r because this good is non-tradeable. Since the demand for each differentiated product does not depend on the net income (wage minus land rent and commuting costs) of each individual, it does not matter if the budget constraint is (2) or (3).

Each service firm determines its price by maximizing (4), using (5) and treating the price index P_r as a parameter. Solving the first-order conditions yield the equilibrium prices of a *non-tradeable* service for a variety in city r, given by:

$$p_r^* = \frac{\alpha(\beta - \gamma)}{\beta + (\beta - \gamma)} \equiv p^* \tag{6}$$

which is the same in both cities.[5] Hence, the consumer surplus generated by any variety at the equilibrium market price p^* is equal to:

$$S^*(v) = \frac{q^*(v)\,[p(0) - p^*]}{2} = \frac{(\alpha - p^*)^2}{\beta} = \frac{\alpha^2\beta^2}{[\beta + (\beta - \gamma)]^2 + \beta}$$

where $p(0)$ is the inverse demand when $q_r(v) = 0$. Note that the consumer surplus $S^*(v)$ for any variety is the same regardless of the city in which consumers live because all varieties are available everywhere at the same price. However, the consumer surplus generated by *all* varieties available in a city, i.e. $n_r S^*$, changes with the supply of varieties in city r. Without loss generality, we then set $S^* = 1$. This assumption does not affect qualitatively the properties of the spatial equilibria but greatly simplifies the algebra.

Urban labor market and equilibrium wages Because labor is the only factor of production, the number of varieties available in each city is proportional to the mass of individuals living *and* working in this city. More precisely, the labor market-clearing conditions imply

$$n_r = \frac{L_r}{\phi} \tag{7}$$

Urban labor markets are local and the equilibrium wage is determined by a bidding process in which firms compete for workers by offering them higher wages until no firm can profitably enter the market. In other words, operating profits are completely absorbed by the wage bill. This is a free-entry condition that sets profits (4) equal to zero so that, using (5) and (6), we find that the equilibrium wage paid by service firms established in city r is equal to:

$$I_r^* = \varphi(L_r + C_r) \tag{8}$$

where

$$\varphi \equiv \frac{(p^*)^2}{\phi} = \frac{1}{\phi}\left[\frac{\alpha(\beta - \gamma)}{\beta + (\beta - \gamma)}\right]^2 \tag{9}$$

Observe that φ corresponds to the real labor productivity. In accordance with empirical evidence, the equilibrium wage increases with population size ($N_r = L_r + C_r$). However,

[5] Note that our model does not capture the pro-competitive effects generated by the agglomeration of firms.

the equilibrium wage is unaffected by the residential location of each worker within the city. It is also worth stressing that the equilibrium wage (I_r^*) rises with product differentiation (low γ) and labor productivity (low fixed requirement in labor ϕ).

Land market and equilibrium land rents Let us first determine the equilibrium land rent for the *workers*. From the budget constraint (2), we obtain:

$$q_0 = I_r^* - R_r^w(x) - tx - \xi C_r + \overline{q}_0 - \int_0^{n_r} p_r(v) q_r(v) \mathrm{d}v$$

By plugging this value and the equilibrium quantities and prices (5) and (6) into the utility (1), we obtain:

$$V_r^w(x) = n_r^* + I_r^* - R_r^w(x) - tx - \xi C_r + \overline{q}_0 \tag{10}$$

Because of the fixed lot size assumption (normalized to 1), the value of the consumption of the nonspatial goods $\int_0^{n_r} q_r(v) p_r(v) \mathrm{d}v + q_0$ at the residential equilibrium is the same regardless of the worker's location. Using (2), this implies that the total *urban costs*, $UC_r^w(x) \equiv R_r^w(x) + tx$, borne by a worker living at location x in city r, is constant whatever the location x.

Since *criminals* do not commute to the CBD, which implies that their utility does not depend on location x, we have: $R_r^c(x) = R_r^c$. In equilibrium, since it is costly for workers to be far away from the CBD, they will bid away criminals who will live at the city fringe, paying the opportunity cost of land R^a so that $R^c = R^a$. Without loss of generality, the opportunity cost of land is normalized to zero, i.e. $R^a = 0$.

For workers, given $V_r^w(x)$, the equilibrium land rent in the city must solve $\partial V_r^w(x)/\partial x = 0$ or, equivalently, $\frac{\partial R_r^w(x)}{\partial x} + t = 0$, whose solution is $R_r^w(x) = r_0 - tx$, where r_0 is a constant.[6] Because the opportunity cost of land R^a is equal to zero, it has to be that $R_r^w(L_r/2) = 0$ (see Figure 1) so that $r_0 = L_r/2$. As a result, the equilibrium land rent for workers is equal to:

$$R_r^w(x) = t\left(\frac{L_r}{2} - x\right) \tag{11}$$

and the urban costs borne by a worker are given by:

$$UC_r^w = t\frac{L_r}{2}. \tag{12}$$

[6] We could easily extend the model to take into account the fact that workers residing further away from criminals experience lower negative externalities. For example, if we assume that $\xi(x) = \xi_0 + \xi_1 x$ so that the criminals steal less to workers residing closer to the center, we can show that the results remain qualitatively the same.

4 Criminal activities when city choices are exogenous

Assume that workers do not choose in which city they live and let $\theta_r = C_r/N_r$ be the share of criminals in city r and $\lambda_r = N_r/N$ the share of individuals living in city r. Hence, we have

$$C_r = \theta_r \lambda_r \quad \text{and} \quad L_r = (1 - \theta_r)\lambda_r$$

In this section, because location choices are exogenous, θ_r is endogenously determined for any given population size λ_r. An individual becomes criminal in city r if and only if $V_r^c - V_r^w > 0$, where V_r^c and V_r^w are the utility of a criminal and a worker living in city r evaluated at the equilibrium prices. Plugging the equilibrium land rent (11) into (10), we obtain:

$$V_r^w = n_r + I_r^* - \xi C_r - tL_r/2 + \overline{q}_0 \tag{13}$$

From the budget constraint of criminals, (3), we obtain:

$$q_0 = \xi L_r - R^a + \overline{q}_0 - \int_0^{n_r} p_r(v) q_r(v) dv$$

By plugging this value and the equilibrium quantities and prices (5) and (6) into the utility (1) and adding the cost c of committing crime, we obtain:

$$V_r^c = n_r + \xi L_r - c + \overline{q}_0 \tag{14}$$

Thus, the value of c making a marginal individual indifferent between committing a crime and working is $\widetilde{c}_r$ and is given by

$$\widetilde{c}_r = (\xi - \varphi)\,\lambda_r + \frac{t(1-\theta_r)}{2}\lambda_r \tag{15}$$

where φ is defined by (9). Hence, because of the uniform distribution of c, the fraction of criminals in city r is $\theta_r = \widetilde{c}_r$. The equilibrium share of criminals θ_r^* is thus given by:

$$\theta_r^* = \frac{\lambda_r t + 2\lambda_r\,(\xi - \varphi)}{\lambda_r t + 2} \tag{16}$$

It is easily verified that $\theta_r^* < 1$ if and only if $\lambda_r < 1/(\xi - \varphi)$. A sufficient condition is $\xi < 1 + \varphi$. We thus assume throughout that:

$$\xi < 1 + \varphi \tag{17}$$

Moreover, $\theta_r^* > 0$ if and only if $t > 2(\varphi - \xi)$. In this context, it is easily checked that $\partial\theta_r^*/\partial t > 0$ as soon as $\theta_r^* < 1$, which is guaranteed by (17). This means that, for a given population size λ_r, *higher commuting costs lead to more criminal activities in each city.* Indeed, since the total urban cost increases with commuting costs, the net wages of workers is reduced, which, in turn, leads to a larger fraction of individuals committing crime. This implies that a transport policy that aims at improving access to jobs (lower t) would reduce criminal activities in the short run. We will investigate this issue below. In addition, because $\partial^2\theta_r^*/\partial\lambda_r\partial t > 0$, the impact of commuting costs on criminal activities is higher when the city size increases. This is because urban costs are positively correlated with population size and thus the effect of commuting costs on land rents is higher in larger cities.

Furthermore, $\partial\theta_r^*/\partial\gamma > 0$, which means that the mass of criminals decreases with more differentiated products (lower γ). Indeed, when γ decreases, φ increases, meaning that the revenue per worker is higher for firms because there is less price competition and thus workers obtain higher wages, which deters criminality. A similar effect can be found for the labor productivity $(1/\phi)$ since $\partial\theta_r^*/\partial\phi > 0$.

If we now look at the effect of city size on criminal activities, we see that $\partial\theta_r^*/\partial\lambda_r > 0$. In particular, it is easily seen that $\theta_1^* > \theta_2^*$ if and only if $\lambda_1 > 1/2$. Thus, *a larger population in a city triggers more criminals in this city.* The number of criminals decreases in the smaller city but increases in the large city when agglomeration takes place.

Since there are two cities, $\lambda_1 = \lambda$ and $\lambda_2 = 1-\lambda$ (with $\lambda = N_1/N$), the mass of criminals with respect to the relative size of cities is given by:

$$C(\lambda) = \lambda\theta_1^* + (1-\lambda)\theta_2^* = \frac{[t + 2(\xi - \varphi)]\,[(4-t)\lambda^2 - (4-t)\lambda + 2]}{(2+\lambda t)[2+(1-\lambda)t]} \tag{18}$$

It is straightforward to check that $\partial C(\lambda)/\partial\lambda > 0$ as long as $\lambda \geq 1/2$. This means that, when the size of the population in the first city is more than 50 percent, then the total crime in the economy increase with λ. There is a $U-$shape relationship between total crime C and λ as illustrated in Figure 2. In our model agglomeration is defined by $\lambda \neq 1/2$ and the farther away is λ from 1/2, the more there is agglomeration. We have thus the following result that formalizes the claim made in Introduction.

Proposition 1 *Agglomeration raises criminal activities in the economy.*

[*Insert Figure* 2 *here*]

If a federal planner wants to minimize total crime $C(\lambda)$, then it will be optimal to have two symmetric cities for which $\lambda = 1/2$. Agglomeration increases total crime because of *multiplier effects*. As a result, it is better to have two cities of equal size than a city with 70 percent of people and another one with 30 percent (i.e. $\lambda = 0.7$). This is because, in bigger cities, people are more induced to be criminals since they experience higher urban costs (land rents and commuting costs) and obtain higher proceeds from crime (see (15)). However, they also obtain a higher wage. Proposition 1 shows that the former effect dominates the latter and thus total crime increases with larger cities. This gives a microfoundation to the empirical result found in Glaeser and Sacerdote (1999).

5 Criminal activities and urbanization

Let us now endogeneize the location choice of all individuals. The timing of the model is as follows. In the first stage, households choose in which city they will reside without knowing their type c but anticipating (with rational expectations) the average total population of criminals. The assumption that types are revealed only after location choices has been made to take into account the relative inertia of the land market compared to the crime and labor markets. In the second stage, types (or honesty parameters) are revealed and individuals decide to commit crime or not. In the third stage, goods are produced, workers participate in the labor market while criminals participate in the crime market and all consume the two types of goods. Observe that in the second stage, workers are stuck in their initial locations (decided in the first stage) and cannot relocate themselves. They then decide to become criminal or not.

5.1 Location choices

Consider now the location choice of individuals. The location of individuals is driven by the inter-city difference in their expected utility. Before knowing their c, the expected utility of living in city r is given by (using (13) and (14)):

$$\mathbf{EV}_r = \int_0^{\widetilde{c}_r} V_r^c \mathrm{d}c + \int_{\widetilde{c}_r}^1 V_r^w \mathrm{d}c = \theta_r^2/2 + V_r^w(\theta_r)$$

where $\widetilde{c}_r = \theta_r$ and θ_r is given by (16). We have $\lambda_1 + \lambda_2 = 1$ so that λ is the endogenous share of individuals located in city r. Note that this expected utility is based on θ_r, the average proportion of criminals in city r. Note also that, though the individual demands (5) are unaffected by income, the migration decision takes income into account. Everything else equal, workers are drawn by the higher wage city. The population becoming larger, the local demand for the services is raised, which attracts additional firms. In addition, households are attracted by larger cities to access to more varieties. However, the competition for land among workers increases land rent and commuting costs, which both increase with population size. These different mechanisms interact with the decision to become a criminal and, in turn, the level of agglomeration.

Hence, the spatial difference in the expected utility $\mathbf{EV}_1 - \mathbf{EV}_2 \equiv \Delta\mathbf{EV}$ is given by:

$$\Delta\mathbf{EV}(\lambda, \theta_1, \theta_2) = \left(\lambda - \frac{1}{2}\right)\Gamma(\lambda, t) \tag{19}$$

where

$$\Gamma(\lambda, t) \equiv \frac{-t(1+\xi)(\xi-\varphi)(1-\lambda)\lambda + 2 - 2\xi - 2\xi^2 + 4\varphi + 2\varphi\xi}{(1+\lambda t/2)[1+(1-\lambda)t/2]} + \frac{2(\xi-\varphi+t/2)^2[t(1-\lambda)\lambda+1]}{(1+\lambda t/2)^2[1+(1-\lambda)t/2]^2}$$

We would like now to analyze the equilibrium of this economy, which is defined so that no individual (worker or criminal) has an incentive to change location (or city).

Definition 2

(*i*) *An equilibrium arises at* $0 < \lambda^* < 1$ *when the utility differential* $\Delta\mathbf{EV}[\lambda^*, \theta_r(\lambda^*)] = 0$, *or at* $\lambda^* = 1$ *when* $\Delta\mathbf{EV}[1, \theta_r(1)] > 0$ *or at* $\lambda^* = 0$ *when* $\Delta\mathbf{EV}[0, \theta_r(0)] < 0$.

(*ii*) *An interior equilibrium is stable if and only if the slope of the indirect utility differential* $\Delta\mathbf{EV}$ *is strictly negative in a neighborhood of the equilibrium, i.e.,* $d\Delta\mathbf{EV}[\lambda^*, \theta_r(\lambda^*)]/d\lambda < 0$ *at* λ^*.

(*iii*) *An fully agglomerated equilibrium (i.e. when* $\lambda^* = 1$ *or* $\lambda^* = 0$*) is stable whenever it exists.*

It is well-known that new economic geography (NEG) models typically display several spatial equilibria (Fujita and Thisse, 2013). In such a context, it is convenient to use stability as a selection device since an unstable equilibrium is unlikely to happen. This is what is exposed in Definition 2 where an interior equilibrium is stable if, for any marginal deviation away from the equilibrium, the incentive system provided by the market brings the distribution of individuals back to the original one. In (*ii*), we give the conditions for which the equilibrium is stable.

5.2 Spatial equilibria

Let us now investigate in more detail all the possible equilibria. First, *full dispersion* ($\lambda^* = 1/2$) is always an equilibrium whatever the value of t since $\Delta\mathbf{EV}(1/2) = 0$. Second, there is an equilibrium with *full agglomeration* ($\lambda^* = 0$ or $\lambda^* = 1$) if and only if $\Delta\mathbf{EV}(1) > 0$ and $\Delta\mathbf{EV}(0) < 0$. Using (19), it is easily checked that these two conditions are satisfied if and only if $t < \underline{t}(\xi)$, where

$$\underline{t}(\xi) \equiv 2\left[-\xi^2 + \xi\varphi + \varphi + (1+\varphi-\xi)\sqrt{1+(1+\xi)^2}\right] \tag{20}$$

In Figure 3, we have depicted $\underline{t}(\xi)$, which is a non-linear curve that increases and then decreases up to $\underline{t}(\xi) = 0$ for which $\xi = \overline{\xi} \equiv \sqrt{3+4\varphi+\varphi^2} - 1$, with $\overline{\xi} \in (\varphi, \varphi+1)$. Notice that $\underline{t}(\xi) > 2(\varphi - \xi)$ for all $t, \xi > 0$, which guarantees that $\theta_r^* > 0$ where there is no full agglomeration. Let us now study *partial agglomeration* ($\lambda^* \in (1/2, 1)$), which occurs when $\Gamma(\lambda^*) = 0$. This is the case when $t > \underline{t}(\xi)$ so that we have $\Delta\mathbf{EV}(1) < 0 < \Delta\mathbf{EV}(0)$.

[*Insert Figure* 3 *here*]

5.3 Stability analysis

Let us now look at the stability of the interior equilibria since full-agglomeration equilibria are always stable (see Definition 2). An interior solution $\lambda = \lambda^*$ is stable if and only if

$$\left.\frac{\mathrm{d}\Delta\mathbf{EV}}{\mathrm{d}\lambda}\right|_{\lambda=\lambda^*} = \Gamma(\lambda^*, t) + \left(\lambda^* - \frac{1}{2}\right)\frac{\mathrm{d}\Gamma}{\mathrm{d}\lambda} < 0$$

We have two types of interior solutions: *full dispersion* with $\lambda^* = 1/2$ and *partial agglomeration* with $\lambda^* \in (1/2, 1)$.

Stability for a full dispersion equilibrium In Appendix 1, we determine the conditions under which a full dispersion equilibrium ($\lambda^* = 1/2$) is a *stable* equilibrium (or equivalently $\Gamma(1/2, t) < 0$). From Appendix 1 (Lemma 8), we can conclude that $\lambda^* = 1/2$ is a stable spatial configuration (i) when $\xi \geq \overline{\xi}$ regardless of commuting costs (t); (ii) when $\overline{\xi} > \xi > \varphi$ if and only if $t > \overline{t}_2$; (iii) when $\xi = \varphi$ if and only if $t > \overline{t}_3$; (iv) when $\xi < \varphi$ if and only if $\overline{t}_1 > t > \overline{t}_2$ and $\varphi > \xi > \max\{\underline{\xi}, 0\}$, $\overline{t}_1$, $\overline{t}_2$, and $\overline{t}_3$ being defined in Appendix 1.

Some comments are in order. First, full dispersion is more likely to occur when commuting costs are high enough (like in NEG models, see Gaigné and Thisse, 2013) and when the amount stolen by criminals (ξ) is high enough. Second, a same set of parameters may yield two stable spatial equilibria ($\lambda_r^* = 1/2$ or $\lambda_r^* = 1$ when $t < \overline{t}$ and $\xi \geq \overline{\xi}$ for example). In other words, different levels of criminal activity may emerge for the same economic conditions.

Stability for a partial agglomeration equilibrium In a partial agglomeration equilibrium $\lambda^* \in (1/2, 1)$ such that $\Gamma(\lambda^*, t) = 0$ is stable if and only if $\mathrm{d}\Gamma/\mathrm{d}\lambda < 0$ when $\lambda = \lambda^*$. Note that $\Gamma(\lambda^*, t) = 0$ has at most one solution when $\lambda \in (1/2, 1)$. Let λ^i be the implicit solution of $\Gamma(\lambda, t) = 0$. If $\Gamma(1/2, t) < 0$ (where $\Gamma(1/2, t) = \mathrm{d}\Delta\mathbf{EV}/\mathrm{d}\lambda$ when $\lambda^* = 1/2$) and $\Delta\mathbf{EV}(1) > 0$, then λ^i exists but is unstable. By contrast, if $\Gamma(1/2, t) > 0$ and $\Delta\mathbf{EV}(1) < 0$, then λ^i exists and is stable. In other words, an *asymmetric spatial configuration* emerges when commuting costs take intermediate values. In addition, under this spatial configuration, we have

$$L_1^* - L_2^* = \left(\lambda^* - \frac{1}{2}\right)\Lambda(\lambda^*)$$

where

$$\Lambda \equiv \frac{2(1 - \xi + \varphi) + t(\varphi - \xi)(1 - \lambda^*)\lambda^*}{(1 + t\lambda^*/2)[1 + t(1 - \lambda^*)/2]}$$

and where $L_1^* = L_2^*$ when $\lambda^* = 1/2$ and $L_1^* > L_2^*$ when $\lambda^* > 1/2$ because $\partial\Lambda(\lambda)/\partial\lambda > 0$. In other words, the large city hosts more workers and more criminals. It is also worth stressing that, ex post, *workers living in the smaller city are better off than workers living in the larger city* (ex ante they all have the same expected utility). Indeed, because $\theta_1^* > \theta_2^*$ and $\Delta\mathbf{EV}(\lambda) = 0$ when $1/2 < \lambda^* < 1$, then $V_1^w < V_2^w$.

The following proposition summarizes all our main findings whereas Figure 3 displays the spatial equilibria in the $t - \xi$ space.

Proposition 3 *There are three stable spatial equilibria with respect to commuting costs:*

(*i*) *If $\Gamma(1/2, t) < 0$, i.e. when commuting costs are high enough, there are two identical cities in population size, $\lambda_1^* = \lambda_2^* = 1/2$, and in share of criminals, $\theta_1^*(1/2) = \theta_2^*(1/2)$.*

(*ii*) *If $\Gamma(1/2, t) > 0$ and $t > \underline{t}$, i.e. when commuting costs take intermediate values, there is a large city and a small city where the former has more criminals and more workers than the latter, $1/2 < \lambda_r^* < 1$, $L_1 > L_2$ and $\theta_1^* > \theta_2^*$.*

(*iii*) *If $t < \underline{t}$, i.e. when commuting costs are low enough, there is a single city .*

6 Impact of commuting costs on criminal activities

We now analyze the impact of commuting costs (t) on the criminal activity when there is *free mobility* between the two cities. This parameter can be interpreted as a more efficient transport policy or a better access to jobs.

We have seen in Section 4 that the impact of a reduction in commuting costs on total criminal activities was positive when the location choice of individuals is exogenous (i.e. when λ was given). This is not true anymore when individuals choose location and, in fact, the total effect is ambiguous. Indeed, at any given location of households, lower commuting costs reduce the number of criminals in each city ($\partial C_r/\partial t > 0$) because $\partial \theta_r/\partial t > 0$ regardless of city r. On the other hand, the location of individuals adjusts in the long run to a change in commuting costs. More precisely, falling commuting costs promotes agglomeration ($\partial \lambda/\partial t < 0$) and, in turn, more crimes are committed in the larger city ($\partial C_1/\partial \lambda > 0$) while the number of crime in the small city shrinks ($\partial C_2/\partial \lambda < 0$). As a result, the long-run effect associated with falling commuting costs on criminal activity is ambiguous. Even though lower commuting costs induce higher legitimate net income for all workers, they also promote higher levels of agglomeration.

Consider first that the economy shifts from full dispersion to full agglomeration due to lower commuting costs. Under these spatial configuration, we have

$$C(1/2) = \frac{t/2 + \xi - \varphi}{2 + t/2} \quad \text{for } \Gamma(1/2, t) < 0 \text{ and } \quad C(1) = \frac{t/2 + \xi - \varphi}{1 + t/2} \quad \text{for } t < \underline{t}.$$

For example, it appears $C(1/2, t = \overline{t}_2) > C(1, t = \underline{t})$ if and only if $\varphi > \widehat{\varphi}$ where

$$\widehat{\varphi} \equiv \frac{2(5 + 3\xi)\sqrt{2\xi + \xi^2 + 2} - 5\xi^2 - 15\xi - 14}{1 + \xi}$$

and $\widehat{\varphi}$ is positive and increases with ξ. Hence, a shift from dispersion to agglomeration due to lower commuting costs may give rise to a decline in criminal activity. The final effect is that there are less criminal activities in the economy (the former effect dominates the latter effect).

In addition, $C(\lambda)$ reaches its minimum value when $t \leq t_{\min}$ ($C = 0$). It is straightforward to check that $t_{\min} < \underline{t}$ so that $C = 0$ may occur when $\lambda^* = 1$ and not when when $\lambda^* = 1/2$.[7] Therefore, improving access to jobs by reducing commuting costs can be a relevant policy tool in reducing crime.

When t decreases when partial agglomeration occurs, the degree of agglomeration (λ^*) increases gradually so that the relationships between $C(\lambda^*)$ and t is ambiguous when $\overline{t} > t > \underline{t}$. Because λ^* is highly nonlinear, we need to perform some numerical simulations. These simulations reveal a $U-$shaped relationship between $C(\lambda^*)$ and t. There exists a threshold value $\widehat{t}$ such that $C(\lambda^*)$ decreases with a reduction in commuting costs when $\overline{t} > t > \widehat{t}$. However, whether criminal activity may fall in the economy when t moves from

[7] Indeed, $\Delta(1) = 1 + \xi$ when $t = t_{\min}$ and $\mathrm{d}\mathbf{\Delta EV}(\lambda^*)/\mathrm{d}\lambda > 0$ at $\lambda^* = 1/2$ when $t = t_{\min}$.

$\bar{t}$ to $\underline{t}$, crime increases occur in the larger city due to a larger population size. Figure 4 displays the relationship between crime and commuting costs.

$$[Insert\ Figure\ 4\ here]$$

To summarize,

Proposition 4 *When there is no intercity-mobility, decreasing commuting costs always increase total crime. When there is free mobility between the two cities, reducing commuting costs increases total crime is more likely to occur if φ is low and ξ high ($\varphi < \widehat{\varphi}$).*

Indeed, when t decreases, there will be more agglomeration, which leads to two opposite effects. On the one hand, the urban costs in the big city will increase compared to the small city and thus more people decide to become criminal. On the other hand, real wages increases in the big city because of a bigger market size, which reduces the number of criminals. The net effect is ambiguous. When φ is high and ξ low, the latter effect dominates the former one for a large range of commuting costs while, we have the reverse result, when φ is low and ξ is high. Remember that the incentive to become a criminal is relatively strong when real labor productivity (φ) is low or crime productivity (ξ) is high.

7 Police resources and crime

Assume now that the number of active criminals in city r is given by $(1 - a_r)C_r$ where a_r is the share of criminals in jail or equivalently the individual probability of being caught (by the law of large number). This share of active criminals depends on the resources affected by local government to fight criminal activity. We consider that the probability of arresting a criminal is an increasing function of the *per capita* public resources, denoted by T_r. These public expenditures are financed by a local head tax paid by workers (τ_r). More precisely, we assume that $a_r \equiv f(T_r)$, where $T_r = \tau_r L_r / \lambda_r$, which are the total resources per capita invested in police for the local government. For a same level of tax revenue, the probability to be arrested is lower in a larger a city. We also assume that $f(0) = 0$, $f'(T_r) > 0$ and $f''(T_r) < 0$. Increasing the resources devoted to police and decreasing the population size raises the probability of arresting a criminal.

The timing is now follows. In the first stage, individuals freely choose which city to reside in, anticipating the head tax they will pay and the wage they will earn. In second stage, each governments chooses a head tax to maximize the welfare of the representative worker. Last, in the third stage, types (or honesty parameters) are revealed and individuals decide to commit crime or not while product, land and labor markets clear. As usual, the game is solved by backward induction.[8]

7.1 Taxation and share of criminals

Let us solve stage 3 where crime is decided for given λ_r and τ_r. Using (14), the indirect utility of a criminal located in city r is now given by

$$V_r^c = n_r + [1 - f(T_r)]\,\xi L_r - c + \overline{q}_0 \tag{21}$$

whereas, using (13), we have:

$$V_r^w = n_r + I_r^* - [1 - f(T_r)]\,\xi C_r - tL_r/2 + \overline{q}_0 - \tau_r \tag{22}$$

which is the indirect utility of a worker living in city r. Indeed, the number of active criminals is $1 - a_r$ since $a_r \equiv f(T_r)$ represents of the fraction of criminal in jail (incapacity effect). Note that the effect of the local head tax τ_r on V_r^w is ambiguous since there is direct negative effect and an indirect positive effect through $f(T_r)$. For the criminals, we assume, for simplicity, that only the consumption of the numeraire is affected if she is arrested. Using (21) and (22) and the fact that $L_r = (1 - \theta_r)\,\lambda_r$, $L_r + C_r = \lambda_r$, and $I_r^* = \varphi\lambda_r$, the value of c making an individual indifferent between committing a crime and working is now given by

$$\widetilde{c}_r = \left(\frac{t}{2} + \xi - \varphi\right)\lambda_r + \tau_r - \xi f(T_r)\lambda_r - \theta_r\frac{t\lambda}{2}.$$

Because of the uniform distribution of c, we assume the following sufficient condition $\partial\widetilde{c}_r/\partial\theta_r < 0$ for all $\theta_r \in [0, 1]$ to ensure that there exists an equilibrium share of criminals in each city. We thus assume throughout that

$$t\lambda/2 - \xi\lambda_r f'(.)\tau_r > 0 \tag{23}$$

holds. In addition, $\theta_r < 1$ if and only if $\widetilde{c}_r(\theta_r) < 1$ or, equivalently, $(\xi - \varphi)\,\lambda_r + \tau_r < 1$ (if $\theta_r = 1$ then $f(T_r) \to 0$, *i.e.* the probability to be arrested is close to zero when there

[8]Note that the specification of governments' objective is a controversial issue in our case because individuals are mobile among cities (Scotchmer, 2002; Cremer and Pestieau, 2004) and they can work or be criminals. We consider that local goverments disregard the indirect utility of criminals. We also avoid the difficulty associated with the mobility of individuals because governments know who their workers are, and thus may maximize the welfare of workers because the size of the population is exogenous (because of our timing).

is no worker because there is no public resource). Notice also that $\theta_r > 0$ requires that $\widetilde{c}_r(\theta_r = 0) > 0$ or, equivalently, $(t/2 - \varphi)\lambda_r + \tau_r > 0$ (if $\theta_r = 0$ then $f(T_r) \to 1$, i.e. the probability to be arrested is close to one when a worker becomes a criminal if there is no criminal). As a result, when $1 - (\xi - \varphi)\lambda_r > \tau_r > (\varphi - t/2)\lambda_r$, the equilibrium share of criminals is given by $\widetilde{c}_r = \theta_r$, or equivalently,

$$\theta_r^*(\tau_r)[1 + t\lambda_r/2] + \xi f(T_r)\lambda_r - (t/2 + \xi - \varphi)\lambda_r - \tau_r = 0 \tag{24}$$

Lemma 5 *There exists a unique minimum $\tau_r = \tau_r^c > 0$ implicitly defined by*

$$\frac{\partial \theta_r^*}{\partial \tau_r} = \frac{-\xi\lambda_r f'(.)(1-\theta_r^*) + 1}{1 + t\lambda_r/2 - \xi\lambda_r f'(.)\tau_r} = 0 \tag{25}$$

Furthermore, for a given city size λ_r, a decrease in commuting costs reduces the share of criminals in each city, i.e.

$$\frac{\partial \theta_r^*}{\partial t} = \frac{(1-\theta_r^*)\lambda_r/2}{1 + t\lambda_r/2 - \xi\lambda_r f'(T_r)\tau_r} > 0$$

Proof. See Appendix 2.

A rise in the tax rate has an ambiguous effect on the share of criminals in each city. On the one hand, it increases the probability of arresting a criminal so that less individuals have an incentive to become a criminal. On the other hand, it directly reduces the legitimate net income for all workers, making the criminal activity more attractive. Hence, there is an U-shaped relationship between θ_r^* and τ_r (see Figure 5) and θ_r^* reaches its minimum value when $\tau_r = \tau_r^c$. Indeed, starting from low levels of tax rate, higher tax burden reduces the share of criminals in the city. Above the critical value of tax rate (τ_r^c), criminal activities raise with tax burden. In addition, as expected, regardless of tax rate prevailing in each city, the share of criminals in the city is reduced when there is a decrease in commuting cost, as expected.

[*Insert Figure* 5 *here*]

It is worth stressing that the tax rate maximizing the probability of arresting a criminal τ_r^a is identical to the tax rate maximizing the tax revenue and is given by

$$\frac{\mathrm{d}a_r}{\mathrm{d}\tau_r} = f'(.)\frac{\mathrm{d}T_r}{\mathrm{d}\tau_r} = f'(.)\left(1 - \theta_r^* - \tau_r\frac{\partial\theta_r^*}{\partial\tau_r}\right) = 0$$

Starting from low tax rate, higher tax rates raise public resources per capita and, in turn, increase the probability of arresting a criminal (see Figure 5). Beyond τ_r^a, a rise in tax rate reduces the revenue from the tax because the number of taxpayers (workers) reaches low values (a variant of the Laffer curve).[9] It is straightforward to check $\mathrm{d}a_r/\mathrm{d}\tau_r > 0$ when $\tau_r = \tau_r^c$ so that $\tau_r^a > \tau_r^c$. Hence, *a local government maximizing the public resources to fight criminal activity induces more criminals and tax burden than a local authority minimizing the number of criminals.*

7.2 Tax Policy, police and criminal activity.

Let us now solve stage 2 where the two governments set a tax rate that maximizes the welfare of the representative worker (or, equivalently, the median voter), given by (22). By rising its tax rate, the local government increases the welfare of workers by reducing the share of criminals who are in jail and decreases the welfare by raising the tax burden and land rents. In addition, remember that, if the fraction of criminal in jail increases with τ_r, its effect on the share of criminal in the city is ambiguous. Using $C_r = \theta_r^* \lambda_r$, $L_r = (1-\theta_r^*)\lambda_r$, $n_r = L_r$ and $I_r^* = \varphi\lambda_r$ as well as $V_r^w = V_r^c(\widetilde{c}_r)$, (22) can be written as

$$V_r^w = (1-\theta_r^*)\lambda_r - \theta_r^* + \xi\left[1 - f(.)\right](1-\theta_r^*)\lambda_r.$$

The first order condition is given by:

$$\frac{\mathrm{d}V_r^w}{\mathrm{d}\tau_r} = -\left\{\lambda_r + 1 + \xi\left[1-f(.)\right]\lambda_r\right\}\frac{\partial\theta_r^*}{\partial\tau_r} - \xi\lambda_r(1-\theta_r^*)f'(.)\frac{\mathrm{d}T_r}{\mathrm{d}\tau_r}$$

The equilibrium tax rate is given by τ_r^d such that $\mathrm{d}V_r^w/\mathrm{d}\tau_r = 0$. Since $\partial\theta_r^*/\partial\tau_r = 0$ implies that $\xi\lambda_r f'(T_r)(1-\theta_r^*) = 1$, we have

$$\left.\frac{\mathrm{d}V_r^w}{\mathrm{d}\tau_r}\right|_{\tau_r=\tau_r^c} = -\left.\frac{\mathrm{d}T_r}{\mathrm{d}\tau_r}\right|_{\tau_r=\tau_r^c} = -(1-\theta_r^*) < 0$$

As a result, $\tau_r^d < \tau_r^c$ and, when $\tau_r = \tau_r^d$, $\partial\theta_r^*/\partial\tau_r < 0$. Hence, at given city size λ_r, the tax rate maximizing the utility of the median voter is lower than the tax rate inducing the minimum value of the share of criminals (see Figure 5) because, at τ_r^c, taxes are too high. We can also conclude that $\mathrm{d}T_r/\mathrm{d}\tau_r > 0$ when $\tau_r = \tau_r^d$. Hence, the equilibrium tax rate is in the upward-sloping portion of the curve.

In addition, there exists an interior solution which is positive if and only if

$$\left.\frac{\partial V_r^w}{\partial\tau_r}\right|_{\tau_r=0} = -(1+\lambda_r+\xi\lambda_r)\left.\frac{\partial\theta_r^*}{\partial\tau_r}\right|_{\tau_r=0} - \xi f'(0)(1-\theta_r^*)^2\lambda_r > 0$$

[9] Note that $\mathrm{d}^2a_r/\mathrm{d}\tau_r^2$ when $\tau_r = \tau_r^a$.

where $\partial\theta_r/\partial\tau_r < 0$ when $\tau_r = 0$ and $\xi f'(0)(1-\theta_r^*)\lambda_r > 1$. As a consequence, $\tau_r^d > 0$ if and only if

$$\frac{\xi\lambda_r^2 f'(0)\left[1-(\xi-\varphi)\lambda_r\right](1+2\xi-\varphi)}{1+\lambda_r+\xi\lambda_r} > (1+t\lambda_r/2)$$

where $1-(\xi-\varphi)\lambda_r > 0$. Hence, the local governments are more likely to levy taxes to fight criminal activity when commuting costs are low enough. Furthermore lower commuting costs impact directly the share of criminals in each city and through a change in tax policy. Indeed, we have

$$\frac{\mathrm{d}\theta_r^*}{\mathrm{d}t} = \frac{\partial\theta_r^*}{\partial t} + \frac{\partial\theta_r^*}{\partial\tau_r}\frac{\partial\tau_r^d}{\partial t} = \frac{\partial\theta_r^*}{\partial t} + \frac{\partial\theta_r^*}{\partial\tau_r}\left(-\frac{\partial^2 V_r^w}{\partial t\partial\tau_r}/\frac{\partial^2 V_r^w}{\partial\tau_r^2}\right)$$

Because $\frac{\partial^2 V_r^w}{\partial\tau_r\partial t}$ is highly non linear, we perform some numerical simulations to study the sign of $\frac{\partial\tau_r^d}{\partial t}$. Figure 6 displays the results of the simulations.[10] From our simulations, it appears that $\mathrm{d}\tau_r^d/\mathrm{d}t < 0$ and, in turn, $\mathrm{d}\theta_r^*/\mathrm{d}t > 0$. Hence, local governments adjust downward their tax rate when commuting costs decline, leading to less criminals in each city (see Figure 5). This means that, when commuting costs are low, the optimal tax rate τ_r^d chosen by the local government is closer to τ_r^c, the tax rate that minimizes total crime. The intuition of this result is straightforward. Because the incentive to become a criminal is lower and the workers' welfare increase when commuting costs decline, each local government can increase its tax rate so that the share of criminal shrinks.

[*Insert Figure* 6 *here*]

To summarize,

Proposition 6 *Assume no inter-city mobility. Then, lower commuting costs make the impact of police on crime more efficient.*

7.3 Crime, location and taxes

In short run (no inter-city mobility), we have seen that more police resources reduce criminal activities, especially when commuting costs are low (6). However, in the long run, population location reacts to a change in tax rates and crime rates. Therefore, we need to solve the first stage of our game. The impact of tax rate on expected utility is as follows:

$$\left.\frac{\mathrm{d}\mathbf{EV}_r}{\mathrm{d}\tau_r}\right|_{\tau_r=\tau_r^d} = \theta_r^*\left.\frac{\mathrm{d}\theta_r^*}{\mathrm{d}\tau_r}\right|_{\tau_r=\tau_r^d} + \left.\frac{\mathrm{d}V_r^w}{\mathrm{d}\tau_r}\right|_{\tau_r=\tau_r^d} = \theta_r^*\frac{-\xi\lambda_r f'(.)\,(1-\theta_r^*)+1}{1+t\lambda_r/2-\xi\lambda_r f'(.)\tau_r^d} \equiv \Omega_r < 0$$

[10] We consider that $f(.) = \sqrt{\tau_r(1-\theta_r^*)}$ as well as $\varphi = 2\xi = 2$ and $\lambda = 0.8$.

Hence, when each local government applies its best fiscal strategy to fight criminal activity, the expected utility declines in each city. However, the magnitude of this negative effect rises when the city size increases. Indeed, we obtain

$$\frac{\mathrm{d}\Omega_r}{\mathrm{d}\lambda_r} = \frac{\mathrm{d}\theta_r^*}{\mathrm{d}\lambda_r} \times \left.\frac{\mathrm{d}\theta_r^*}{\mathrm{d}\tau_r}\right|_{\tau_r=\tau_r^d} + \theta_r^* \times \left.\frac{\mathrm{d}^2\theta_r^*}{\mathrm{d}\lambda_r \mathrm{d}\tau_r}\right|_{\tau_r=\tau_r^d}$$

where, by using (24),

$$\frac{\mathrm{d}^2\theta_r^*}{\mathrm{d}\lambda_r \mathrm{d}\tau_r} = \frac{-\xi f'(.)\left(1-\theta_r^*-\tau_r\right)-t/2}{\left[1+t\lambda_r/2-\xi\lambda_r f'(.)\tau_r\right]^2} < 0.$$

and $\mathrm{d}\theta_r^*/\mathrm{d}\tau_r < 0$ when $\tau_r = \tau_r^d$ (see Section 7.2). Hence, dispersion is strengthened due to a tax policy and, thus, criminal activity declines. As a result, in the long run, *falling commuting costs trigger dispersion through higher tax rates in the larger city, thus inducing less criminals in the economy.*

Proposition 7 *When there is free inter-city mobility, decreasing commuting costs in each region leads to a higher increase of police resources (or tax rate) in the bigger region, which, in turn, raises dispersion and reduces total crime.*

This proposition means that the most efficient way of reducing crime is to use both a transportation and crime policy by reducing commuting costs and increases policy resources (or tax rate).

8 Concluding remarks

This paper provides the first model of agglomeration and crime in a general equilibrium framework. We develop a two-city model where both firms and individuals (workers and criminals) are freely mobile between and within the cities. We deliver a full analytical solution that captures in a simple way how interactions between the land, product, crime and labor market yield agglomeration and criminal activity. Our model takes into account the following fundamental aspects of urban development: larger cities are associated with higher crime rate, higher nominal wages, more product varieties and higher housing and commuting costs.

First, we show that different stable spatial equilibria emerge. When commuting costs are high enough, then an equilibrium will full dispersion in population size and share of criminals between the two identical cities prevails. When commuting costs take intermediate values, there is a large city and a small city where the former has more criminals and more workers than the latter. Finally, when commuting costs are low enough, there is a single city.

Second, when there is no intercity-mobility, we show that decreasing commuting costs always increases total crime. On the contrary, when there is free mobility between the two cities, we find that, if the real labor productivity is high (low) and crime productivity low (high), then reducing commuting costs reduces (increases) total crime.

Finally, we take into account how the optimal resources are affected to each local government to fight criminal activity and assume that the probability of arresting a criminal is an increasing function of the per capita public resources. Each local government sets a tax rate that maximizes the welfare of the representative worker (i.e. the median voter). By rising its tax rate, each local government increases the welfare of workers by reducing the share of criminals who are in jail and decreases the welfare by raising the tax burden and land rents. This means that the effect of the tax on the share of criminal in the city is ambiguous. When there no inter-city mobility, we show that lower commuting costs make the impact of police on crime more efficient. On the contrary, when there is free inter-city mobility, decreasing commuting costs in each region leads to a higher increase of police resources (or tax rate) in the bigger region, which, in turn, raises dispersion and reduces total crime.

References

[1] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2006), "Who's who in networks. Wanted: the key player," *Econometrica* 74, 1403-1417.

[2] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2010), "Delinquent networks," *Journal of the European Economic Association* 8, 34-61.

[3] Bayer, P., Hjalmarsson, R. and D. Pozen (2009), "Building criminal capital behind bars: Peer effects in juvenile corrections," *Quarterly Journal of Economics* 124, 105-147.

[4] Baum-Snow N. and R. Pavan (2011), "Understanding the city size wage gap," *Review of Economic Studies* 79, 88-127.

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

[6] Calvó-Armengol, A. and Y. Zenou (2004), "Social networks and crime decisions: The role of social structure in facilitating delinquent behavior," *International Economic Review* 45, 935-954.

[7] Case, A.C., and L.F. Katz. 1991. "The company you keep: The effects of family and neighborhood on disadvantaged youths," NBER Working Paper No. 3705.

[8] Conley, J.P. and P. Wang (2006), "Crime and ethics," *Journal of Urban Economics* 60, 107-123.

[9] Cremer, H. and P. Pestieau (2004), "Factor mobility and redistribution," In: J.V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics, Vol. 4*, Amsterdam: North-Holland, pp. 2529-2560.

[10] Deutsch, J., Hakim, S. and J. Weinblatt (1987), "A micro model of the criminal's location choice," *Journal of Urban Economics* 22, 198-208.

[11] Duranton and Puga (2004), "Micro-foundations of urban agglomeration economies," In: J.V. Henderson and J-F Thisse (Eds.), *Handbook of Urban and Regional Economics*, *Vol. 4*, Amsterdam: North Holland, pp. 2063-2117.

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

[13] Fujita, M. and J.-F. Thisse (2013), *Economics of Agglomeration, Second edition*, Cambridge: Cambridge University Press.

[14] Gaigné, C. and J.-F. Thisse (2013), "New economic geography and cities," In: Fisher, M. and P. Nijkamp (Eds.), *Handbook of Regional Science*, Berlin: Springer Verlag, forthcoming.

[15] Glaeser, E.L. (1998), "Are cities dying?" *Journal of Economic Perspectives* 12, 139-160.

[16] Glaeser, E.L., Kolko, J. and A. Saiz (2001), "Consumer city," *Journal of Economic Geography* 1, 27-50.

[17] Glaeser, E.L. and B.I. Sacerdote (1999), "Why is there more crime in cities," *Journal of Political Economy* 107, S225-S258.

[18] Glaeser, E.L., Sacerdote, B. and J. Scheinkman (1996), "Crime and social interactions," *Quarterly Journal of Economics* 111, 508-548.

[19] Handbury, J. and D.E. Weinstein (2011), "Is new economic geography right? Evidence from price data," NBER Working Paper No. 17067.

[20] Ihlanfeldt, K.R. (2002), "Spatial mismatch in the labor market and racial differences in neighborhood crime," *Economics Letters* 76, 73-76.

[21] Kahn, M.E. (2010), "New evidence on trends in the cost of urban agglomeration," In: E.L. Glaeser (Ed.), *Economics of Agglomeration*, Chicago: The University of Chicago University Press, pp. 339-354.

[22] Kling, J.R., Ludwig, J. and L.F. Katz (2005), "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment," *Quarterly Journal of Economics* 120, 87-130.

[23] Ludwig, J., Duncan, G.J. and Hirschfield, P. (2001), "Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment," *Quarterly Journal of Economics* 116, 655-679.

[24] Melitz, M.J. and G.I.P. Ottaviano (2008), "Market size, trade, and productivity," *Review of Economic Studies* 75, 295-316.

[25] O'Sullivan, A. (2000), *Urban Economics*, Fourth edition, New York: Irwin.

[26] Patacchini, E. and Y. Zenou (2012), "Juvenile delinquency and conformism," *Journal of Law, Economic, and Organization* 28, 1-31.

[27] Ottaviano, G.I.P., Tabuchi, T. and J.-F. Thisse (2002), "Agglomeration and trade revisited," *International Economic Review* 43, 409-436.

[28] Scotchmer, S. (2002) Local public goods and clubs. In: A. Auerbach and M. Feldstein (eds.). *Handbook of Public Economics, Vol. 3*. Amsterdam: North-Holland, 1997-2042.

[29] Tabuchi, T. and J.-F. Thisse (2002), "Taste heterogeneity, labor mobility and economic geography," *Journal of Development Economics* 69, 155-177

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

[31] Zenou, Y. (2003), “The spatial aspects of crime,” *Journal of the European Economic Association* 1, 459-467.

Appendix

8.1 Appendix 1: Conditions under which $\lambda^* = 1/2$ is a spatial equilibrium

Let $\Gamma(1/2,t) \equiv \frac{\mathrm{d}\Delta\mathbf{EV}}{\mathrm{d}\lambda}\big|_{\lambda=1/2}$ be the slope of $\Delta\mathbf{EV}$ at $\lambda = 1/2$ where

$$\Gamma(1/2,t) = \frac{(1+\xi)(\varphi-\xi)t^2 + 4[3\xi(\varphi-\xi)-(2-\varphi-\xi)]t - 16(\xi-\overline{\xi})(\xi+2+\overline{\xi})}{4(1+t)^3}$$

with

$$\overline{\xi} = \sqrt{3+4\varphi+\varphi^2} - 1 \in (\varphi, \varphi+1)$$

where $\Gamma(1/2,0) = 4(\xi-\overline{\xi})(\xi+2+\overline{\xi})$ and

$$\frac{\mathrm{d}\Gamma(1/2,t)}{\mathrm{d}t}\bigg|_{t=0} = (7+3\varphi)\left[\xi - \frac{3\varphi+8}{3\varphi+7}(\varphi+1)\right] < 0 \tag{26}$$

which is negative as long as $\xi < \varphi + 1$. In addition, $\Gamma(1/2,t) = 0$ when $t = \overline{t}_1$ or $t = \overline{t}_2$ with

$$\overline{t}_1 \equiv 2\frac{(2-\varphi-\xi) - 3\xi(\varphi-\xi) + (2+\varphi-\xi)\sqrt{(1+\xi)^2 + 4(\xi^2-\xi\varphi-\varphi)}}{(1+\xi)(\varphi-\xi)} \tag{27}$$

$$\overline{t}_2 \equiv 2\frac{(2-\varphi-\xi) - 3\xi(\varphi-\xi) - (2+\varphi-\xi)\sqrt{(1+\xi)^2 + 4(\xi^2-\xi\varphi-\varphi)}}{(1+\xi)(\varphi-\xi)} \tag{28}$$

where $2+\varphi-\xi > 0$ when $\xi < 1+\varphi$, $\overline{t}_2 = 0$ when $\xi = \overline{\xi}$ and $\overline{t}_2 > 0 > \overline{t}_1$ when $\xi > \varphi$.

As a result,

(i) When $\xi \geq \overline{\xi}$, the function $\Gamma(1/2,t)$ is concave with $\Gamma(1/2,0) \leq 0$ and $\Gamma(1/2,t) < 0$ as long as $t > 0$. Indeed, if $t = 0$ then $\Gamma(1/2,t) \leq 0$ (when $\xi \geq \overline{\xi}$) and $\mathrm{d}\Gamma(1/2,t)/\mathrm{d}t < 0$ when $t = 0$. By implication, $\lambda^* = 1/2$ is a stable spatial configuration regardless of $t \geq 0$ when $\xi \geq \overline{\xi}$.

(ii) When $\overline{\xi} > \xi > \varphi$, the function $\Gamma(1/2,t)$ is still concave where we have now $\Gamma(1/2,0) > 0$ and $\overline{t}_2 > 0 > \overline{t}_1$. Thus, there exists a single positive value of t ($\overline{t}_2$) such that $\Gamma(1/2,\overline{t}_2) = 0$ and $\Gamma(1/2,t) < 0$ if and only if $t > \overline{t}_2$. Hence, when $\overline{\xi} > \xi > \varphi$, $\lambda^* = 1/2$ is a stable spatial configuration if and only if $t > \overline{t}_2$.

(iii) When $\xi = \varphi$, there exists a single positive value of t ($4(1+\varphi)/(1-\varphi)$) such that $\Gamma(1/2,t) = 0$ and $\Gamma(1/2,t) < 0$ if and only if $t > 4(1+\varphi)/(1-\varphi)$. Hence, when $\xi = \varphi$, $\lambda^* = 1/2$ is a stable spatial configuration if and only if $t > \overline{t}_3$ with

$$\overline{t}_3 \equiv \frac{4(1+\varphi)}{1-\varphi}$$

(iv) When $\varphi > \xi$, we have $\bar{t}_1 > \bar{t}_2$ and the function $\Gamma(1/2, t)$ becomes convex. Hence, $\Gamma(1/2, t) < 0$ if and only if $\bar{t}_1 > t > \bar{t}_2$. Note that $\bar{t}_2 > 0$ if and only if $\xi > \max\{\underline{\xi}, 0\}$ with

$$\underline{\xi} \equiv \frac{2\varphi}{5} - \frac{1}{5} + \frac{2}{5}\sqrt{\varphi^2 + 4\varphi - 1}$$

where $\underline{\xi} = \varphi$ when $\varphi = 1$ and $\underline{\xi} < \varphi$ if and only if $\varphi < 1$. Hence, $\lambda^* = 1/2$ is a stable spatial configuration if and only if $\bar{t}_1 > t > \bar{t}_2$ and $\varphi > \xi > \max\{\underline{\xi}, 0\}$. Note that that $\varphi > \max\{\underline{\xi}, 0\}$ if and only if $\varphi < 1$. Hence, $\lambda^* = 1/2$ is not a spatial equilibrium when $\varphi > 1$ or when $\xi < \max\{\underline{\xi}, 0\}$.

To sum up,

Lemma 8 *Full dispersion ($\lambda = 1/2$) is a stable spatial configuration*

(i) when $\xi \geq \bar{\xi}$ regardless of commuting costs t;

(ii) when $\bar{\xi} > \xi > \varphi$ if and only if $t > \bar{t}_2$;

(iii) when $\xi = \varphi$ if and only if $t > \bar{t}_3$;

(iv) when $\xi < \varphi$ if and only if $\bar{t}_1 > t > \bar{t}_2$ and $\varphi > \xi > \max\{\underline{\xi}, 0\}$.

8.2 Appendix 2: Proof of Lemma 5

The crime rate is given by (24) where

$$\theta_r^*(0) = \frac{(t/2 + \xi - \varphi)\lambda_r}{1 + t\lambda_r/2} > 0$$

By differentiating (24), we easily obtain:

$$\frac{\partial \theta_r^*(\tau_r)}{\partial \tau_r} = \frac{-\xi\lambda_r f'(T_r)\,(1-\theta) + 1}{1 + t\lambda_r/2 - \xi\lambda_r f'(T_r)\tau_r} = 0$$

which is (25). Denote by

$$N \equiv -\xi\lambda_r f'(T_r)\,(1-\theta) + 1$$

and by

$$D \equiv 1 + t\lambda_r/2 - \xi\lambda_r f'(T_r)\tau_r$$

Observe that

$$\left.\frac{\partial \theta_r^*(\tau_r)}{\partial \tau_r}\right|_{\tau_r=0} < 0$$

$$\frac{\partial^2\theta_r^*}{\partial\tau_r^2}=\frac{\left[-\xi\lambda_r^3 f''(T_r)\left(1-\theta\right)^2+\xi\lambda_r^2 f'(T_r)\frac{\partial\theta_r^*}{\partial\tau_r}\right]D+\xi\lambda_r^2\left[f''(T_r)\left(1-\theta\right)\lambda_r\tau_r+f'(T_r)\right]N}{D^2}$$

We look for a minimum and thus $\frac{\partial\theta_r^*(\tau_r)}{\partial\tau_r}=0$, which implies that $N=0$ and $\frac{\partial\theta_r^*}{\partial\tau_r}=0$. Thus

$$\frac{\partial^2\theta_r^*}{\partial\tau_r^2}=\frac{-\xi\lambda_r^3 f''(T_r)\left(1-\theta\right)^2}{D}>0$$

and thus there is a unique minimum that we denote by $\tau_r=\tau_r^c$.

Furthermore, by differentiating (24), we have

$$\frac{\partial\theta_r^*(\tau_r)}{\partial t}=\frac{\left(1-\theta_r^*\right)\lambda_r/2}{1+t\lambda_r/2-\xi\lambda_r f'(T_r)\tau_r}$$

which is positive when (23) holds.

Finally, τ_r^c is defined by:

$$-\xi\lambda_r f'(\tau_r^c L_r)\left(1-\theta\right)+1=0$$

which is equivalent to

$$\tau_r^c=\frac{1}{L_r}f'^{-1}\left(\frac{1}{\xi\lambda_r\left(1-\theta\right)}\right)>0$$

which proves that $\tau_r^c>0$. ▮

Figure 1: Equilibrium land use in city *r*

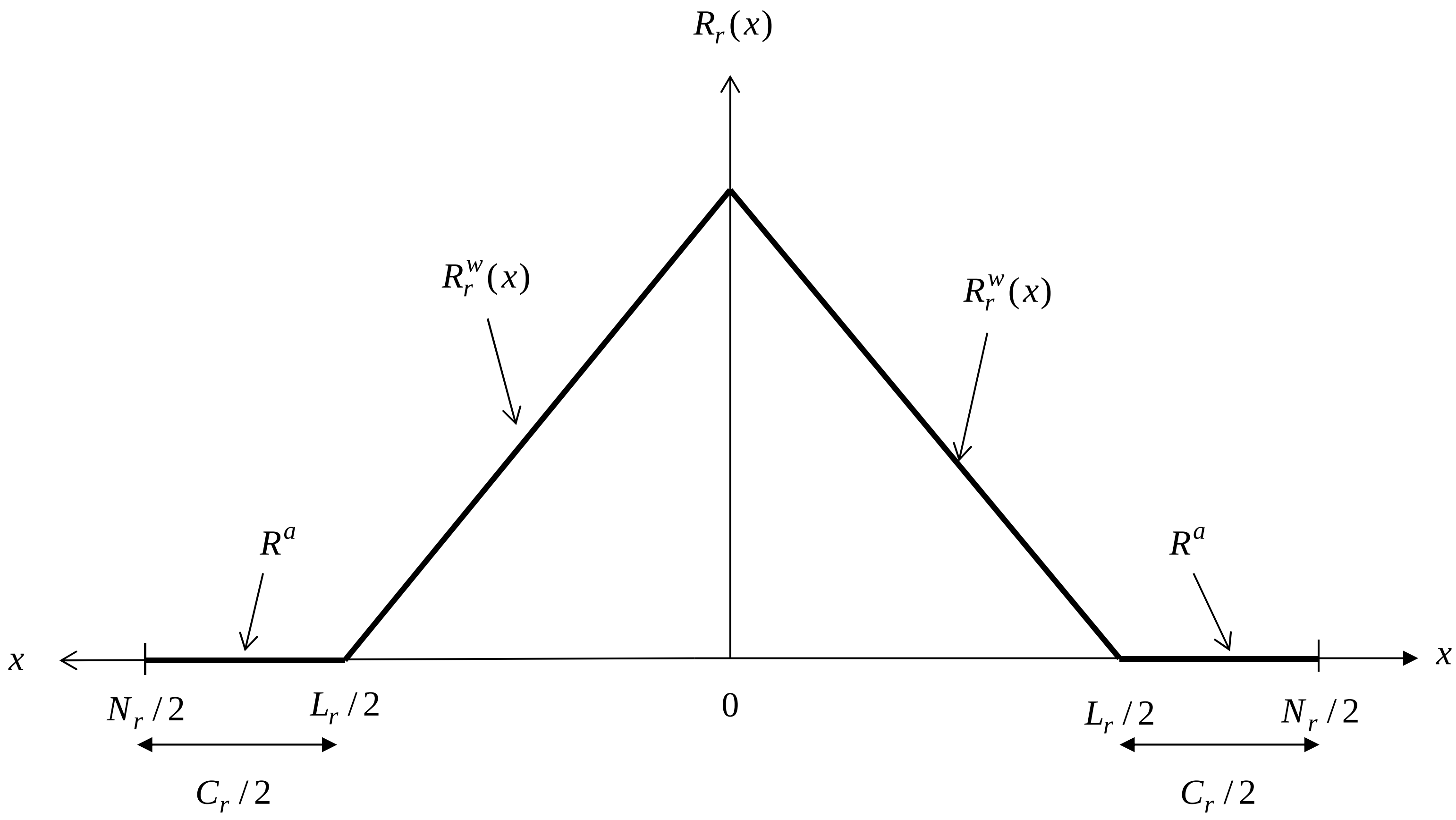


Figure 2: Relationship between total crime and agglomeration when individuals' location choices are exogenous

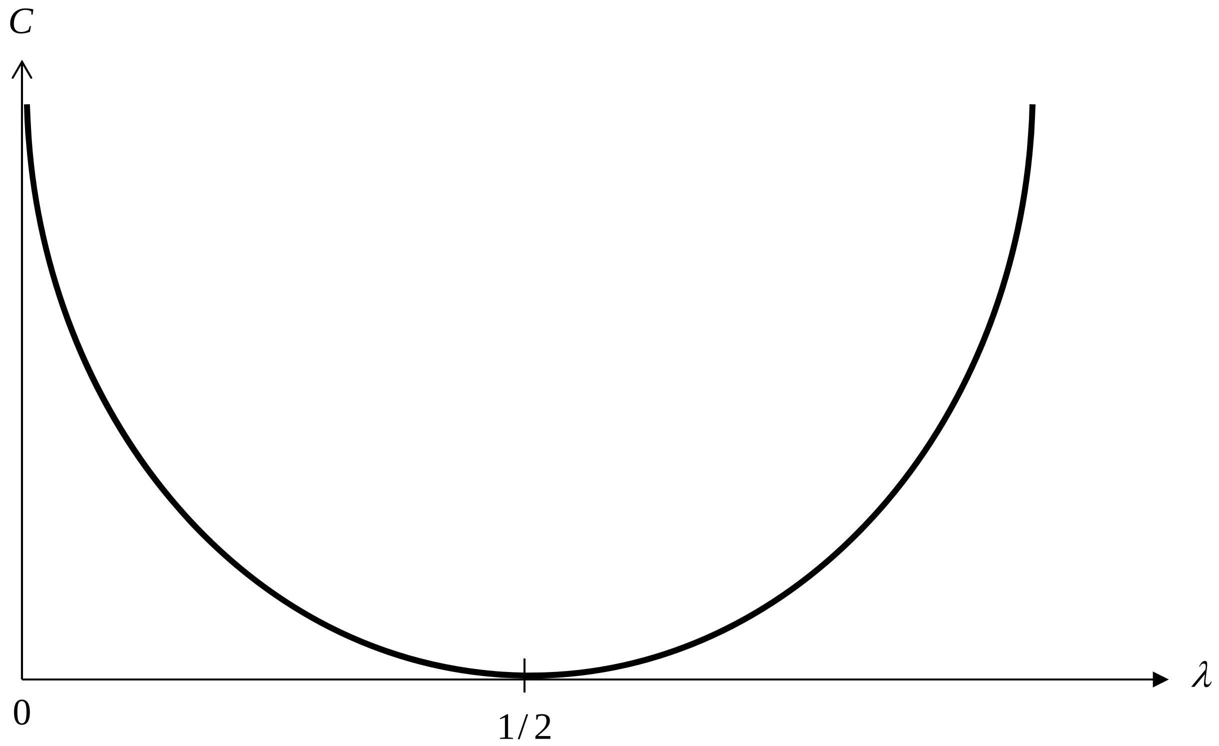
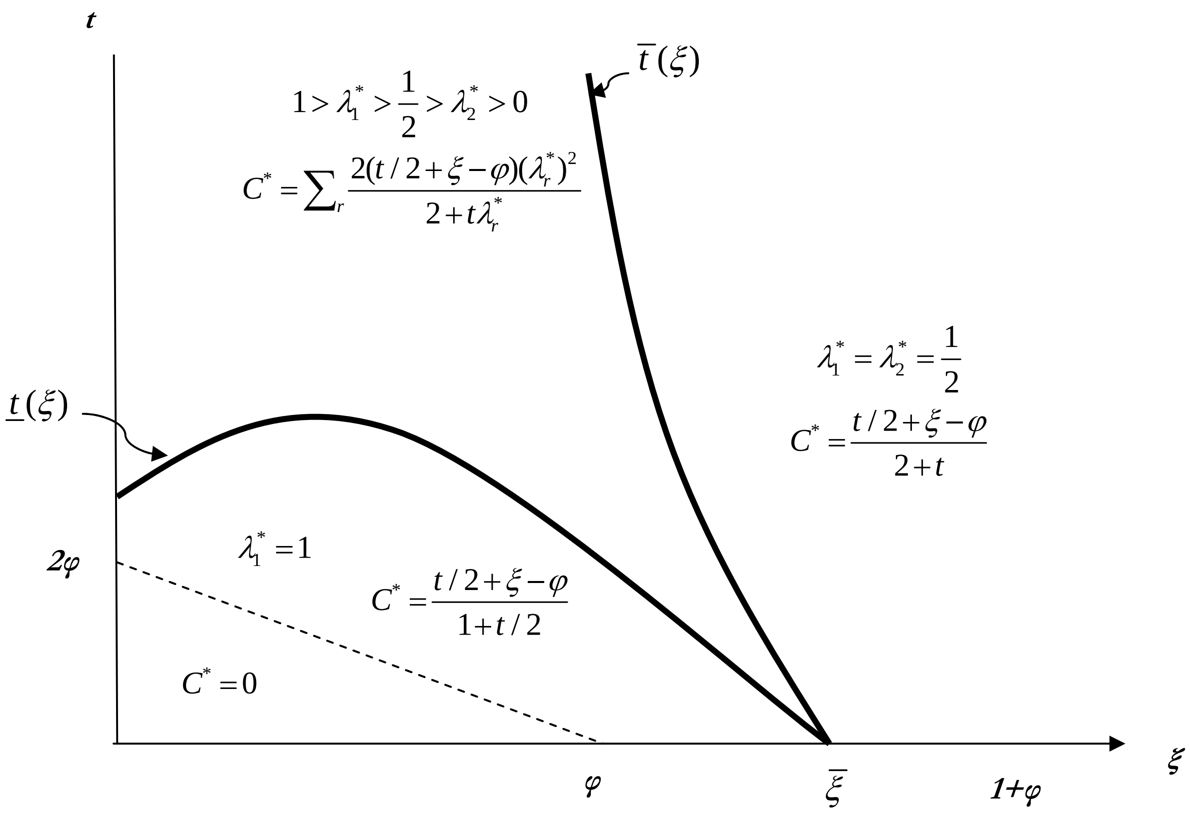


Figure 3: Spatial equilibria and crime in (*t,ξ)*-space.

(a) high φ (φ>1)



(b) low φ (φ<1)

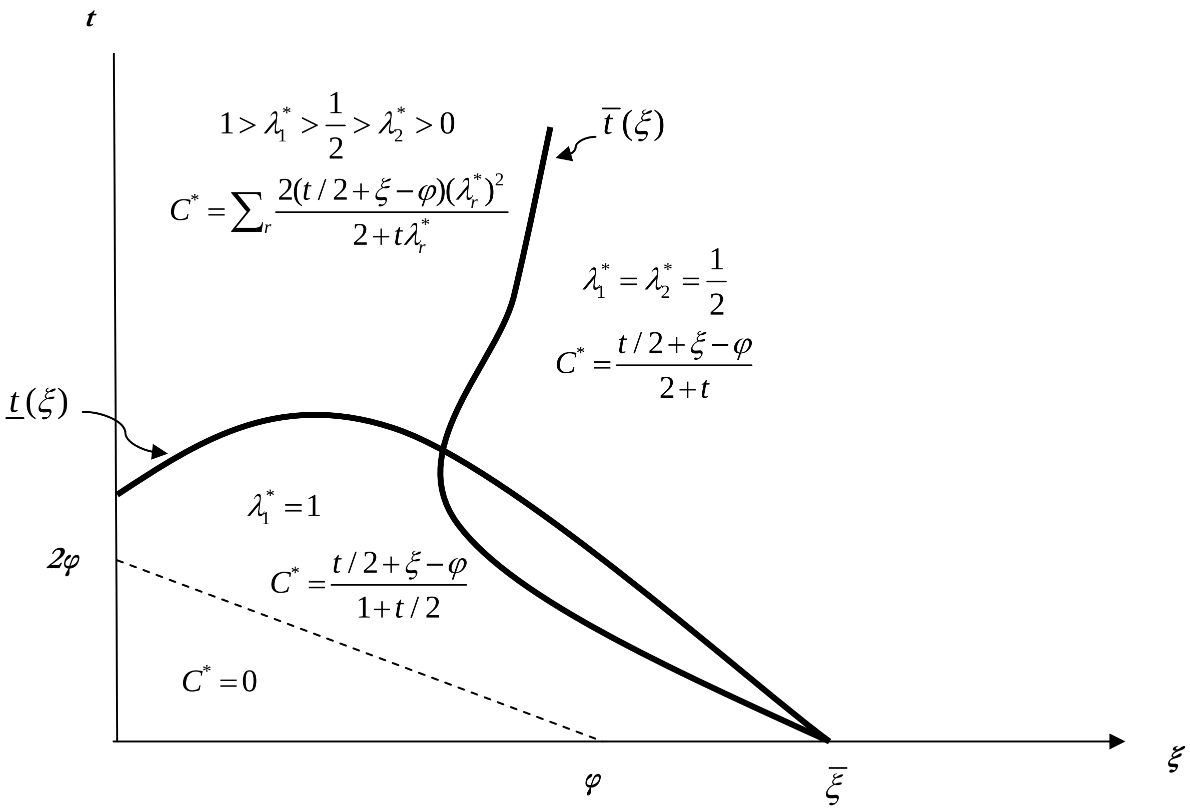
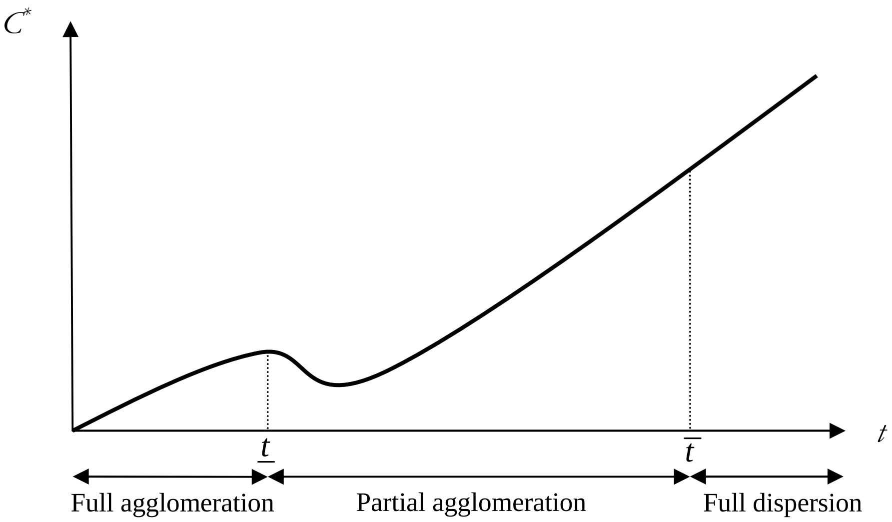


Figure 4: The impact of commuting costs on criminal activities

(a) high φ ($\varphi > \hat{\varphi}$)



(b) low φ ($\varphi < \hat{\varphi}$)

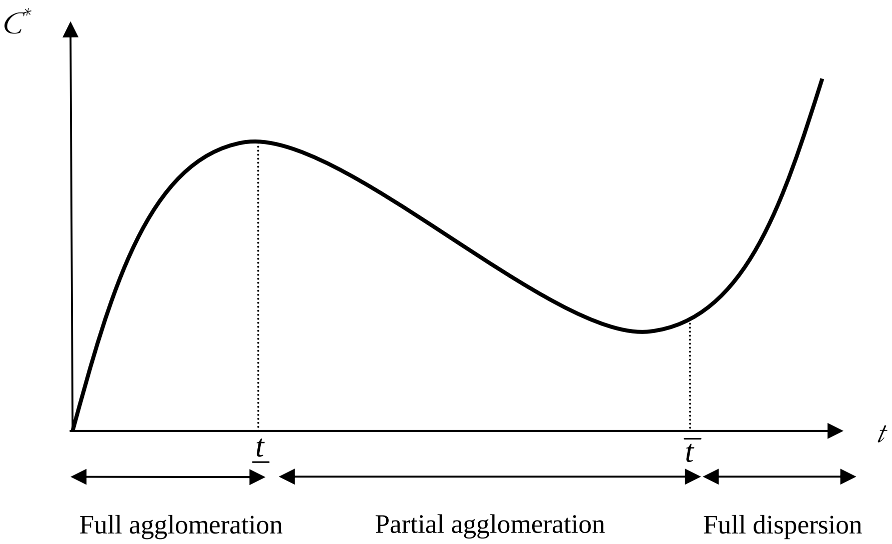


Figure 5: Tax Policy and Criminal activity

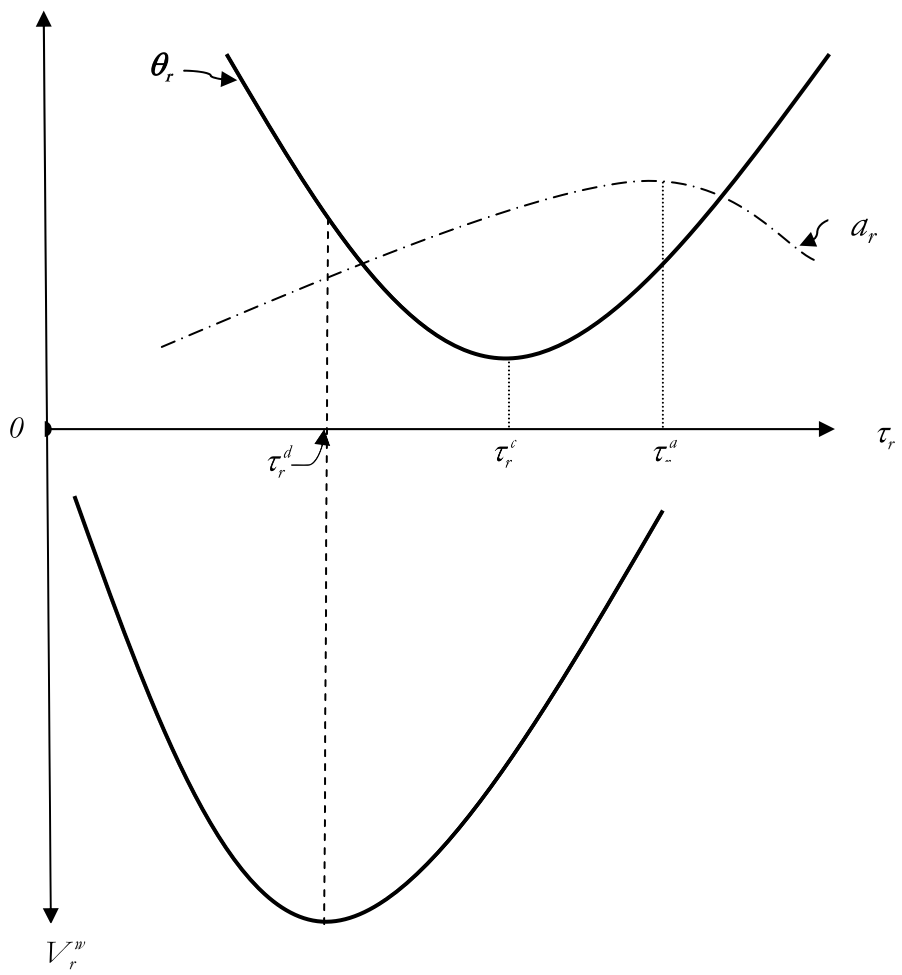


Figure 6: Equilibrium tax rate and commuting costs