SERIE DOCUMENTOS DE TRABAJO

FACULTAD ECONOMÍA Y NEGOCIOS UNIVERSIDAD DE CHILE

DEPARTAMENTO DE ECONOMÍA

SDT 348

IDENTIFYING THE DETERMINANTS OF CRIME OCCURRENCE AND THE DETERRING IMPACT OF POLICE: EVIDENCE ACROSS CHILEAN HOUSEHOLDS

Autores: José Miguel Benavente , Javier Turén

Santiago, Marzo de 2012

Identifying the determinants of crime occurrence and the deterring impact of police: Evidence across Chilean households

José Miguel Benavente[a], Javier Turén[b]

[a]*Bussiness School*
Adolfo Ibañez University
Santiago, Chile
[b]*Department of Economics*
University of Chile
Santiago, Chile

Abstract

We analyze the socio-economic and demographic determinants of crime across Chilean households. In particular, we are interested on the impact that police has on deterring crime. Novel evidence is presented and an instrumental variable correction is performed to avoid the typical reverse causality problem of police on crime. We use multilevel probit and count models to estimate different crime equations. Results indicate that socioeconomic and demographic characteristics have heterogeneous impact on crimes. In terms of police deterrence effects, our results reveal that the number of police officers has no impact on crimes suffered by families (except for burglary) while the true impact of police is determined by the workload that police must face. According to the results, a 10% increase in the workload rate (per 100,000 residents), would raise the crime rates by around 10%.

1. Introduction

The main purpose of this paper is to contribute to the existing literature in criminology through an identification of the determinants of crime occurrence and to address the importance of police in reducing crime occurrence. The existing literature has study these two topics separately.

Empirical attempts to disentangle household from demographic and effort effects on crime levels is scarce. And for less developed countries where crime rates are relatively high, there is no empirical evidence at all. We develop a methodological framework which allows us to perform this task considering a particular middle-income country: Chile.

[1]First draft. Do not cite. This work was partially supported by the Chilean CONICYT Anillo project ACT87 "Quantitative methods in security" (`www.ceamos.cl`). We want to thank Juan Jose Price for his helpful asistance, CONICYT and also other members of CEAMOS for their comments. All errors are ours.

The main contribution of this paper is twofold. Firstly, we attempt to identify the socioeconomic and demographic factors that make a person eligible for a crime and, secondly, we evaluate the deterrence power that police force exerts on crime at a household level. Evidence is presented for a large set of specifications, starting with a measure for all crimes and continuing with results across each type of crime.

2. Related literature

The relationship between police and crime has received a lot of attention over the years. Nevertheless, the existing evidence doesn't show any relationship between crime rates and police force (Cameron (1988) and Eck and Maguire (2000)). Recent scholar responses to this counter intuitive result have focused on methodological issues especially on how the endogeneity problem has been tackled.

This problem arises because governments or local authorities tend to allocate more police officers and resources in cities or areas where crime rates are high. Then a positive correlation is observed between the level of crime and the amount of police resources, leading to inconsistent estimators of the true impact of police DOTACION on crime rates. Levitt's (1997) pioneer study uses an instrumental variable specification in order to tackle this problem and found a negative casual effect from police to crime, even though his conclusion remains a bit controversial (see McCrary's (2002) comment and Levitt's (2002) response).

Recent response to the endogeneity problem is the use of external events in order to isolate the causal effect of police on crime rates. Di Tella and Schargrodsky (2004) rely on a quasi-experimental approach and find a large deterrence effect of police on crime in Buenos Aires. These authors use the terrorist attack on the main Jewish center in Buenos Aires as an external event, a consequence of which was that more police officers were deployed to this institution. They use this event to isolate causality effects of police on crime, showing that motor vehicle theft decreased significantly near the main Jewish center compared to what happened several blocks away where no extra police officers were deployed. In the same line, Draca, Machin and Witt (2008) use the terrorist attack that took place in London in 2005, after which authorities decided to increase the number of police officers in central London. The authors show how during the time of the redeployment the crime decreased substantially in central compared to outer London.

While the previous studies propose proper corrections to the endogeneity problem, the existing evidence about the impact of police on crime rates is only at a city-level basis. The importance of police in reducing the occurrence of crimes committed at a household level still remains unclear. To address this question, it is important to bear in mind that criminology is basically a multilevel challenge. As suggested by Nieuwbeerta et al. (2008) the likelihood of being a victim of a crime is directly influenced by socio-economic, neighborhood and demographic characteristics as well. On the other hand, Buonanno and Montolio (2008) show that lagged crime rate, clearance rate, urbanization rate and the ratio of foreigners to total population show a positive correlation with crime rates after controlling for other individual characteristics.

Multilevel models have gained popularity since they may disentangle individual from household and more aggregate effects including police efforts. Par, Felson and Ouimet (2007) look at how the type of crimes and the characteristics of the neighborhood in which they occur affect the likelihood that they will be cleared by the police. Using a multilevel model these authors show how police is more effective in preventing crimes in small communities rather than large urban areas and also in communities with significant level of poverty. Also Rosenfeld, Fornango and Rengifo (2007) propose a multilevel procedure to analyze the effects of order maintenance policing on violent crime trends in New York City. They found a modest impact of this new policy with small crime-reduction effects only in homicide and robbery rates.

Current research on the effect of police on crime rates has to deal with the above mentioned problems. On the one hand, to ensure that changes in police force do not respond to changes in crime activity (endogeneity problem). On the other hand, to isolate the effect of aggregate and police effort variables from those that may affect crime on an individual basis, in particular, those related with the amount of police officers that is necessary to significantly affect the fear of crime. So far scholars have relied on quasi experimental designs or long time series to tackle the first problem; we instead rely on an instrumental variable approach. We address the second problem by including in the econometric analysis data at the household level; this is a second source of novelty in the empirical analysis of crime and its determinants in the Chilean context. But before presenting our methodological framework, we describe the data used in the next section.

3. The Data

The data used in this paper come from the National Public Safety Survey issued by the Chilean National Institute of Statistics during 2003. This survey has 15,508 observations of Chilean households and is nationally representative[2]. The survey asks questions about fear of crime (feeling of insecurity), victimization in the household and within its members, types of crimes suffered, reaction

[2]The survey has observations of 75 *large* provinces (in terms of inhabitants) across the thirteen Chilean regions.

towards crime and perception about delinquency. This information is complemented with socio-economic characteristics of each respondent, such as age, gender, total income of the household, and level of education.

We complement this database with information coming from the 2002 National Census. This allows us to include demographic characteristics for each province, such as, population, surface, population density and urban areas. On top of that we add information related to the number of sworn police officers, number of arrests and crime reports for every province. This information comes from the Chilean Police Department and the Ministry of Interior. Thus, after combining all these databases, we have information across individuals, households and provinces.[3]

A first look at the data presented in Table 1, shows an heterogeneous pattern in terms of crime frequency. On the one hand, larceny and theft represent almost 40% of the total crimes that occurred in Chile during 2003 while rape, followed by motor vehicle theft, are the less common crimes. This heterogeneity may be explained - following Becker (1977), by differences in expected outcomes related with types of crime. But it could also be related to household socioeconomic as well as individual characteristics that affect the likelihood of being a victim. As previously discussed, police efforts may also play a role. Deterrence could have heterogeneous impacts depending on the type of crime. For example, rape figures could be inelastic to the number of policemen compared to assaults and robberies that normally occur outside residencies. But more importantly, figures may hide those who have been offended in more than one occasion, an issue that we discuss in the following section.

In order to have a better understanding of these linkages, in Table 2 we present a detailed description of the available information. For comparability purposes we have defined several dummy variables where a value of 1 is assigned if the interviewed person claims that she or he has suffered a crime and 0 otherwise. Then a vulnerability index of 0.27 means that 27% of interviewed people say that during 2003 they have suffered at least one crime. Figures presented in Table 2 show that larceny, followed by burglary and theft from a motor vehicle, are the most frequent crimes.[4] These numbers are very high since they suggest that at least for these crimes almost 10% of the population has suffered at least one of them.

But what is more striking is that there are people who declare having been attacked on more than one occasion. By using an cumulative index, which takes into account whether the same respondent has been a victim of a crime more

[3]For a detailed description of the used databases, see Benavente and Turen (2010?)

[4]The difference with the previous table is that here we are not weighting for the number of crimes that the same person has suffered.

than once, Table 2 shows that there are people who have suffered up to six crimes during the year 2003. In the case of crimes related with property some people claim having been offended up to three times, while in the case of violent ones, some of them suffered up to four times during the same year. Therefore, aggregate figures may hide the fact that a person could have been offended more than once as the impact of increasing police force may be different when considering this cumulative situation.

The evidence coming from the survey clearly shows that certain types of crimes occur more frequently than others. Apparently each type of person has specific and personal characteristics that make them more vulnerable to certain types of crimes; this is precisely one of the main focus of this paper. Table 1 shows the types of crime and the frequency of occurrence displayed in the survey.

The first three crimes of table 1 are the so-called *property crimes*. Among them are: motor vehicle theft, theft from motor vehicle (thefts of motor vehicle parts and accessories) and burglary. This group of crimes accounts for almost 34% of total crimes. The remaining crimes can be classified as crimes that occur to people, as mentioned the largest number of cases are concentrated in this group. The most frequent crime is larceny (40%) followed by theft from motor vehicle (19%), burglary (14%) and robbery (14%). The diversity of crimes occurrence suggests the existence of different factors that make a certain crime more likely to occur than others, and also that the deterrence effects of the police force might be heterogeneous across different types of crime.

We use these data to define two measures of crime. The first measure can be referred to as a Vulnerability Index, that is the crime rate at an aggregate (regardless of the type of crime) and disaggregate level (separated by type of crime). For this purpose, the focus is the most reported crimes, leaving aside Embezzlement, Bribery and Rape. The second variable is a count variable, which indicates the number of times a certain person has been a victim of a crime. These two measures of crime are listed in table 2 and are going to be considered as dependant variables in the econometric models that are shown in the following section.

Characteristics at the individual level along with characteristics at the province level are also presented in table 2. Variables at the individual-household level include education, age, sex (1 if male, 0 if female), participation in the labor market (1 if participates, 0 otherwise) and the logarithm of the mean income level of the household where the person lives. The province level variables are the size (in terms of population), a dummy for type of area (1 if urban, 0 if rural), the percentage of families with low income, the population density, the percentage of males between 15-30 years old (Buonanno and Montolio, 2008) and two variables to identify the importance of police in deterring crime: the number of police officers in each province and the ratio between the number of crimes and the number of police officers in each province, i.e. Workload. (These last two variables are measured per 100,000 habitants). These are going to be the independent variables of our econometric specifications.

4. Econometric specification

4.1. Reverse causality problem

Among the province level variables, the number of police officers in each area is certainly one of the factors we are most interested in. Unfortunately, the decision of how many police officers are allocated to a certain area is not random. Moreover, there exist a positive correlation between the number of police officers and the number of crimes in a city or province. In other words, communities with high crime rates tend to have more police resources than others. The positive correlation is clear: more police are deployed in areas with increasing crime rates, failing to fulfill the required orthogonality condition between these two variables, obtaining biased and inconsistent estimators of the coefficient of interest. This problem is known as the "reverse causality" or "endogenous regressor", and is quite common in this type of studies.

To correct for the simultaneity between police force and crime we propose an instrumental variable approach. With at least one variable (instrument) that affects the number of police officers in each area, but does not affect crime rates, we can partially tackle this problem. The search for this kind of variable needs a full understanding of how the Chilean Police Department determines the number of police officers in each province. In 1998, the Chilean Police started a plan to strengthen the police efficiency, by focusing the police resources on prevention and surveillance. The plan proposed a sector monitoring system, where the surveillance of each sector is responsibility of a specific police headquarter. The amount of resources allocated to each headquarter (including the number of police officers) is determined by the demand of police services that each area has. The demand is calculated following three main points: prevention, police procedures and supervision[5]. The prevention tasks are related to variables such as: the size (in terms of population), the squared miles and the crime rate of each city or province. The police procedures include those referred to arrests, crime reports, car accidents occurred in the community, among others. Finally, supervision is related to reach the highest level of coverage and surveillance of the zone, according to the existing resources.

Following the previous description, an equation to determine the number of officers in each province for the year 2003, is proposed. The independent variables are: size, square miles, a dummy that is 1 if the area is urban and

[5]The Chilean Police Department uses two additional criteria to determine demand: Extraordinary and Warrants services, but we don't consider these two variables because we lack the data.

0 otherwise, the crime report rate, the effectiveness rate (the ratio of arrests to crime reports) and the number of car accidents in the city/province. The last three variables are measured for the previous year, i.e. 2002. Thus, with the lagged variables we can identify a response in the number of police officers deployed in each province, related to previous events. The results are presented in Table 3.

Broadly, the results support the principles followed by the Chilean Police Department to determine the allocation of police officers in each province. Moreover, the analysis shows that variables such as the squared miles of the city, being a urban area and the past rates of crime and car accidents help to explain the current numbers of police officers in a certain province. Two interesting results arise: firstly, the relationship between size and the number of police officers is negative and, secondly, the resulting number of police allocated to a certain area is a function of: past effectiveness, the number of crimes that were reported in the past, and the past car accidents.

By using the auxiliary regression we are able to predict the number of police officers in each area, a variable that is then used to correct for the reverse causality problem (Wooldridge, 2005).

4.2. Multilevel methodology

As mentioned, the main purposes of this paper are to identify the individual determinant of crime occurrence and to quantify the impact of the police force in deterring crime. For doing so, we consider separately:

1. The likelihood that a person is a victim of a crime.
2. The number of times that a person has been victim of a crime.

As already mentioned, we are not going to use aggregate level data; instead, we propose a multilevel approach. Specifically, we propose a two level model in which households (level 1) are clustered in provinces (level 2). The household level model is (see Cameron and Trivedi (2005)):

$$y_{ij} = x_{ij}\beta_j + \epsilon_{ij}, \quad \epsilon_{ij} \sim \Gamma(0, \sigma_\epsilon^2) \tag{1}$$

And the level 2 or city/province level model is:

$$\beta_j = \gamma + v_j, \quad v_j \sim N(0, \Sigma) \tag{2}$$

Where i indexes persons and j indexes provinces. The dependent variable y_{ij} is going to be a discrete variable or a count variable depending on the type of estimation performed[6]. The explanatory variables are represented by x_{ij} and

[6]The type of estimation is also going to change the Γ distribution function.

we allow that the regression parameters β vary by group (or cluster). It is clear from the second level equation that we build a model with random intercepts, allowing the intercept to take on different values for each level 2 unit in the data. Moreover, the random effects are characterized by Σ which accounts for the conditional variance across each province. By combining (1) and (2) we obtain:

$$y_{ij} = x'_{ij}\gamma + x'_{ij}v_j + \epsilon_{ij} \tag{3}$$

This is what we refer to as the crime equation. For identification issues, the errors ϵ_{ij} and v_j are assumed IID. Regarding this last issue the estimation of the crime equation (3) can be conducted by maximum likelihood. Besides, thanks to the IID and distribution assumptions about the error terms, the estimation would yield unbiased parameters:

$$y_{ij} \sim \Gamma(x'_{ij}\beta_j, \sigma^2_\epsilon)$$

$$\beta_j \sim N(\gamma, \Sigma)$$

.

4.3. Multilevel Probit Model

Let's define v^*_{ij} a variable that represents the *vulnerability degree* of a person i that lives in city/province j. That vulnerability degree is related to the personal and environmental characteristics of each individual. Clearly, we don't observe this variable, but we do know whether it exceeds some critical threshold ψ (latent variable):

$$v_{ij} = \begin{cases} 1 & v^*_{ij} > \psi \\ 0 & v^*_{ij} \leq \psi \end{cases}$$

If a person was a victim of a crime, then obviously his or her "vulnerability degree" was overcome. Therefore, the dependent variable (v_{ij}) takes the value 1 if the person i claims to have been a victim of any crime in the province or neighborhood where he/she lives and zero otherwise[7]. Finally, the estimation of equation (3) assumes $y_{ij} = v_{ij}$ and Γ takes the form of the logistic cdf.

[7]This last point is very important. The multilevel scheme was proposed mainly for addressing the importance of both socio-economic and demographic variables. In this line, we truncate the binary dependent variable to show only crimes that were committed in the city or province where the person lives. For example, a person can be victim of an assault outside his work. If his work is in another neighborhood (a neighborhood that could have different characteristics compared to his home environment), then it would be wrong to attribute the level two characteristics to that likelihood. In other words, the demographic variables of his home neighborhood are not correlated with the probability of being victim of an assault outside his job. To avoid this source of bias, we truncate the dependent variable to show only crimes that a a person suffered in his home neighborhood or province.

We start by estimating the crime equation considering the full sample of crimes, without making any distinction, and then we present evidence for each type of crime. The variables used in each of the estimations are described in table 2. In order to highlight the importance of the endogeneity problem, we present the results with and without the instrumental variable correction.

The results of the multilevel probit regression are presented in table 4. Each column shows the estimated results where the two measures of police efficiency are contrasted. In columns 1 and 2 the results are calculated without the correction for endogeneity, while columns 3 and 4 come from the specifications that consider the instrumental variable. The evidence suggests that the socio-economic variables (educational level, labor participation and household income) have a positive impact on the likelihood of being a victim of a crime. Also, we see that there is a negative relationship between age and crime, indicating that younger people are more vulnerable to crime, and that gender differences apparently have no effects. In terms of demographic variables, our results support most of the findings of the existing literature (Glaeser and Sacerdote (1999) and Pare, Felson and Ouimet (2007)), finding a positive relationship between size and crime (crime tends to concentrate in large populated cities or provinces). Living in urban and poor areas (poverty is defined in terms of the ratio of low income households to the total number of households) seems to have no effect on crime likelihood. On the other hand, a person that lives in an area that has a high ratio of young males to total population tends to have a higher likelihood of becoming a victim of crime. The evidence about density is not clear, showing a positive impact only in the last specification, whose results are presented in the last column. The differences between the estimated coefficients for the number of police officers in columns 1 and 3, are quite enlightening. If the model had been estimated without considering the endogeneity problem, the conclusion would be that the number of police indeed has a negative and significant deterrence effect on crime. But once we correct for endogeneity, this deterrence effect becomes non significant. The effect of workload on crime likelihood is significant and goes in the same direction in both specifications but it is twice as high in the IV specification. Apparently, the relevant variable to identify the impact of police on crime is not the number of police officers, but the workload that each police headquarter must face in each province.

Apparently, the determinants of the vulnerability degree differ systematically across the different types of crimes. Therefore, the estimated deterrence effects and the efficiency of the police could vary significantly across offences. In order to test for this possibility we estimate a Multilevel probit model for each of the types of crimes presented in table 2. The results allow the comparison of magnitudes, signs and significance of the coefficients, while any potential aggregation bias would be omitted (Cherry and List (2002)). The results are presented in table 5.

Columns 1 and 2 of table 5 indicate that the likelihood of Motor Vehicle

Thefts is affected only by the household income and the percentage of young males in the province. The number of police officers doesn't have a significant impact while the workload rate coefficient is positive although only slightly significant. Theft from motor vehicle is more likely to happen to young men with higher levels of education that belong to high income households (maybe this is simply because they are the ones who own motor vehicles that are of the interest of thieves, unless we accept the possibility that they take less care of their cars). Also, this type of crime is concentrated among big cities or provinces that have a significant presence of young men. The police variable is only significant for the workload rate. For burglary many interesting results arise: homes inhabited by people with higher education are more likely to suffer a burglary (clearly, this result addresses the strong relationship between the education level and the income of a person or family). Burglary tends to be higher in large cities rather than in small provinces. Surprisingly the number of police officers hasn't got any significant impact on crime likelihood, except in the case of burglary; in this case a true deterrence effect seems to exist. Finally, the workload rate seems to have a significant effect on burglary.

With respect to larceny, the level of education appears to have a positive impact on the likelihood of being a victim of such a crime while both age and gender seem to have a negative effect on it. Thus, younger people and women are more exposed to suffer larceny. Also, larceny tends to concentrate in bigger cities or provinces, and interestingly the density variable reports a positive impact in more dense areas. Finally, workload continues showing a significant positive impact on the dependent variable. The results for robbery appear in columns 9 and 10 of table 5. The evidence shows that older people and richest families are less likely to suffer a robbery. This last result is interesting because it supports our intuition regarding the importance of considering each type of crime separately; in fact, while, as already reported, income exerts a positive impact on property crime rates, this result does not hold for robberies. Another interesting result is that the coefficient for the urban/rural dummy is positive and significant, which means that robbery is a more severe problem in urban rather than rural areas. As we can see in column 10, the coefficients for size, density and workload are again positive and highly significant. The last two columns show a negative impact of education and age on the likelihood of suffering assaults. It is remarkable how, in the case of all violent crimes, age always has a negative and statistically significant impact on the crime likelihood. As in the case of robberies, the coefficient on household income is negative and highly significant, and assaults are more likely to occur in bigger areas. Finally, and as in all the other forms of crime, the workload rate has a positive and significant impact in the likelihood of being assaulted.

The results suggest that the number of police officers in a certain area has no effect in deterring crime; the only case where the number of police officers appear to have a negative (although only slightly significant) impact is burglary. Conversely, the evidence supports a strong effect of the workload rate on crime rates; the probability of suffering any type of crime is higher in areas where the proportion of crimes to police officers is high. As already mentioned, individual and demographic characteristics that make certain crimes more likely to occur do not hold across all different types of crime. The heterogeneous effects arise because of the differences between each crime and the situation in which they occurred, therefore, the evidence points to certain characteristics that criminals identify in the victims that lead them to commit certain crimes more frequently than others.

4.4. Multilevel Count Model

Police resources may affect crime in several ways. The impact of police in the likelihood that a person suffers a crime may be just one dimension. Moreover, one may think that the chances of suffering a crime could behave as a random process, where at the end luck is the only reason for being a victim of a crime. In that line, perhaps the impact of police on crimes, is trying to minimize the frequency of occurrence of such crimes. In this section we extend our previous analysis to address this question, using a count multilevel model.

We estimate equation (3) assuming that y_{ij} is now a count variable that represents the total number of crimes that a person has suffered during a period of a year. As we are estimating a count model, Γ should behave as a Poisson cdf. Nevertheless, our dependent variable presents an over-dispersion problem (i.e. the variance is greater than the mean[8]). Consequently, for in order to obtain an unbiased estimator we assumed that Γ takes the form of a Negative Binominal distribution[9].

The estimation procedure considers the same independent variables than before, but unfortunately we are not able to conduct the estimations for each type of crime. This is because the survey only asks individuals whether they have suffered each of the previously mentioned crimes but doesn't ask them whether they have been a victim of the same crime more than once. As a result, the count variable accounts for the total number of crimes that a certain individual

[8]This is a straightforward result that is presented in table 2, where for all the count dependent variables the standard deviation is much greater than the mean.

[9]The negative binomial distribution is the most commonly used to solve for over dispersion. It's quite common that for count variables one has a mixture of Poisson and gamma distributed data. For that reason, the expected average outcome of the resulting distribution follows a Poisson distribution ($\lambda = \mu$) but its variance follows a gamma distribution ($Var(\lambda) = \mu^2/k$). Then, a subset of a gamma mixture of Poisson distribution yields the negative binomial distribution (MacDonald and Lattimore 2010). Therefore, the expected mean of the outcome of counts (Y) follows a Poisson distribution and the variance is equal to the Poisson and gamma distribution :

$$E(Y) = \mu$$
$$Var(Y) = \mu + \mu/k^{-1}$$

has suffered, despite the fact that he/she could have suffered a particular crime more than once. Nevertheless, we are still able to separate the counts between property and violent crimes. The results are presented in table 6.

According to the results, the education, age and size coefficients have the same sign for every specification. The level of education (a very good proxy for the income or wealth) and living in big provinces or areas, increase the number of potential crimes that a person would suffer in his lifetime. The negative impact of age indicates that the frequency of crimes decreases, as the person gets older. The household income is a positive determinant of property crimes while it decreases the frequency of violent crimes. Apparently, higher level of wealth necessarily expose richer families to suffer property crimes such as motor vehicle theft or burglary even though that same wealth allows them to prevent violent crimes more effectively. For example, wealthier families tend to use their private cars for transportation rather than taking the bus or the subway, avoiding typical environments for robbery or larceny. Interestingly, only density appears to affect (positively) the frequency of violent crimes while other demographical characteristics such as urban areas or the percentage of families with low income don't have a direct impact.

The number of police officers deployed in a certain area doesn't appear to have any significant impact on the frequency of crimes, contrary to the workload rate, which shows a positive and significant impact. Since in this specification the police variables are in logarithm the estimated coefficients can be interpreted as semielasticities. A 10 percent increase in the workload rate (every 100,000 inhabitants) raises the number of crimes committed to a person by around 10 percent. This result holds both for total crimes as property and violent crimes considered separately. The presented evidence strengthens our conclusions about the real deterrence effect of the police force.

5. Conclusions

In this paper we have developed a methodological framework to identify the individual and demographic determinants of crime occurrence in Chile and to quantify the impact of the police force in deterring crime. This framework has allowed us to disentangle individual from household and more aggregate effects including police effort levels. We have also relied on an IV approach to tackle the problem of reverse causality between crime rates and the number of officers allocated to a certain area.

We have defined two measures of crime: a Vulnerability Index, and a count variable that measures the number of times a person has been a victim of a crime. In our search for explanations to the observed variation in these variables across different types of crimes we have taken a look at individual characteristics as well as variables measuring police effort.

The evidence suggests that vulnerability is higher for those living in large populated cities and areas where the ratio of young males to total population is high. On the other hand, gender differences and the distinction between urban and rural areas seem to have no effect on crime likelihood. Regarding the impact of police on crime, and once we control for endogeneity, it appears that the relevant variable to identify the effect of police force is not the number of police officers, but the workload that each police headquarter must face in each province. This is valid for both the likelihood of crime and the frequency of certain types of crime.

The important points to highlight are: firstly, controlling for reverse causality between police force and crime rates is extremely important if we want to obtain consistent estimators, and, secondly, when studying the determinants of crime it is important to control for socioeconomic and demographic factors at the individual and household level; crime is by nature a multilevel phenomenon, so to rely only on aggregate data is not enough.

6. References

[1] Buonanno, P. and Montolio, D. 2008. "Identyfing the socio-economic and demographic determinants of crime across Spanish provinces", *International Review of Law and Economics*, 28, 89-97.

[2] Cameron, S. 1998. "The Economics of Crime Deterrence: A survey of Theory and Evidence",*Kyklos*, 41, 301-323.

[3] Cameron, C. and Trivedi, P, 2005. "Microeconometrics", Cambridge University Press, 845-8.

[4] Cherry, T. and List, J, 2002. "Aggregation bias in the economic model of crime", *Economic Letters*, 75, 81-86.

[5] Di Tella, R and E. Schargrodsky, 2004. "Do Police Reduce Crime? Estimate Using the Allocation of Police Forces After a Terrorist Attack", *American Economic Review*, 94, 115-133.

[6] Draca, M, Machin, S and Witt, R, 2008. "Panic on the streets of London: police, crime and the July 2005 terror attacks", *CEPDP*, 852. Centre for Economic Performance, London School of Economics and Political Science, London, UK.

[7] Eck, J. and E. Maguire, 2000. "Have Changes in Policing Reduced Violent Crime: An assessment of the Evidence", in A.Blumstein and J. Wallman (eds) *The Crime drop in America*, Cambridge University Press: New York.

[8] Glaeser, E. and Sacerdote, B, 1999. "Why Is There More Crime in Cities?", *Journal of Political Economy*, 107, 225-258.

[9] Levitt, S. 1997. "Using Electoral cycles in Police Hiring to Estimate the Effects of Police on Crime", *American Economic Review*, 87, 270-290.

[10] Levitt, S. 2002. "Using Electoral cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply", *American Economic Review*, 92, 1244-50.

[11] MacDonald, P. and Lattimore, P. 2010. "Count Models in Criminology", *Handbook of Quantitative Criminology*, Springer.

[12] McCrary, J, 2002. "Using Electoral cycles in Police Hiring to Estimate the Effects of Police on Crime: Comment", *American Economic Review*, 92, 1236-43.

[13] Nieuwbeerta, P. McCall, PL. Elffers, H. and Wittebrood, K, 2008. "Neighborhood characteristics and individual homicide risks: effects of social cohesion, confidence in the police, and socioeconomic disadvantege, *Homicide Studies*, 12(1),90-116.

[14] Pare, P-P. Felson, RB. and Ouimet, M, 2007. "Community Variation in Crime Clearance: A Multilevel Analysis with Comments on Assesing Police Performance", *Journal of Quantitative Criminology* ,23(3), 243-258.

[15] Rosenfeld, R. Fornango, R. and Rengifo, AF, 2007. "The impact of order-maintenance policing in New York City homicide and robbery rates: 1988-2001. *Criminology* 45(2), 355-384.

Table 1: Types of Crimes

Type of Crime	Frequency	% of total crimes
Motor vehicle theft	122	1.3
Theft from motor vehicle	1770	18.81
Burglary	1285	13.65
Larceny - Theft	3742	39.76
Robbery	1275	13.55
Assault	845	8.98
Embezzlement	175	1.86
Bribery	175	1.86
Rape	22	0.23
Total	9411	100

Table 2: Descriptive statistics for variables

Variable	Mean	SD	Min	Max
Dependent Variables				
Vulnerability Index	0.27	0.44	0	1
Motor Vehicule Theft	0.01	0.07	0	1
Theft from motor vehicle	0.07	0.25	0	1
Burglary	0.08	0.27	0	1
Larceny - Theft	0.11	0.31	0	1
Robbery	0.04	0.19	0	1
Assault	0.04	0.18	0	1
Total number of crimes	0.35	0.65	0	6
Total number of property crimes	0.2	0.39	0	3
Total number of violent crimes	0.2	0.46	0	4
Independent variables				
Individual level predictors (N=16204)				
Education	10.02	4.09	0	17
Age	42.9	17.7	15	98
Sex (male = 1)	0.45	0.49	0	1
Household Income (Log)	12.88	0.533	11.7	14.74
Participation (participates = 1)	0.501	0.5	0	1
Province level predictors (N = 75)				
Size	153,011	99,302.4	31,516	492,915
Urban (urban=1)	0.73	0.44	0	1
% Families with low income	0.59	0.19	0.059	0.91
Density	3,132.3	4,658.3	0.97	15,667
% Males between 15-30 years old	0.3138	0.0511	0.181	0.473
Police officers (per 100,000 residents)	138	61.6	51	438
Crimes / Police officers (per 100,000 residents)	0.511	0.32	0.06	1.61

Table 3: Auxiliary regression results, Police Officers (2003)

Variables	
Size	-0.000112*
	(-0.00006)
Square Miles	0.00141***
	(0.000217)
Urban	20.94**
	(10.35)
Crime Report Rate (2002)	0.0128***
	(0.00262)
Effectiveness Rate (2002)	15.94
	(24.09)
Car Accidents Rate (2002)	0.909***
	(0.299)
Constant	77.27***
	(15.25)
Observations	75
R^2	0.414

Notes: All the variables except Size, Square Miles and Urban are per 100,000 residents. Robust standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Multilevel probit model for crime ocurrence likelihood (All crimes)

Variables	ML-Probit	ML-Probit	ML-IV-Probit	ML-IV-Probit
Individuals (Level 1)				
Education	0.0345***	0.0347***	0.0343***	0.0344***
	(0.00558)	(0.00557)	(0.00558)	(0.00556)
Age	-0.0107***	-0.0106***	-0.0107***	-0.0107***
	(0.00118)	(0.00118)	(0.00118)	(0.00118)
Sex	-0.0253	-0.0254	-0.0250	-0.0251
	(0.0387)	(0.0387)	(0.0387)	(0.0387)
Household Income (Log)	0.0789*	0.0783*	0.0777*	0.0800*
	(0.0410)	(0.0409)	(0.0410)	(0.0409)
Participation	0.118***	0.117***	0.117***	0.116***
	(0.0397)	(0.0396)	(0.0397)	(0.0396)
Province (Level 2)				
Size (Log)	0.191***	0.350***	0.214***	0.470***
	(0.0639)	(0.0576)	(0.0681)	(0.0630)
Urban	0.0148	0.0541	-0.0330	0.0859
	(0.0975)	(0.0814)	(0.105)	(0.0788)
% Families with low income	0.153	0.226	0.224	0.141
	(0.198)	(0.166)	(0.209)	(0.160)
Density (Log)	0.0160	0.0154	0.0173	0.0276**
	(0.0153)	(0.0129)	(0.0158)	(0.0125)
% Males between 15-30 years old	1.473**	0.922	1.585**	1.086*
	(0.717)	(0.620)	(0.735)	(0.588)
Police officers	-0.00127**		-0.000283	
	(0.000610)		(0.00106)	
Crimes / Police officers		0.648***		1.160***
		(0.107)		(0.170)
Constant	-4.729***	-6.992***	-5.165***	-8.726***
	(0.937)	(0.868)	(1.004)	(0.948)
Observations	15,508	15,508	15,508	15,508

Notes: The variables: Police officers and Crime/Police Officers are per 100,000 residents. Robust standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5: Multilevel IV-Probit model for different crime ocurrence likelihood)

	Property Crimes								Violent Crimes			
Variables	Motor vehicle theft		Theft from motor vehicle		Burglary		Larceny/Theft		Robbery		Assault	
Individuals (Level 1)												
Education	0.0496	0.0510	0.0617***	0.0626***	0.0279***	0.0276***	0.0413***	0.0412***	0.0101	0.0105	-0.0240*	-0.0228*
	(0.0341)	(0.0341)	(0.0100)	(0.0100)	(0.00901)	(0.00899)	(0.00782)	(0.00780)	(0.0128)	(0.0128)	(0.0135)	(0.0136)
Age	-0.00580	-0.00548	-0.00495**	-0.00479**	-0.00241	-0.00235	-0.00690***	-0.00686***	-0.0212***	-0.0210***	-0.0211***	-0.0210***
	(0.00723)	(0.00725)	(0.00208)	(0.00208)	(0.00189)	(0.00189)	(0.00164)	(0.00164)	(0.00275)	(0.00275)	(0.00294)	(0.00294)
Sex	0.0870	0.0830	0.178***	0.176***	-0.0470	-0.0469	-0.116**	-0.116**	0.122	0.120	0.105	0.101
	(0.231)	(0.231)	(0.0677)	(0.0677)	(0.0628)	(0.0628)	(0.0537)	(0.0537)	(0.0850)	(0.0850)	(0.0925)	(0.0925)
Household Income (Log)	0.741***	0.758***	0.696***	0.697***	0.0186	0.0242	0.0304	0.0335	-0.203**	-0.198**	-0.552***	-0.554***
	(0.253)	(0.254)	(0.0760)	(0.0759)	(0.0661)	(0.0659)	(0.0573)	(0.0571)	(0.0892)	(0.0891)	(0.0898)	(0.0899)
Participation	0.188	0.183	0.0913	0.0898	0.0404	0.0361	0.0642	0.0629	0.0853	0.0812	0.0469	0.0502
	(0.242)	(0.243)	(0.0704)	(0.0704)	(0.0643)	(0.0643)	(0.0549)	(0.0549)	(0.0877)	(0.0877)	(0.0945)	(0.0946)
Province (Level 2)												
Size (Log)	0.0733	0.414	0.257**	0.456***	0.287***	0.623***	0.113	0.360***	0.154	0.431***	0.240**	0.347***
	(0.253)	(0.291)	(0.1000)	(0.109)	(0.0906)	(0.0963)	(0.0865)	(0.0884)	(0.103)	(0.107)	(0.0992)	(0.116)
Urban	-0.198	-0.0274	-0.221	-0.0241	-0.131	-0.112	-0.0763	0.0242	0.437***	0.515***	-0.0463	0.187
	(0.406)	(0.373)	(0.156)	(0.138)	(0.141)	(0.117)	(0.133)	(0.111)	(0.168)	(0.152)	(0.154)	(0.143)
% Families with low income	-0.158	-0.400	0.166	-0.0422	0.292	0.352	0.0735	-0.00214	-0.0144	-0.0457	0.619*	0.399
	(0.795)	(0.761)	(0.305)	(0.276)	(0.284)	(0.245)	(0.264)	(0.224)	(0.315)	(0.279)	(0.322)	(0.311)
Density (Log)	-0.0480	-0.0140	-0.0233	-0.00749	-0.00843	-0.000370	0.0422**	0.0533***	0.0495**	0.0619***	-0.00724	-0.00192
	(0.0557)	(0.0568)	(0.0226)	(0.0214)	(0.0213)	(0.0189)	(0.0198)	(0.0175)	(0.0226)	(0.0208)	(0.0219)	(0.0223)
% Males between 15-30 years old	6.154**	5.227*	2.358**	1.832*	1.709*	1.094	1.595*	1.097	1.081	0.361	0.588	0.149
	(2.764)	(2.681)	(1.075)	(1.016)	(0.977)	(0.877)	(0.921)	(0.818)	(1.054)	(0.967)	(1.050)	(1.036)
Police officers	0.00115		0.00173		-0.00270*		-0.000490		-0.00116		0.00329**	
	(0.00372)		(0.00149)		(0.00150)		(0.00134)		(0.00150)		(0.00142)	
Crimes / Police officers		1.737**		1.116***		1.250***		1.108***		1.155***		0.788***
		(0.801)		(0.295)		(0.254)		(0.237)		(0.294)		(0.299)
Constant	-17.92***	-22.73***	-16.13***	-18.74***	-6.500***	-11.38***	-4.616***	-8.083***	-2.608	-6.523***	0.960	-0.139
	(4.586)	(5.017)	(1.599)	(1.693)	(1.432)	(1.479)	(1.314)	(1.327)	(1.748)	(1.773)	(1.725)	(1.867)
Observations	15,508	15,508	15,508	15,508	15,508	15,508	15,508	15,508	15,508	15,508	15,508	15,508

Notes: The variables: Police officers and Crime/Police Officers are per 100,000 residents. Robust standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: Multilevel IV-Count model for frequency of crime ocurrence

Variables	Total		Property		Violent	
Individuals (Level 1)						
Education	0.0302***	0.0304***	0.0398***	0.0402***	0.0217***	0.0220***
	(0.00448)	(0.00447)	(0.00629)	(0.00627)	(0.00578)	(0.00576)
Age	-0.00815***	-0.00809***	-0.00317**	-0.00308**	-0.0118***	-0.0117***
	(0.000939)	(0.000939)	(0.00131)	(0.00131)	(0.00122)	(0.00122)
Sex	0.00660	0.00689	0.0518	0.0515	-0.0141	-0.0146
	(0.0304)	(0.0304)	(0.0430)	(0.0430)	(0.0391)	(0.0390)
Household Income (Log)	0.0544*	0.0560*	0.306***	0.308***	-0.134***	-0.131***
	(0.0328)	(0.0327)	(0.0469)	(0.0469)	(0.0412)	(0.0410)
Participation	0.0612*	0.0597*	0.0622	0.0600	0.0508	0.0507
	(0.0313)	(0.0313)	(0.0445)	(0.0444)	(0.0401)	(0.0401)
Neighborhood (Level 2)						
Size (Log)	0.195***	0.398***	0.260***	0.496***	0.152**	0.334***
	(0.0577)	(0.0519)	(0.0731)	(0.0717)	(0.0659)	(0.0604)
Urban	-0.0526	0.0606	-0.205*	-0.0824	0.00449	0.153**
	(0.0919)	(0.0650)	(0.118)	(0.0879)	(0.104)	(0.0772)
% Families with low income	0.173	0.0888	0.196	0.108	0.134	0.0220
	(0.173)	(0.131)	(0.223)	(0.180)	(0.195)	(0.154)
Density (Log)	0.0119	0.0228**	-0.0128	-0.000687	0.0309**	0.0422***
	(0.0126)	(0.0101)	(0.0161)	(0.0139)	(0.0142)	(0.0118)
% Males between 15-30 years old	1.335**	0.930*	1.922**	1.418**	0.992	0.644
	(0.595)	(0.480)	(0.762)	(0.656)	(0.669)	(0.557)
Police officers (Log)	0.0321		0.0481		0.106	
	(0.156)		(0.199)		(0.172)	
Crimes / Police officers (Log)		0.984***		1.093***		0.966***
		(0.140)		(0.192)		(0.162)
Constant	-2.969**	-5.695***	6.760	2.328	0.100	-2.031**
	(1.264)	(0.779)	(232.3)	(155.3)	(1.454)	(0.944)
Observations	15,508	15,508	15,508	15,508	15,508	15,508

Notes: The variables: Police officers and Crime/Police Officers are per 100,000 residents. Robust standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.