PARDEE RAND GRADUATE SCHOOL

This PDF document was made available from www.rand.org as a public service of the RAND Corporation.

Jump down to document ▼

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world.

Support RAND

Browse Books & Publications

Make a charitable contribution

For More Information

Visit RAND at www.rand.org

Explore Pardee RAND Graduate School

View document details

DISSERTATION

Police Effectiveness

Measurement and Incentives

Ben A. Vollaard

RAND PARDEE RAND GRADUATE SCHOOL

Summary

Although providing safety to citizens is one of the most important tasks of the government, policy makers tend to have very limited information on how to achieve this goal. This dissertation addresses the policy question how to assess police performance and how to use performance assessments to strengthen incentives for improvement of police performance.

The first part of this dissertation presents an econometric analysis of the effect of police on crime and public disorder. The police funding formula is modeled to identify the endogenous variation in police levels across municipalities and over time. The remaining variation is used to identify the effectiveness of the police. The analysis is based on individual data from a Dutch victimization survey unique in its size, duration and scope. By using victimization data, this study provides evidence on the effects of police on public disorder rather than crime only, circumvents measurement error common to police statistics, and controls for both individual and municipality characteristics. The estimation results suggest significantly negative effects of police on property crime, violent crime and public disorder. Additionally, experience of public disorder is found to be mostly a characteristic of the municipality in which someone lives, with little variation across individuals in a municipality, whereas victimization of property crime and particularly violent crime varies across individuals rather than municipalities. The findings also provide evidence that greater police protection allows people to move around more freely, which is an additional benefit of higher police levels not reflected in a decline in victimization rates.

The second part discusses the tradeoffs related to two alternative approaches to holding police forces accountable to results. Since many countries have decentralized the organization of their police, the analysis focuses on the interaction between the national government and the regional police forces. The first approach to performance assessment, objective performance assessment, defines 'good performance' in advance. There is a clear rule that defines when a police force receives a performance reward. In the second approach, subjective performance assessment, subjective judgment closes the gap between performance data and a judgment of a force's contribution to the objective of police work. In both approaches, much of the same information could be used, the crucial difference is whether the reward is based on a rule determined ex ante or on subjective judgment ex post. Given the multidimensional nature of police work and the prevalence of non-discrete outcomes, the first approach proves to be difficult. The government faces an unfavorable tradeoff between power of incentives on the one hand and distortion of effort and discretionary power on the other hand. Subjective performance assessment may provide a way out by alleviating concerns about distortion of effort while leaving more room to respond to (changing) regional conditions. A case study of the Netherlands illustrates the unfavorable tradeoff related to objective performance assessment.

Contents

Acknowledgements

My mother's returning question 'Did you get some sun, my son?' helped me to keep a healthy balance between work and leisure. My parents and five brothers helped me to fully enjoy the hours away from work with hikes, dinners, concerts and drinks so as to be full of energy when returning to the books. The discussions with and encouragements by *Merel* were a great stimulus to bring this work to a close. I am greatly indebted to all of you.

Ben Vollaard, Scheveningen, January 2006.

Chapter 1. Introduction

I. Rationale

Although fighting crime and maintaining public order are two of the most important tasks of the government, policy makers tend to have very limited information on how to achieve this goal. The political debate about the government's imperative 'to do something about crime' tends to revolve around a few readily available statistics like the number of police officers. After all, a straightforward way of signaling that fighting crime is a political priority is to channel more resources towards the police, the most visible of public organizations that is supposed 'to protect and serve' us.

An ever increasing amount of public resources is spent on the police. The police budget grew from €150 per 1,000 population in 1985 to €220 in 2005 (adjusted for inflation). Some €3.6 billion was spent on the Dutch police in 2005. Reliable information on the way these resources are being spent is not available, and hard evidence on how much of an impact police work is having on safety is virtually nonexistent. Whether a greater number of police personnel results in less crime and public disorder – or is even the best way of achieving greater safety, these questions are seldom addressed.

An example illustrates the current state of affairs. In the absence of good indicators of police performance, at the very least we would like to know how police resources are being allocated to different activities like criminal investigations and street patrol. Such information is simply not available, which shows that matters of resource allocation are not on the top of the decision makers' minds. Even when asked by members of Parliament, the Interior Secretary could not answer questions about the time police officers spend patrolling the streets. Not surprisingly, research shows that resource management within the police is based on *ad hoc* or mechanistic rules rather than a systematic analysis of which activities are most needed (Netherlands Court of Audit, 2003a). For instance, the level of crime in a region does not affect the share of resources spent on crime fighting by the police force in that region. Consequently, the chance of getting caught as a criminal is higher in low-crime regions than in high-crime regions (Vollaard, 2004). The lack of even the most basic statistics about police activities indicates the low importance attached to understanding and improving police performance. After all, insight in performance and its determinants is a necessary first step towards improving performance.

The rise of a political voice for widely-shared public concerns about safety in 2001, Pim Fortuyn, has forced the government to leave its 'hands off approach' to the police. The national government faces the challenge of making a gesture beyond just increasing police resources. Of course, the lack of insight in police performance provides a shaky basis for any

policy initiatives. The national government is in a disadvantage given the information advantage of the police forces and their habit of following their own course. After a long period of negotiations with the chiefs of police, the Balkenende administration introduced performance contracts for each of the police forces in 2003. The contracts include performance targets for a few items that could be reliably measured – mostly thanks to third parties like the prosecution council – such as the number of fines and criminal cases. If the forces meet the targets by 2006, they will receive a modest financial bonus. At this time, it is unclear whether the incentive scheme is – or even could be – effective.

Although the new financial incentives are a leap in the dark, awareness of the need for reform and willingness to actually change things are the first necessary steps towards a more accountable and better performing police. This dissertation is aimed to contribute to the foundations of this policy of reform.

II. Research questions

This study addresses the current demand from policy makers within the national government to measure and improve police performance. More specifically, we address two research questions:

> *How can we assess performance of the police forces?*

> *How can the national government use performance assessments to strengthen incentives for improvement of police performance?*

Clearly, these two questions are strongly related: to be able to provide incentives, the government needs to have some idea of how well the police are doing and whether their performance is improving or not. Below, we describe the research methods used to address these two research questions.

III. Method

The method of choice to assess police performance depends on the level of analysis. At the national level, police effectiveness can be identified with the help of econometric analysis (see Levitt, 2002 for example). Major challenges are to single out exogenous variation in police levels, and to use crime data that is comparable across localities and over time. We model the police funding formula to identify the endogenous variation in police levels, and use the remaining variation to identify the effect of police on crime. As source of crime data, we use individual data from a Dutch victimization survey unique in its size, duration and scope.

At the regional level, the number of unknowns exceeds the number of knowns; econometric analysis is not possible. Here, the challenge is not empirical, but conceptual: how do we get from information about police activities to an overall assessment of police performance? Based on the economics literature, we discuss the tradeoffs between alternative approaches of assessing police performance – and their consequences for incentive schemes. To provide a practical understanding of the issues, we conduct a case study on the Dutch experiences with assessing and rewarding police performance.

Thus, addressing the research questions implies two separate steps. First, chapter 2 includes an econometric analysis of the effectiveness of the Dutch police at the national level. Second, chapter 3 presents a conceptual analysis of the approaches to assess and reward police performance at the regional level.

IV. Scope

To keep the analysis tractable, we had to decide which issues were truly essential to address the research questions and which issues are outside the scope. There are several areas that would justify separate studies.

First, in the empirical analysis, we do not explore what actually goes on inside police departments. We take on the challenge to estimate the effect of police on crime. A next step is to identify the effectiveness of specific policing methods (see Wittebrood and Van Beem, 2004 for a meta study of the existing empirical evidence within the Dutch context). Second, a cost-benefit analysis of the expenditures on the police is beyond the scope of this study. Valuing the benefits of lower crime and public disorder demands a separate research effort (see Dubourg, Hamed and Thorns, 2005 for instance). Third, data limitations force us to focus on the effects of policing on crime, public disorder and precautionary measures in our empirical analysis. We leave other police activities like traffic management for future research. By the same token, we do not look into murder and crimes against businesses and non-resident tourists, because these crimes are not included in the survey. Fourth, our goal is to identify the impact of police on crime, not to explain changes in the crime rate. We only include control variables such as unemployment in the analysis to prevent a bias in estimating police effectiveness – not to analyze which factors have contributed to the decline in victimization of crime.

V. Contribution

This study contributes to both the scientific literature and to policy making. Chapter 2 contributes to the empirical literature on police effectiveness in two ways. First, the analysis is based on a victimization survey as source of crime data rather than on police statistics, the default in the literature. By using victimization data, we provide evidence on the effects of

police on public disorder and precautionary measures rather than 'hard crime' only, we circumvent measurement error common to police statistics, and we are able to control for both individual and municipality characteristics. Moreover, empirical analyses based on non-American data are virtually non-existent. Additionally, we employ a research design that has not yet been used in the empirical literature on police effectiveness. Rather than directly singling out exogenous variation in police levels, as in the instrumental variable approach, we identify endogenous variation related to the variables in the police funding formula, and use the remaining variation to identify the effect of police on crime.

Chapter 3 contributes to the literature on management of the police by focusing on what the police should do rather than what the police should not do. Thus far, when design of incentives for the police was subject of analysis, policy analysts tended to focus on corruption and racial profiling, rather than on the desirable outcomes of police work such as fighting crime. With this study, we aim to contribute to redressing this imbalance in the literature.

We contribute to policy making in two ways. First, we evaluate the effectiveness of an important part of the public sector: the police. Although recent foreign research repeatedly shows that the police are effective in lowering crime, these findings have not found their way to the policy debate in the Netherlands. Conducting an evaluation with Dutch data is of particular relevance since police personnel per capita in the Netherlands has been expanded by an unprecedented 20 percent over the last ten years. The Kok I and Kok II administrations promised to make our country safer, now the data and methods are available to evaluate the effects of this policy initiative.

Second, the analysis of incentives for police forces provides inputs for the evaluation of another major policy initiative: the introduction of performance contracts in 2003. Ticking off whether a police force meets a number of performance targets may be the most straightforward, but not the most productive way of holding the police accountable to results. In our analysis, we identify a number criteria that can be used to tradeoff alternative approaches to assessing and rewarding police performance. The conceptual framework allows us to analyze the value added of performance contracts versus other designs of incentive schemes. Since the contracts have only been concluded three years ago, it is still too early to draw hard conclusions about their impact.

Chapter 2. Estimating police effectiveness with individual victimization data

I. Introduction

Methodological advances and the crime drop of the 1990s observed in many countries have spurred a renewed interest in identifying the causal effect of police on crime. New research designs are employed to break through the simultaneity between police and crime levels and to address omitted variable bias. Di Tella and Schargrodsky (2004) and Klick and Tabarrok (2005) use shocks to police presence related to terrorist attacks and terrorist alerts to identify police effectiveness. Corman and Mocan (2000, 2005) use high frequency data to escape the (slow) adjustment in allocation of police resources to crime rates. Levitt (2002) employs an instrumental variable design to identify changes in police levels that are not related to changes in crime rates. These recent studies consistently find a substantial negative effect of police on crime recorded by the police.

So far, the literature is exclusively based on police statistics as source of crime data. In this chapter, we switch the perspective from offenders to victims of crime, using data from the Dutch Victimization Survey for the period 1996-2004. During this period, the Netherlands experienced unprecedented growth in police levels. Using individual victimization data to identify the effect of the growth in police numbers has a number of advantages. First, whereas recorded crime statistics indicate the formal aspects of police performance over time and across jurisdictions, the victimization survey also reflects police activities that primarily rely on informal policing methods. The survey includes many dimensions of 'quality of life' in the neighborhood not covered in police statistics, including types of public disorder such as vandalism, littering and graffiti. By using victimization data, we provide evidence on the effects of police on people's experience of quality of life rather than crime only. As Weisburd and Eck (2004) point out in their review of the existing empirical literature, no evidence on the effect of police on public disorder is available to date.

Additionally, we circumvent several types of measurement error common to police statistics that may result in biased estimates. Recorded crime statistics are subject to changes in reporting and recording behavior that are hard to control for with cross-sectional and time-series dummy variables (Dryden Witte and Witt, 2001). The effect of more police officers on the percentage of crimes reported may be limited (Levitt, 1998), but there are many other factors that may bias both cross-section and time-series analysis based on recorded crime. Sources of bias in recorded crime include changes in policing priorities and administrative practices, the introduction of new information technology within the police organization, and changes in citizen concerns about specific crimes (Wittebrood and Junger, 2002). Only studies analyzing the effect of police for a short period of time in a single city are not likely to

suffer as much from measurement error in police statistics (for example, see DiTella and Schardgrodsky, 2004 and Klick and Tabarrok, 2005).

The victimization survey also includes a rich and easily accessible source of individual background characteristics of respondents (both victims and non-victims). When using data on offenders, individual characteristics can only be collected through painstaking efforts related to combining police statistics with sources such as school records and draft registration records. Because of the costs related to such data collection, the resulting data tend to be limited to one locality, such as the Philadelphia cohort of young men studied by Tauchen et al. (1994). As we will show in the analysis, individual characteristics are particularly important for explaining violent crime.

We apply a novel approach to addressing endogenous variation in police levels in non-experimental data. There are two ways of identifying variation in police levels that is exogenous to crime. In the conventional approach, as applied by Levitt (2002), an instrumental variable related to policy (police levels) but not to the outcome of interest (crime) is used to isolate exogenous variation. Our approach works the other way round. In the Netherlands, the police budget is distributed across municipalities on the basis of a funding formula. The formula encompasses a number of municipality characteristics such as the number of immigrants and housing density. By explicitly modeling the variables that are known to be part of the police funding formula, we are able to distinguish between endogenous and exogenous variation in police levels. The endogenous variation is related to the variables from the funding formula. The remaining variation is exogenous to the relation between police and crime and can be used to identify the effect of police on crime and public disorder.

The rest of the chapter is organized as follows. Section II describes our data. In section III, we discuss the empirical strategy to address unobserved heterogeneity, simultaneity and measurement error. Section IV presents the estimation results. Section V concludes.

II. Data

For our analysis, we use the Dutch Victimization Survey (PMB). The PMB is a repeated cross-section telephone survey that is unique in its sampling size. Whereas the US National Crime Victimization Survey covers 1 out of 2,000 of the population above 12 years of age, the PMB covers some 1 out of 200 of the population above 15 years of age. As a result of PMB's sample size, victimization rates can be disaggregated at the level of municipalities.

The PMB contains detailed information on victimization of crime, experience of public disorder and preventative measures. For every survey wave, respondents have been selected at random from the total population over 15 years of age. Per police region (and sometimes

smaller areas), the interviewers used stratified sampling. A minimum of 1,000 respondents were interviewed in each of the 25 police regions.

Table 1	Sample statistics (1996-2004)		
Deterrence measure		Mean	Standard deviation
Police personnel per 100,000 population		273	103
Victimization of crime			
Bicycle theft (per household)		0.11	0.31
Theft from car (per household)		0.07	0.25
Theft of car (per household)		0.01	0.09
Burglary (per household)		0.06	0.23
Threat with violence		0.05	0.22
Assault		0.01	0.09
Robbery with violence		0.00	0.06
Frequent experience of public disorder			
Littering		0.29	0.46
Graffiti		0.14	0.35
Nuisance from youth		0.12	0.33
Harassment (in public spaces)		0.03	0.18
Public intoxication		0.08	0.27
Vandalism		0.19	0.39
Nuisance from drug users		0.07	0.25
Frequently taken preventative measures			
Drive or walk round to avoid unsafe places		0.09	0.29
Leaving valuable properties at home to prevent theft		0.17	0.38
Not allowing children to go out because of safety reasons (per household)		0.17	0.38
Individual and municipality characteristics			
Male		0.54	0.50
Age		47.8	17.4
Primary school or basic vocational training only		0.24	0.43
Student		0.05	0.22
Immigrant		0.04	0.04
Employed		0.57	0.50
Housewife		0.20	0.40
Single household		0.22	0.42
Children[a]		0.15	0.23
Terraced house		0.57	0.50
Average moving mobility[b]		0.11	0.03
Average number of shops[c]		0.03	0.01
Average length of roadways (km)		0.08	0.06

Notes: (a) Share of children in total household size. (b) Number of persons moved per 1,000 population: number of persons moved is equal to the people moving within a municipality plus the sum of half of the people moving into the municipality and half of the people moving out of the municipality. (c) Number of commercial services establishments.

All victimization is observed at the individual level, only some property crimes are measured at the household level (bicycle theft, burglary, car theft, theft from cars) as well as one preventative measure (not allowing children to go out). Respondents are interviewed in the first ten weeks in the year a survey is held, with the victimization reflecting the twelve months preceding the interview date. This means that a substantial part of the observed victimization

occurs in the year preceding the interview year. For this reason, we use crime and public disorder data from the survey in year t for observations in year $t–1$.

In addition to the PMB, data on police resources were obtained from the Dutch Interior Department. Historical series of police levels are only available at the regional level. For all 25 police regions, growth in police personnel outstripped growth in population in this period. The total number of police personnel in full time equivalents increased from 39,462 in 1996 to 47,342 in 2004. We obtained data for some of the funding formula variables from the Dutch Central Bureau of Statistics.

We have pooled PMB data for the years 1996-2004. During this period, the PMB survey was conducted every other year. Thus our sample includes five waves. Table 1 summarizes the key variables we use for the empirical analysis. Background characteristics include age, gender, education level, ethnicity, housing type, household size, and income status (employed, student, housewife, or else).

We estimate the effect of police on victimization of crime and experience of public disorder at the level of individual respondents, taking into account municipality specific effects. In the pooled sample, we have 644 municipalities, with 599 individual observations on average. Table 2 shows the distribution of individual observations over municipalities. For the vast majority of municipalities, we have sufficient individual observations to obtain reliable municipality specific effects.

Table 2   Distribution of respondents across municipalities (1996-2004)

Number of individual respondents per municipality	Percent of municipalities	Percent of respondents
< 50	11	1
50 – 100	13	2
100 – 200	19	5
200 – 300	13	5
300 – 400	6	4
400 – 500	8	6
500 – 1000	19	23
> 1000	11	55
Total number	644	385,543

Note: As a result of amalgamations, the number of municipalities declined from 625 in 1996 to 483 in 2004. We coded a subset of all amalgamations to limit the number of municipalities with a very low number of observations.

III. Research design

Heterogeneity across municipalities

We start off by modeling the chance of becoming victim of crime or experiencing public disorder *y* as follows:

(1) $\quad \Pr(y_{ijt} = 1 \mid \boldsymbol{\alpha}, \boldsymbol{X}) \quad = \quad \alpha_j + X_{ijt}\,\boldsymbol{\beta} + \gamma\, lnp_{jt-1} + \varepsilon_{ijt}$

In this Linear Probability Model (LPM), α_j indicates an effect on *y* that is specific for municipality *j*. X_{ijt} is a matrix representing the characteristics of individual *i* living in municipality *j* at time *t*, with $\boldsymbol{\beta}$ as a vector describing the effects of *X*. lnp_{jt-1} represents the logarithmic value of the number of police personnel per capita in municipality *j* at time *t-1*, and γ is a parameter describing the effect of *lnp*. For ease of exposition, we initially assume that the number of police personnel per capita is observed at the level of municipalities. We return to this issue at the end of this section. Finally, ε_{ijt} is the error term of the model. Initially, we assume this error term to be identically and independent distributed, with mean zero.

We include individual characteristics X_{ijt} to prevent estimation bias through (observable) factors that affect both police and victimization levels (we also include household size and composition for crimes measured at the household level). Individuals with a higher probability of victimization may also enjoy greater police protection. Ignoring the correlation between police and victimization levels would bias the estimated effect of police on crime towards zero.

As we show in section IV, victimization is not determined by individual characteristics alone. In particular, experience of disorder like vandalism and public intoxication is mostly determined by the characteristics of the municipality in which someone lives rather than by his or her individual characteristics like age and sex. Therefore, we need to include municipality-specific characteristics that co-determine police and victimization levels as well.

Since the Netherlands uses a funding formula to distribute police resources across municipalities, municipality-specific characteristics that may be related to both police and victimization levels are all observable (see the text box below). By including the variables from the formula that are related to crime rates, we are able to control for the fact that municipalities with persistently higher crime rates enjoy higher levels of police staffing. We include the following seven variables from successive funding formulas in our estimation equation (all per population): population younger than 20 (we include the complete demographic structure), number of employed (the survey focuses on employment rather than unemployment), number of immigrants, number of shops, number of houses (the survey includes information on housing type), length of roadways, number of moves and the product

of housing density and number of houses. Statistics on cafe personnel and parking places are no longer available. By explicitly modeling budget policy, we aim to eliminate correlation between *lnp* and *α*, the municipality specific effect.

The police funding formula

The police budget has been distributed by means of a formula since the 1960s. Until 1985, distribution of resources was based on the number of population. That year, the number of population per house was added to redistribute resources from rural to urban municipalities. To bring the distribution of the police budget further in line with differences in workload of the police between urban and rural municipalities, a new funding formula was introduced in 1989. The variables in the formula encompassed number of population below 20, immigrant population and unemployment, number of houses, number of personnel in service sector and personnel in cafes, and the number of parking places. In the 1990s, rural municipalities complained that they were no longer able to meet minimum standards of emergency assistance. To meet their demand for a greater share of police resources, a new formula was introduced in 1996. The formula encompasses number of population and immigrants, moving mobility, number of houses and shops, the product of housing density and the number of houses, and length of roadways. This formula has been in use largely unchanged since then.

As we will show later in this section, not all variation in police levels is explained by the crime-related variables of the funding formula. Redistribution of police personnel across municipalities takes long to crystallize out, both in the case of uneven growth in new personnel and in the case of shifts in existing resources between municipalities. New personnel needs to be hired and trained and transfers of personnel are slow. Moreover, next to the funding formula, additional funding is provided to some police regions – while they maintain the freedom to spend these resources in a way they see fit. These extra allowances are meant to foresee in additional workload stemming from police tasks such as protection of embassies and international organizations, Rotterdam-harbor and Schiphol Airport, border patrol and policing inland waterways.

To control for the possibility that variation in police levels – after controlling for funding formula variables – may still be crime-related, we include municipality characteristics by averaging each individual control variable at the level of the municipality. We also include the average police level in a municipality. Year fixed effects are included to focus on differences between municipalities rather than capturing the correlation between national trends in police and victimization levels.

Including municipality averages is in line with the modified random effects (RE) model developed by Mundlak (1978). Within this approach, the correlation between the municipality effect and the time varying observables is specified as a linear function of municipality averages:

(2) $\Pr(y_{ijt} = 1) = \alpha_{jt} + X_{ijt}\boldsymbol{\beta} + \gamma\, lnp_{jt-1} + \varepsilon_{ijt}$

with the auxiliary regression:

(3) α_{jt} = $\delta_1 X_{.j} + \delta_2 \, lnp_{j.} + \eta_{jt}$

where $X_{.j}$ represent characteristics of municipality j, and δ_1 describes the effect of these characteristics on α. $lnp_{j.}$ is the average value of police levels per municipality. η_{jt} represents the remaining yearly municipality effect. We assume this variable to be independent and identically distributed. By adding average values of X and lnp as a set of controls for unobserved heterogeneity, we disentangle the well known 'within' from the 'between' estimators of both coefficients. Thus the coefficient estimates are identified from variation of X and lnp, holding the averages constant.

Modifying the RE model along these lines is related to the empirical literature on cluster or peer effects (Manski, 1993). In this literature, adding group averages has been used to estimate the importance of cluster effects – that is, the effect of the behavior of other people in a reference group on the behavior of an individual. Within the context of our model, however, the interpretation of cluster effects is different. The coefficients for individual characteristics reflect the effect of an individual's behavior on the risk of victimization. The coefficients for the municipality averages reflect both cluster and sorting effects. Cluster effects occur if reference group behavior affects the individual risk of victimization. Think of vandalism indiscriminately affecting people living in a certain municipality. Sorting effects stem from individuals with similar characteristics concentrating into municipalities with an *a priori* higher risk of victimization. In section IV, we estimate the relative importance of individual versus municipality effects for several types of crime and public disorder.

Given the size of the data set, we expect the LPM to provide consistent estimates of the partial effects for the average values of X and lnp, provided that estimation techniques are robust to heteroskedasticy. Another advantage of using an LPM is that the estimation of municipality specific effects is computationally far more complex in a non-linear setting.

Simultaneity

For our estimates of police effectiveness to be consistent, changes in police personnel should not be correlated with local trends in crime. If they are, the estimated effect of police on crime and public disorder will be biased. A distribution of police resources based on a formula with observable characteristics allows us to address the problem of simultaneity as well.

Simultaneity is not likely to be strong for several reasons. First, the funding formula focuses on overall crime, whereas we relate changes in police levels to trends in specific types of crime and public disorder. There is quite some variation in trends between different crime categories. Second, it takes at least two years to hire and train new personnel. Therefore, budget decisions have a lagged effect on police levels, making it more difficult to respond to differences in crime trends (rare shifts of existing resources between municipalities were also slow to materialize). Third, none of the funding formula variables has been updated every

year; some variables have not been updated at all during 1996-2004. Regular updating of all variables would have resulted in reoccurring major redistributions of police resources across municipalities, which was seen as undesirable. Clearly, infrequent updating of variables impairs a policy response to differing local crime trends.

To control for changes in the distribution of police resources that are related to local crime trends, we allow the funding formula variables in use during the period 1996-2004 as well as the other municipality characteristics to vary over time. Thus, $X_{.j}$ in the auxiliary regression varies from year to year and becomes $X_{.jt}$. Since some of the variables were 'frozen', we take the average value for the number of shops, moving mobility and length of roadways.

Table 3   **The funding formula as predictor of police levels**

Demographic structure		
Age < 20	0.13	(0.47)
20 < Age < 30	− 0.94**	(0.40)
30 < Age < 50	− 0.11	(0.37)
50 < Age < 60	− 0.11	(0.35)
Housing type		
Terraced house	0.57*	(0.29)
Detached house	0.41	(0.27)
Apartment	1.31***	(0.32)
Immigrants	1.54***	(0.44)
Employed	0.12	(0.38)
Housing density * houses	0.14***	(0.03)
Average number of shops	− 0.61	(2.27)
Average number of moves	− 2.84***	(0.93)
Average length of roadways	0.67**	(0.31)
Number of observations	370,437	
R^2	0.74	

Notes: Results for year fixed effects are not reported. Standard errors are adjusted for correlation within police regions. Standard errors are between parentheses. * Statistically significant at the 10-percent level. ** Statistically significant at the 5-percent level. *** Statistically significant at the 1-percent level.

Table 3 shows that the variables from the budget variables together with the national trend explain some three quarter of the variation in police levels. We use the remaining variation to identify the effect of police on crime and public disorder. Thus we assume the variation in police levels that is unrelated to the funding formula variables, the other municipality characteristics and a national trend to be exogenous to the relation between police and crime. An important related assumption is that the police does not change its prioritization of resources targeted at combating specific types of crime and public disorder.

Measurement error

We use municipalities as the relevant geographical or reference unit in our analysis. Thus far, we have abstracted from the fact that police personnel per capita is measured with error – that

is, we observe police personnel per capita for 25 regions, instead of 644 municipalities. Now suppose we define the natural logarithm of police personnel per capita as the sum of a regional component and a municipality specific component. Furthermore, we assume the municipality specific component to have an expected value of zero and to be independent and identically distributed. Then, in order to obtain consistent estimates, the municipality specific component should not be correlated with the regional component. This assumption is common for various applications of stratified sampling, where units in the sample are represented with different frequencies than they are in the population.[1]

Measurement error may also arise with respect to the averages of the other independent variables in our model. Suppose that these variables would be averaged across regions instead of municipalities, then – using similar arguments as in the case of police personnel – consistency is still achieved. This is not an argument to estimate the model using regions as reference groups for all independent variables. Given the number of independent variables in our model, the number of regions is very limited, casting doubts on the efficiency of the estimated parameters of the averaged variables.[2] This is also likely to affect the efficiency of the estimated effect of police on crime. Thus, although we do not observe police personnel at the level of municipalities, this is no reason to switch to regions as cluster groups for the other independent variables as well.

IV. Estimation results

In this section we present the estimation results. We start by presenting the estimation results of the modified RE approach as defined by equations (2) and (3): the 'preferred model'. In order to assess the robustness of our findings, we re-estimate our model using various other specifications. Then, we go into the explanatory power of control variables at the individual versus the municipality level. We also estimate the effect of police on preventative behavior.

Effect of police on crime and public disorder

Table 4 presents the estimation results based on the preferred model.[3] When analyzing property crime, we find the police to be effective in reducing bicycle theft and theft from cars. The effect of police on victimization of burglary and car theft is not statistically significant. The result for car theft is most likely due to low victimization rates (one percent of households) and correspondingly high standard errors. In contrast to police statistics, a victimization survey provides less precise figures for relatively rare types of crimes. Apparently, the police is not very successful in bringing down burglary, a result that stands in

---

[1] This contrasts to situations where the classical errors in variables (CEV) assumption applies. In that case, measurement errors are (fully) correlated with the observed variable, causing estimates to be biased to zero (attenuation bias).
[2] See Wooldridge (2002), who discusses the importance of a high number of clusters to be able to apply panel data methods.
[3] The elasticities are computed as follows: estimated coefficient for *lnp* (as in appendix) divided by the average value of that specific type of crime or nuisance (as in summary statistics).

contrast to other studies that find a statistically significant negative effect of police on burglary of around – 0.30 (Marvell and Moody, 1996, Corman and Mocan, 2002, Klick and Tabarrok, 2005).

We estimate a one percent increase in police personnel to result in a decrease in threat with violence of 0.4 percentage point. We also find a negative effect on assault and robbery with violence, but the effect is not statistically significant. Victimization of both types of crimes occurs even less frequently than victimization of car theft, which results in high standard errors.

We select seven measures of public disorder included in the survey that are available for each wave: littering, graffiti, vandalism, harassment, youth nuisance, public intoxication and drug nuisance. We find the effect of police to be negative and statistically significant for littering, harassment, youth nuisance, public intoxication and drug nuisance. The size of the effect on public disorder is similar to the estimated effect on crime. We find no significant effect of police on graffiti and vandalism.

Table 4   Effect of police on crime and public disorder – implied elasticities

Property crime			Violent crime			Public disorder		
Burglary	– 0.14	(0.14)	Threat with			Littering	– 0.16***	(0.06)
Car theft	0.17	(0.41)	violence	– 0.38**	(0.16)	Graffiti	– 0.07	(0.09)
Theft from car	– 0.31**	(0.14)	Assault	– 0.32	(0.40)	Vandalism	– 0.11	(0.08)
Bicycle theft	– 0.48***	(0.10)	Robbery with			Harassment	– 0.47**	(0.20)
			violence	– 0.37	(0.59)	Youth nuisance	– 0.42***	(0.10)
						Public intoxication	– 0.20*	(0.12)
						Drug nuisance	– 0.36***	(0.13)

Notes: Estimation results for all other variables are included in the appendix. Standard errors are between parentheses. * Statistically significant at the 10-percent level. ** Statistically significant at the 5-percent level. *** Statistically significant at the 1-percent level.

Comparison with other studies

Table 5 shows that our estimated elasticities for property and violent crime are in line with recent studies based on police statistics (we only include estimates statistically significant at the 10-percent level or better). For reasons of comparability, we have lumped together several types of property crimes and violent crimes.

Reliable estimates of the effect of police on violent crime are rare. Measurement error in police statistics on violent crime is a primary reason for the lack of empirical evidence. The Netherlands is no exception. Whereas Dutch victimization data and police statistics for property crime show similar trends, Wittebrood and Junger (2002) find very different trends for violent crime. Police records on violent crime have improved considerably, because police reports are made for an increasing number of notified crimes and more police reports are finding their way into official records. The discrepancy between victimization data and police statistics, the so-called 'dark number', is becoming smaller. The precise trends in recorded

violent crime differ from police force to police force, however, leading to major measurement error. Thus, particularly in the case of violent crime, Dutch police statistics do not provide for a reliable alternative source of crime data.

Table 5	Point estimates for effect of police on property and violent crime			
	Unit of analysis	Property crime	Car theft and theft from cars	Violent crime
This study	Dutch municipalities	– 0.35***[a]	– 0.28**	– 0.31** [b]
Marvell & Moody, 1996	Major US cities		– 0.85**[c]	
Levitt, 2002	Major US cities	– 0.50**		– 0.44*
DiTella & Schargrodsky, 2004	Buenos Aires neighborhoods		– 0.33***[c, d]	
Klick & Tabarrok, 2005	Washington D.C. city districts		– 0.86**	
Corman & Mocan, 2005	New York City		– 0.56**[c]	

Notes: (a) Includes burglary, car theft, theft from car and bicycle theft. (b) Includes threat with violence, assault and robbery with violence. (c) Excludes theft from cars. (d) Deterrence effects only. * Statistically significant at the 10-percent level. ** Statistically significant at the 5-percent level. *** Statistically significant at the 1-percent level.

Although the Netherlands may provide a different setting than the US or Argentina, the results do not suggest that the use of a victimization survey as alternative source of crime data greatly affects the estimated effect of police on crime. If anything, our point estimates are somewhat lower. Thus, our results confirm the existing evidence using a different source of crime data and a different estimation method. Additionally, we extend the evidence to include the effect of police on public disorder – and on preventative measures, which we will discuss later on.

Robustness

To assess the importance of controlling for the policy response to differences in crime (trends) and to validate the robustness of our results, we re-estimate the model using various other specifications. We vary the set of control variables, the method of estimation, and the circumstances under which the police works. Table 6 provides an overview of the results – with the preferred model as our benchmark.

In the second specification, we leave out municipality characteristics that are known to be part of the budget formula. The estimation results show that not explicitly controlling for the relation between budget policy and victimization rates at the municipality level results in an *over*estimation of the effect of police on public disorder, while producing similar estimates for property and violent crime. As we discuss in the next paragraph, municipality characteristics explain most of the variation in public disorder, whereas individual characteristics are of more importance for explaining variation in property and violent crime. Consequently, leaving out key control variables at the municipality level biases the results for public disorder more than the results for crime.

We re-estimate the model using FE estimation in the third specification.[4] In the existing empirical literature on police effectiveness, FE estimation is the default option to address unobserved heterogeneity. In the FE models, differences between municipalities are specified as an effect that is constant over time. A comparison between RE and FE estimation results shows that if we control for all possible municipality characteristics that are constant over time, but do not control for simultaneity, we are likely to *under*estimate police effectiveness, with graffiti and youth nuisance as the exceptions. Apparently, changes in police personnel per capita tend to be positively correlated with trends in crime and public disorder. Thus the overestimation of police effectiveness in specification (ii) is most likely due to time-constant characteristics of municipalities: municipalities with severe and persistent public disorder problems do not enjoy relatively high staffing levels.

Table 6   **Effect of police on crime and public disorder – elasticities for various model specifications**

	Specification (i) Preferred model		Specification (ii) Limited municip. chars.		Specification (iii)		Specification (iv) Urban regions	
Estimation approach	Modified Random Effects (Mundlak)		Modified Random Effects (Mundlak)		Fixed Effects		Modified Random Effects (Mundlak)	
Property crime								
Burglary	− 0.14	(0.14)	− 0.19	(0.14)	0.05	(0.19)	− 0.18	(0.14)
Car theft	0.17	(0.41)	− 0.07	(0.38)	− 0.16	(0.57)	0.13	(0.41)
Theft from car	− 0.31**	(0.14)	− 0.26**	(0.13)	− 0.24**	(0.19)	− 0.37***	(0.14)
Bicycle theft	− 0.48***	(0.10)	− 0.55***	(0.10)	0.26*	(0.14)	− 0.51***	(0.10)
Violent crime								
Threat with violence	− 0.38**	(0.16)	− 0.41**	(0.15)	− 0.42**	(0.21)	− 0.39**	(0.16)
Assault	− 0.32	(0.40)	− 0.34	(0.38)	− 0.48	(0.52)	− 0.34	(0.40)
Robbery with violence	− 0.37	(0.59)	− 0.35	(0.77)	− 0.49	(0.79)	− 0.48	(0.60)
Public disorder								
Littering	− 0.16***	(0.06)	− 0.30***	(0.05)	− 0.20***	(0.07)	− 0.14***	(0.06)
Graffiti	− 0.07	(0.09)	− 0.23***	(0.08)	− 0.32***	(0.12)	− 0.09	(0.09)
Vandalism	− 0.11	(0.08)	− 0.27***	(0.08)	− 0.07	(0.10)	− 0.16**	(0.08)
Harassment	− 0.47**	(0.20)	− 0.75***	(0.18)	1.26***	(0.26)	− 0.48**	(0.20)
Youth nuisance	− 0.42***	(0.10)	− 0.46***	(0.09)	− 0.61***	(0.13)	− 0.44***	(0.10)
Public intoxication	− 0.20*	(0.12)	− 0.22***	(0.12)	0.08	(0.16)	− 0.15	(0.13)
Drug nuisance	− 0.36***	(0.13)	− 0.57***	(0.12)	0.74***	(0.17)	− 0.31**	(0.13)

Notes: Standard errors are between parentheses. * Statistically significant at the 10-percent level. ** Statistically significant at the 5-percent level. *** Statistically significant at the 1-percent level.

The use of non-experimental data allows us to study the robustness of our findings under different conditions. In specification (iv), we vary the degree of urbanization of a municipality in which the police is active. Intuitively, we expect the police to be most effective in fighting crime and public disorder in urban areas. After all, a police officer in a densely populated area can control more people than a police officer in a sparsely population area. To test the

---

[4] The FE model is based on an LPM specification as well. The advantage of using an LPM specification rather than a binary model is that fixed effect (FE) estimation is subject to the incidental parameters problem (see Wooldridge, 2002). The estimation results are robust to the choice between LPM and Probit as well as Logit.

difference in effectiveness between urban and rural regions, we include an interaction term for police levels and average police levels in the four most urbanized 'Randstad' regions in the preferred model.[5] We assume that the impact of all other explanatory variables in our model is equal for regions with a high and low degree of urbanization. The last column of the table shows that the estimation results for urban regions are very similar to the results for the whole country in the first column. Thus effectiveness in fighting crime and public disorder does not differ greatly between urbanized and less urbanized regions, which is in line with Kovandzic and Sloan (2002, p. 73).

Victimization: is it the individual or the municipality?

Since we use individual victimization data, we are able to control for background characteristics both at the level of the individual and of the municipality. The estimation results for the control variables allow us to analyze whether victimization of crime and experience of public disorder is dominated by individual characteristics or by factors at the level of the municipality. Individual characteristics can reflect the extent to which people are willing to take risks for instance. But higher victimization rates may also be related to the municipality in which someone lives. In that case, people experience crime and public disorder regardless of their individual background characteristics.

To see whether individual or municipality factors dominate, we decompose the explained variance in crime and public disorder rates into two parts: the proportion of the explained variance due to differences in background characteristics of individual respondents and the proportion due to differences in characteristics of municipalities (all estimates are based on the preferred model). Table 7 presents the results of the variance decomposition.

Factors at the level of the municipality are particularly important when explaining experience of public disorder. This result makes intuitive sense: there is not much an individual can do about experiencing problems like littering, graffiti and drug nuisance. Public disorder is high in municipalities with a young, low educated, immigrant population outside the labor force living in apartment buildings (tables A.3 and A.4). Public disorder is simply part of living in municipalities with these characteristics.

In the case of property crime and particularly violent crime, individual effects dominate, with robbery with violence as the exception. Especially assault, threat with violence and bicycle theft vary across individuals rather than across municipalities. Thus, compared to public disorder, these types of crime are not as likely to be a characteristic of a municipality. Victimization of property and violent crime decreases with age and education level, and is relatively high among females, immigrants, and people living in apartment buildings (tables

---

[5] The most urbanized regions include: Amsterdam-Amstelland, Rotterdam-Rijnmond, Haaglanden and Utrecht.

A.1 and A.2). These findings underline the importance of controlling for individual background characteristics, especially in the case of violent crime.

Table 7	Variance decomposition for property crime, violent crime and public disorder	
	Explanatory power of individual effects	Explanatory power of municipality effects
Property crime		
Total property crime	0.71	0.29
Burglary	0.51	0.49
Car theft	0.49	0.51
Theft from car	0.47	0.53
Bicycle theft	0.72	0.28
Violent crime		
Total violent crime	0.77	0.23
Threat with violence	0.79	0.21
Assault	0.83	0.17
Robbery with violence	0.23	0.77
Public disorder		
Total public disorder	0.29	0.71
Littering	0.26	0.74
Graffiti	0.29	0.71
Vandalism	0.42	0.58
Youth nuisance	0.48	0.52
Harassment	0.26	0.74
Public intoxication	0.57	0.43
Drug nuisance	0.29	0.71

Police protection and preventative behavior

With a higher level of police protection, citizens need to take fewer self-protective measures such as not venturing out at night or installing a burglar alarm. A relaxation in preventative behavior is an additional gain of higher police levels not reflected in lower victimization rates.[6]

Using the same modified RE approach as before, we test whether precautionary measures are affected by increases in police personnel. The victimization survey includes data on preventative measures. We focus on three measures in the survey that people can easily alter according to local safety conditions: drive or walk round to avoid unsafe places, leaving valuable properties at home to prevent theft, and not allowing children to go out because of safety reasons. These are also measures people decide for themselves, as against measures like additional hinges and locks on doors and windows that the police might advise about.[7] Just like the one-year lag between the effect of police on crime, we assume a one-year lag

---

[6] We do not consider the optimal balance between the two forms of protection.
[7] In the case the police advise citizens on appropriate preventative measures, more police could lead to more private prevention. We focus on measures that are most likely to be individual decisions affected by the degree of police protection rather than by advice from the police.

between higher police presence and changes in preventative behavior. We assume that when making a decision on whether to avoid a certain street for instance, potential victims treat public expenditures on crime control as exogenous. After all, as discussed in section III, our estimation approach addresses simultaneity between police and crime, therefore, we also deal with simultaneity between police and preventative measures related to crime levels.

Table 8	Effect of police on preventative behavior, estimated elasticities		
Drive or walk round to avoid unsafe places		− 0.25**	(0.11)
Leaving valuable properties at home to prevent theft		− 0.16**	(0.08)
Not allowing children to go out because of safety reasons		− 0.08	(0.15)

Notes: Standard error between parentheses. * Statistically significant at the 10-percent level. ** Statistically significant at the 5-percent level. *** Statistically significant at the 1-percent level.

Table 8 provides the estimated elasticities. As expected, more police leads to fewer self-protective measures. A one percent increase in police levels leads to a 0.2 percent decrease in people who frequently drive or walk round to avoid unsafe places and a similar decrease in people who frequently leave valuable properties at home to prevent theft. The effect on parents' behavior towards their children is not statistically significant. Thus, we provide evidence that there is an additional effect of police on preventative behavior, next to the effect of police on victimization rates most frequently reported in the literature.[8]

V. Conclusions

The literature on police effectiveness relies exclusively on police statistics as source of crime data. In this chapter, we switch the perspective from offenders to victims of crime, using data from the Dutch Victimization Survey. This data set covers a wide range of crimes and also types of public disorder for a range of years and provides a great number of individual background characteristics of respondents. When estimating police effectiveness, the survey data allows us to control for both municipality effects and individual characteristics of victims. We model the crime-related variables that are part of the police funding formula to identify the endogenous variation in police levels, and use the remaining variation to estimate the effect of police on crime and public disorder.

We find significantly negative effects of higher police levels on property crime and violent crime. The estimated elasticities for a number of crime categories range from − 0.3 to − 0.5. Using a different source of data and a different estimation method, we confirm the existing empirical evidence – and extend it to the experience of public disorder. We find the police to

---

[8] Philipson and Posner (1996) argue that a decline in preventative behavior due to better police protection partially offsets the effect of police on crime. An increase in the level of public protection, as by hiring more police, will cause the crime rate to fall and thus will lower the demand for private prevention – which will cause the crime rate to rise again, partially undoing the effect of the increase in public protection. Because of same-year simultaneity between crime and preventative behavior, we are not able to test whether fewer preventative measures in response to better police protection lead to higher victimization rates.

have a similar impact on several types of public disorder not included in previous studies: littering, youth nuisance, harassment, public intoxication and drug nuisance. The estimated elasticities range from – 0.2 to – 0.5.

Comparing estimates from different model specifications and different estimation techniques, we show the importance of controlling for simultaneity. We find police levels to be correlated with changes in municipality specific characteristics. This effect would not have been picked up in a fixed effects estimation approach, leading to underestimation of the effect of police.

We find experience of public disorder mostly to be a characteristic of the municipality in which someone lives, with little variation across individuals in a municipality, whereas victimization of property crime and particularly violent crime varies across individuals rather than municipalities. These findings underline the importance of controlling for individual background characteristics, especially in the case of violent crime.

Finally, we provide evidence that greater police protection leads to fewer preventative measures, which is an additional benefit of higher police levels not reflected in a decline in victimization rates.

Appendix

Table A.1 **Effect of police on victimization of property crime - the modified RE model**

	Burglary		Car theft		Theft from car		Bicycle theft	
Ln (police) (t-1)	− 0.008	(0.008)	0.001	(0.004)	− 0.021**	(0.010)	− 0.053***	(0.011)
Individual characteristics								
Male	− 0.005***	(0.001)	− 0.002***	(0.000)	− 0.006***	(0.001)	0.006***	(0.001)
Age < 20	0.039***	(0.004)	0.010***	(0.002)	0.060***	(0.004)	0.138***	(0.006)
20 < age < 30	0.028***	(0.002)	0.007***	(0.001)	0.078***	(0.002)	0.074***	(0.002)
30 < age < 50	0.021***	(0.002)	0.005***	(0.001)	0.049***	(0.002)	0.050***	(0.002)
50 < age < 60	0.018***	(0.001)	0.004***	(0.001)	0.033***	(0.001)	0.032***	(0.002)
Education level 1	− 0.021***	(0.002)	0.001	(0.001)	− 0.020***	(0.002)	− 0.040***	(0.002)
Education level 2	− 0.015***	(0.002)	0.001	(0.001)	− 0.017***	(0.002)	− 0.034***	(0.002)
Education level 3	− 0.011***	(0.002)	0.001	(0.001)	− 0.008***	(0.003)	− 0.028***	(0.003)
Education level 4	− 0.004**	(0.002)	0.001	(0.001)	− 0.005**	(0.002)	− 0.011***	(0.002)
Employed	− 0.001	(0.001)	0.001	(0.001)	0.008***	(0.001)	0.011***	(0.002)
Student	− 0.004	(0.003)	− 0.002	(0.002)	− 0.010***	(0.004)	0.062***	(0.005)
House wife	− 0.005***	(0.001)	0.000	(0.001)	− 0.003**	(0.001)	− 0.007***	(0.002)
Immigrant	0.004**	(0.002)	0.006***	(0.001)	0.017***	(0.003)	0.018***	(0.003)
Terraced house	− 0.009***	(0.002)	− 0.003***	(0.001)	− 0.009***	(0.002)	− 0.018***	(0.002)
Detached house	0.010***	(0.002)	− 0.002**	(0.001)	− 0.005**	(0.002)	− 0.024***	(0.002)
Apartment	− 0.016***	(0.002)	0.001	(0.001)	0.011***	(0.003)	0.007***	(0.003)
Municipality characteristics								
Ln (police) (t-1)	0.004	(0.008)	− 0.002	(0.003)	0.020**	(0.010)	0.014	(0.011)
Male	0.024***	(0.009)	0.006	(0.004)	0.014	(0.010)	− 0.029**	(0.012)
Age < 20	− 0.031	(0.034)	0.004	(0.015)	− 0.122***	(0.038)	− 0.064	(0.044)
20 < age < 30	0.002	(0.015)	0.001	(0.006)	− 0.033**	(0.017)	0.070***	(0.020)
30 < age < 50	− 0.012	(0.014)	0.015***	(0.006)	− 0.011	(0.015)	0.075***	(0.018)
50 < age < 60	− 0.011	(0.014)	0.001	(0.006)	− 0.014	(0.016)	0.006	(0.018)
Education level 1	0.030***	(0.011)	0.006	(0.005)	0.066***	(0.013)	0.031**	(0.015)
Education level 2	0.006	(0.010)	− 0.002	(0.004)	0.022*	(0.012)	− 0.033**	(0.014)
Education level 3	− 0.017	(0.018)	− 0.018**	(0.007)	0.027	(0.021)	− 0.106***	(0.024)
Education level 4	0.063***	(0.013)	0.010*	(0.006)	0.098***	(0.015)	0.024	(0.018)
Employed	− 0.004***	(0.011)	− 0.005	(0.005)	0.026**	(0.013)	− 0.024	(0.015)
Student	0.013**	(0.031)	− 0.011	(0.013)	0.091***	(0.035)	0.217***	(0.041)
House wife	− 0.036***	(0.013)	− 0.004	(0.005)	0.014	(0.014)	− 0.047***	(0.017)
Immigrant	0.000**	(0.019)	0.007	(0.008)	0.065***	(0.023)	− 0.136***	(0.026)
Terraced house	0.100***	(0.013)	0.013**	(0.006)	0.062***	(0.015)	0.125***	(0.017)
Detached house	0.074***	(0.014)	0.017***	(0.006)	0.058***	(0.016)	0.078***	(0.019)
Apartment	0.119**	(0.015)	0.020***	(0.007)	0.101***	(0.017)	0.054***	(0.020)
Housing density	0.009***	(0.002)	0.000	(0.001)	0.023***	(0.003)	0.050***	(0.003)
Shops	0.176*	(0.074)	0.073**	(0.032)	0.217***	(0.085)	− 0.153	(0.100)
Moves	− 0.028***	(0.004)	− 0.006***	(0.002)	− 0.022***	(0.005)	0.012*	(0.007)
Length of roads	− 0.068***	(0.010)	− 0.019***	(0.004)	− 0.067***	(0.011)	− 0.048***	(0.013)
Number of obs.	368,638		307,115		307,114		340,501	

Notes: Results for year fixed effects, household size and composition are not reported. Standard errors between parentheses. * Statistically significant at the 10-percent level. ** Statistically significant at the 5-percent level. *** Statistically significant at the 1-percent level.

Table A.2   Effect of police on victimization of violent crime – the modified RE model

	Threat with violence		Assault		Robbery with violence	
Ln (police) (t-1)	− 0.019**	(0.008)	− 0.003	(0.003)	− 0.001	(0.002)
Individual characteristics						
Male	− 0.034***	(0.001)	− 0.003***	(0.000)	0.000	(0.000)
Age < 20	0.105***	(0.004)	0.032***	(0.002)	0.005***	(0.001)
20 < age < 30	0.058***	(0.002)	0.014***	(0.001)	0.001**	(0.000)
30 < age < 50	0.037***	(0.001)	0.008***	(0.001)	0.000	(0.000)
50 < age < 60	0.022***	(0.001)	0.004***	(0.000)	0.001	(0.000)
Education level 1	− 0.019***	(0.002)	0.000	(0.001)	− 0.001***	(0.000)
Education level 2	− 0.011***	(0.002)	0.001	(0.001)	− 0.001**	(0.000)
Education level 3	− 0.006***	(0.002)	0.001	(0.001)	− 0.001	(0.001)
Education level 4	0.001	(0.002)	0.001**	(0.001)	0.000	(0.000)
Employed	0.001	(0.001)	− 0.001**	(0.001)	0.000	(0.000)
Student	-0.006*	(0.004)	-0.001	(0.002)	− 0.001	(0.001)
House wife	0.000	(0.001)	0.000	(0.000)	0.001*	(0.000)
Immigrant	− 0.014***	(0.002)	0.002***	(0.001)	0.001*	(0.001)
Terraced house	− 0.009***	(0.001)	− 0.003***	(0.001)	− 0.002***	(0.000)
Detached house	− 0.011***	(0.002)	− 0.002***	(0.001)	− 0.002***	(0.001)
Apartment	0.004**	(0.002)	0.000	(0.001)	− 0.001**	(0.001)
Municipality characteristics						
Ln (police) (t-1)	0.012	(0.008)	0.001	(0.003)	0.002	(0.002)
Male	0.000	(0.009)	− 0.001	(0.004)	− 0.001	(0.002)
Age < 20	− 0.027	(0.028)	0.000	(0.012)	0.003	(0.007)
20 < age < 30	0.026**	(0.013)	0.002	(0.006)	0.008***	(0.003)
30 < age < 50	0.017	(0.011)	− 0.007	(0.005)	0.007***	(0.003)
50 < age < 60	0.023*	(0.012)	0.002	(0.005)	0.008**	(0.003)
Education level 1	0.012	(0.010)	0.000	(0.004)	0.006***	(0.002)
Education level 2	0.010	(0.009)	0.000	(0.004)	0.001	(0.002)
Education level 3	0.019	(0.017)	0.013*	(0.007)	0.002	(0.004)
Education level 4	0.035***	(0.012)	0.006	(0.005)	0.008***	(0.003)
Employed	− 0.024**	(0.010)	− 0.001	(0.004)	− 0.003	(0.002)
Student	0.009	(0.028)	− 0.012	(0.012)	− 0.003	(0.006)
House wife	− 0.036***	(0.011)	− 0.003	(0.005)	− 0.004	(0.003)
Immigrant	0.074**	(0.018)	0.010	(0.008)	0.019***	(0.005)
Terraced house	0.015	(0.012)	0.009*	(0.005)	0.007***	(0.003)
Detached house	0.007	(0.013)	0.009	(0.005)	0.006**	(0.003)
Apartment	0.009	(0.014)	0.003	(0.006)	0.011***	(0.003)
Housing density	0.014***	(0.002)	0.005***	(0.001)	0.001**	(0.001)
Shops	0.195***	(0.068)	− 0.009	(0.029)	0.037**	(0.018)
Moves	− 0.011***	(0.004)	− 0.001	(0.002)	− 0.002**	(0.001)
Length of roadways	− 0.029***	(0.009)	− 0.003	(0.004)	0.004	(0.002)
Number of observations	368,666		368,696		368,700	

Notes: Results for year fixed effects are not reported. Standard errors are between parentheses. * Statistically significant at the 10-percent level. ** Statistically significant at the 5-percent level. *** Statistically significant at the 1-percent level.

Table A.3   Effect of police on experience of public disorder (1) – the modified RE model

	Littering		Graffiti		Youth nuisance	
Ln (police) (t-1)	− 0.046***	(0.016)	− 0.009	(0.013)	− 0.051***	(0.012)
Individual characteristics						
Male	0.032***	(0.002)	0.007***	(0.001)	0.009***	(0.001)
Age < 20	0.054***	(0.006)	0.063***	(0.005)	0.085***	(0.005)
20 < age < 30	− 0.003	(0.003)	− 0.003	(0.003)	0.049***	(0.002)
30 < age < 50	0.014***	(0.003)	0.002	(0.002)	0.038***	(0.002)
50 < age < 60	0.034***	(0.003)	0.017***	(0.002)	0.033***	(0.002)
Education level 1	− 0.002	(0.003)	− 0.015	(0.002)	0.028***	(0.002)
Education level 2	0.015***	(0.003)	0.002	(0.002)	0.028***	(0.002)
Education level 3	0.008**	(0.004)	0.005*	(0.003)	0.020***	(0.003)
Education level 4	0.019***	(0.003)	0.010***	(0.002)	0.018***	(0.002)
Employed	− 0.027***	(0.002)	− 0.006***	(0.002)	− 0.015***	(0.002)
Student	0.014***	(0.005)	0.009*	(0.005)	− 0.001	(0.004)
House wife	0.000	(0.002)	0.004***	(0.002)	− 0.007***	(0.002)
Immigrant	− 0.053***	(0.004)	− 0.035***	(0.003)	0.012***	(0.003)
Terraced house	− 0.026***	(0.003)	− 0.020***	(0.002)	− 0.015***	(0.002)
Detached house	− 0.047***	(0.004)	− 0.041***	(0.003)	− 0.033***	(0.003)
Apartment	0.097***	(0.004)	0.059***	(0.003)	0.043***	(0.003)
Municipality characteristics						
Ln (police) (t-1)	− 0.042***	(0.016)	− 0.008	(0.012)	0.012	(0.012)
Male	− 0.004	(0.018)	− 0.013	(0.013)	0.008	(0.014)
Age < 20	− 0.016	(0.060)	− 0.031	(0.045)	0.014	(0.045)
20 < age < 30	0.226***	(0.028)	0.230***	(0.020)	0.054***	(0.021)
30 < age < 50	0.160***	(0.023)	0.175***	(0.017)	0.121***	(0.017)
50 < age < 60	0.032	(0.027)	0.089***	(0.019)	0.084***	(0.020)
Education level 1	0.181***	(0.020)	0.057***	(0.015)	0.175***	(0.015)
Education level 2	0.113***	(0.019)	0.075***	(0.015)	0.102***	(0.014)
Education level 3	0.094***	(0.035)	0.027	(0.026)	0.038	(0.025)
Education level 4	0.070***	(0.025)	0.065***	(0.019)	0.113***	(0.018)
Employed	− 0.017	(0.022)	− 0.052***	(0.016)	− 0.018	(0.016)
Student	0.134***	(0.058)	0.140***	(0.044)	− 0.030	(0.043)
House wife	− 0.084***	(0.023)	− 0.098***	(0.017)	− 0.077***	(0.017)
Immigrant	0.458***	(0.038)	0.117***	(0.028)	0.085***	(0.028)
Terraced house	0.014	(0.026)	0.081***	(0.018)	0.125***	(0.019)
Detached house	− 0.148***	(0.028)	− 0.023	(0.019)	0.076***	(0.020)
Apartment	0.063**	(0.029)	0.182***	(0.021)	0.143***	(0.021)
Housing density	0.074***	(0.004)	0.024***	(0.107)	0.015***	(0.003)
Shops	0.129	(0.144)	0.351***	(0.003)	0.570***	(0.107)
Moves	− 0.079***	(0.009)	− 0.041***	(0.005)	0.028***	(0.009)
Length of roadways	− 0.170***	(0.021)	− 0.150***	(0.150)	− 0.177***	(0.015)
Number of observations	368,439		367,329		367,753	

Notes: Results for year fixed effects are not reported. Standard errors are between parentheses. * Statistically significant at the 10-percent level. ** Statistically significant at the 5-percent level. *** Statistically significant at the 1-percent level.

Table A.4   Effect of police on public disorder (2) - the modified RE model

	Harassment		Drug nuisance		Public intoxication		Vandalism	
Ln (police) (t-1)	− 0.016**	(0.007)	− 0.025***	(0.009)	− 0.016*	(0.010)	− 0.021	(0.015)
Individual characteristics								
Male	0.005***	(0.001)	0.000	(0.001)	0.000	(0.001)	0.011***	(0.001)
Age < 20	0.018***	(0.003)	0.055***	(0.004)	0.108***	(0.004)	0.070***	(0.005)
20 < age < 30	0.014***	(0.001)	0.040***	(0.002)	0.095***	(0.002)	0.011***	(0.003)
30 < age < 50	0.007***	(0.001)	0.041***	(0.002)	0.052***	(0.001)	0.023***	(0.002)
50 < age < 60	0.006***	(0.001)	0.041***	(0.002)	0.037***	(0.001)	0.046***	(0.002)
Education level 1	0.012***	(0.001)	0.020***	(0.002)	− 0.019***	(0.002)	0.056***	(0.003)
Education level 2	0.008***	(0.001)	0.011***	(0.002)	− 0.024***	(0.002)	0.057***	(0.002)
Education level 3	0.004***	(0.002)	0.001	(0.002)	− 0.024***	(0.002)	0.038***	(0.003)
Education level 4	0.006***	(0.001)	0.005***	(0.002)	− 0.014***	(0.002)	0.036***	(0.002)
Employed	− 0.005***	(0.001)	− 0.012***	(0.001)	− 0.003**	(0.001)	− 0.006***	(0.002)
Student	0.003	(0.003)	− 0.009***	(0.003)	0.005	(0.004)	0.004	(0.005)
House wife	− 0.003***	(0.001)	− 0.005***	(0.001)	− 0.005***	(0.001)	− 0.004***	(0.002)
Immigrant	− 0.001	(0.002)	0.018***	(0.003)	0.000	(0.002)	− 0.014***	(0.003)
Terraced house	− 0.015***	(0.001)	− 0.023***	(0.002)	− 0.041***	(0.002)	0.029***	(0.003)
Detached house	− 0.015***	(0.001)	− 0.029***	(0.002)	− 0.040***	(0.002)	− 0.006**	(0.003)
Apartment	0.011***	(0.002)	0.030***	(0.002)	0.022***	(0.002)	0.041***	(0.003)
Municipality characteristics								
Ln (police) (t-1)	0.001	(0.006)	− 0.023***	(0.009)	− 0.025***	(0.009)	− 0.014	(0.015)
Male	− 0.011*	(0.006)	− 0.046***	(0.009)	− 0.014	(0.011)	0.020	(0.017)
Age < 20	− 0.075***	(0.020)	− 0.144***	(0.032)	− 0.113***	(0.037)	0.226***	(0.055)
20 < age < 30	0.042***	(0.010)	0.236***	(0.015)	0.206***	(0.017)	0.141***	(0.026)
30 < age < 50	0.013*	(0.008)	0.121***	(0.012)	0.043***	(0.014)	0.274***	(0.021)
50 < age < 60	0.030***	(0.009)	0.092***	(0.014)	0.030**	(0.015)	0.276***	(0.025)
Education level 1	0.064***	(0.007)	0.207***	(0.010)	0.096***	(0.012)	0.101***	(0.018)
Education level 2	0.035***	(0.007)	0.098***	(0.010)	0.029***	(0.011)	0.071***	(0.017)
Education level 3	− 0.015	(0.012)	− 0.050***	(0.018)	− 0.035*	(0.020)	0.007	(0.032)
Education level 4	0.063***	(0.009)	0.154***	(0.013)	0.050***	(0.015)	− 0.023	(0.022)
Employed	− 0.018***	(0.007)	− 0.161***	(0.011)	− 0.048***	(0.013)	− 0.147***	(0.020)
Student	0.068***	(0.021)	0.211***	(0.031)	0.148***	(0.035)	− 0.055	(0.053)
House wife	− 0.002	(0.008)	− 0.069***	(0.012)	− 0.004	(0.014)	− 0.159***	(0.022)
Immigrant	0.063***	(0.014)	0.292***	(0.022)	− 0.139***	(0.022)	0.392***	(0.034)
Terraced house	0.057***	(0.008)	0.139***	(0.014)	0.054***	(0.015)	0.158***	(0.023)
Detached house	0.047***	(0.009)	0.139***	(0.014)	0.082***	(0.017)	0.027	(0.025)
Apartment	0.078***	(0.010)	0.109***	(0.015)	− 0.014	(0.017)	0.287***	(0.026)
Housing density	0.021***	(0.002)	0.053***	(0.003)	0.058***	(0.003)	− 0.059***	(0.004)
Shops	0.135***	(0.049)	0.877***	(0.074)	1.741***	(0.094)	− 0.394***	(0.127)
Moves	− 0.022***	(0.002)	− 0.051***	(0.004)	0.093***	(0.010)	− 0.059***	(0.006)
Length of roads	− 0.016***	(0.006)	− 0.019*	(0.010)	− 0.026**	(0.012)	− 0.452***	(0.018)
Number of obs.	357,943		362,362		364,609		359,522	

Notes: Results for year fixed effects are not reported. Standard errors are between parentheses. * Statistically significant at the 10-percent level. ** Statistically significant at the 5-percent level. *** Statistically significant at the 1-percent level.

24

Chapter 3. Incentives for police forces: trading off two alternative approaches

I. Introduction

Finding out how well the police perform without being knowledgeable about police work. That is the challenge facing governments trying to steer the police forces towards optimal performance. Putting a reward on good performance requires data about police activities and translating these data into a judgment of overall performance. No assessment will provide an unambiguous picture of police performance: a measure of true value added is not available at the level of individual police forces.

Approaches to holding the police accountable to results fall into two categories: objective and subjective performance assessment. In the first case, 'good performance' is defined in advance by choosing performance measures and a rule for reward (or penalty). The incentive is unambiguous. The role of government is to fix the rule for reward into some sort of contract with the police forces. Dutch police forces receive a financial reward when they meet a number of performance targets, for example. In contrast, in the case of subjective performance assessment, performance is evaluated afterwards. The role of the government is to organize the assessment by evaluators – and to make sure that the assessment is unbiased. For example, inspectors from the British HM Inspectorate of Constabulary conduct performance reviews of police forces and compose evaluation reports with recommendations. In both approaches, much of the same information could be used, the crucial difference is whether the reward is based on a rule determined *ex ante* or on subjective judgment *ex post*.

In this chapter, we discuss the tradeoff between these two alternative incentive schemes, based on the economics literature on incentives. Thus far, this literature has rarely been used to analyze incentives for the police (see Polinsky and Shavell, 2000 for an overview of the economics of public law enforcement). In the case that incentives for the police are subject of analysis, economists tend to focus on things the police should not venture into like corruption and racial profiling (see Polinsky and Shavell, 2001 and Grogger and Ridgeway, 2004, for instance). A rare exception is Prendergast (2001) who discusses the effect of media attention to police violence on crime fighting efforts of the Los Angeles Police Department. In his book about his time at the New York Police Department, Bratton (1998) points out that a focus on preventing excesses such as corruption alone is no guarantee for better overall police performance. Therefore, in this chapter, we explicitly address the question how to devise incentives aimed at improving the desirable outcomes of police work.

We focus on the interaction between the national government and the (regional) police forces. In this area, several countries are revising their policies. For example, England and Wales recently introduced an annual performance monitor of the police forces, the 'Policing

Performance Assessment Framework', and the Netherlands introduced performance contracts, which we will discuss in this chapter. We focus on performance assessments as a way of improving police performance, leaving labor conditions and other means of improving performance outside the scope.

The chapter is structured as follows. After defining the objective of police work, we discuss why incentives for police forces could be necessary. Next, based on three key dimensions of performance rewards, we discuss the tradeoffs governments face when designing incentive schemes. Then we present a case study of the Netherlands. The last section concludes.

II. The objective of police work

The government's objective for the police can be defined in general terms as the reduction of crime, promotion of safety and reduction of disorder in an efficient, effective and fair and decent manner. Efficiency is related to resource management. Working efficiently means allocating resources timely to places and times where they are most needed. Effectiveness is related to identifying and implementing best practices. Working effectively implies deploying resources in a way that they have most impact.

The imperative of fair and decent policing can be seen as a constraint to maximizing efficiency and effectiveness. The rules of fair and decent policing include limitations to the powers of investigations and the use of force and more informal rules such as providing courtesy in public contacts. These standards of conduct are constraints as they tend to slow down police work. Without such constraints, police officers would search houses without a search warrant, seek out criminals without reading them their rights, allowing them to call their attorneys, or releasing them in response to a writ of habeas corpus. Such police work would be efficient and effective, but it does not appeal to the idea of fair and decent policing.

In this chapter, we focus on efforts aimed at improving efficiency and effectiveness of policing. As stated in the introduction, so far most of the literature focused on the constraints to police work. We focus on what the police should do rather than on what the police should not do.

III. The need for a performance reward

In addition to providing the necessary resources to the police, why would the government want to reward good (or penalize bad) performance? After all, people working for the police are motivated to do their work well by factors that are intrinsic to them. People may choose to work for the police because of an inherent tendency to seek out challenges, to extend and exercise their capacities, to explore, and to learn. Similarly, the police forces can be seen as 'intrinsically motivated' to realize the government's objective in their specific region. The question is why the government needs to provide *external* incentives: positive or negative

'reinforcers' of specific types of behavior that are external to the agent. Three rationales for external incentives are the lack of rewards, distortions in incentives through environmental factors and coordination problems.

Lack of rewards

If a police force performs better than another police force, there is no automatic payoff for the efforts that went into performance improvement. The decentralized value assessments that are generally performed by market participants – customers deciding what a product is worth for instance – are disabled by the fact that the police do not engage in market transactions.[9] Police forces have a monopoly position in their region. In contrast to schools and hospitals, citizens cannot switch to a better performing police force other than by moving somewhere else. Police forces are not engaged in a struggle for customers or market share. They are not under a credible threat of substantial budget cuts for poor performance for the very same reason. By cutting the budget of forces that did not perform satisfactorily, citizens would pay twice for poor police management in their region. The lack of consumer choice puts a constraint on monetary rewards and penalties.

A reward for good performance is important since improving performance is a process of trial and error that comes with high costs. It takes considerable investment in data systems, entry of data, methods to analyze the data, and informal exchanges of experiences within and between forces. Based on the findings, painful changes may need to be made in the organization. The accompanying costs can either be financial (e.g. investment in technologies for more efficient scheduling) or intangible (e.g. stress stemming from effective personnel management). In the process, enterprising individuals within forces may have to overcome bureaucratic barriers and a continuous, strong pressure to respond to the issues of the day. When the efforts that go into improving performance go unnoticed, the drive to identify best practices may be weakened (Prendergast, 2002, p. S117).

Distorted incentives through environmental factors

Empirical studies of police behavior show that the external incentives facing the police are not always in line with the objectives of police work. An important reason is that the incentives are often the unintended consequence of factors in their environment. For instance, Prendergast (2001) sees a reduction in the Los Angeles Police Department's crime-fighting efforts as an attempt to avoid further negative media coverage and investigations. This change in strategy followed a flood of negative news about the LAPD's use of force. 'Nobody notices when the police let criminals walk, but everybody notices it when the police has to use force to arrest that same criminal' (Prendergast, 2001). Well-designed incentive schemes may be instrumental in offsetting the bias in behavior resulting from environmental factors such as media attention.

---

[9] We do not discuss the optimal level of public enforcement versus private enforcement, for a discussion and an overview of the literature see Polinsky and Shavell (2000).

Coordination problems

Coordination between police forces is an important part of improving performance. By working together, the forces can share information and best practices. All forces would be better off by doing so. Think of compatible information and communication systems and platforms for exchanging ideas and experiences. Organizing such a coordinated effort is complex. For example, in the Netherlands, there are 25 regional police forces, in England and Wales 43. In advance, it is hard to identify who exactly benefits how much from each action. For individual police forces, it is unclear what exactly they will get back from costly coordination of their efforts. The uncertain rewards at the level of the police force and the complex way to internalize them may prevent forces to work together. The drive to improve performance may be weakened if such coordination problems block a combined effort to realize the objective of police work.

Conclusion

To conclude, due to factors external to the relationship between the government and the police forces, the level of effort that the police forces put in to achieve the objective of police work may be lower than optimal and may also be misdirected. External incentives would be instrumental in sustaining the motivation to put in high effort to work efficiently (by putting the resources where they are most needed) and effectively (by deploying resources in the most effective way). That way, the government can provide a stimulus, necessary to signal exceptional performance and to overcome the multiple barriers to action. In this sense, external incentives from the government do not replace but complement intrinsic motivation.

IV. Qualities of incentive schemes

In line with the economics literature on incentives, the government can be seen as a principal and the police forces as agents. The principal commissions the agent to complete a task, in this case the objective of police work. As discussed in the previous section, there are several reasons for the government to provide incentives to the police forces. The principal-agent literature deals with the design of incentive schemes in the absence of an ideal performance measure.

Ideally, an incentive scheme is based on a measure of value added. In our case, the ideal performance measure would indicate the contribution of a police force to the reduction of crime and disorder and promotion of safety. A measure of value added would unambiguously signal a police force's level of effort relative to other police forces. Perfect monitoring of effort makes the design of an incentive scheme trivial. The reward can be directly tied to the desired outcomes of police work.

In practice, a measure of value added is not available at the level of police forces. There are just too many other actors and factors affecting the outcomes of police work to identify the contribution of the police. Measures of police output like the number of criminal

28

investigations do not provide a complete picture of a police force's performance either. Without a measure of value added, the police forces have an information advantage over the government. They have a better idea of their effort level than the government.

In the next sections, we discuss the tradeoffs related to two alternative approaches to performance rewards in the absence of the ideal performance measure. The tradeoffs are based on three qualities of incentive schemes: the signal-to-noise ratio of the performance assessment, the discretionary power left to the police forces, and distortion of effort. In this section, we introduce these three concepts.

Signal-to-noise ratio

In the absence of a measure of value added, the government needs to fall back on measures that indicate the level of effort with some uncertainty. Such measures are 'noisy', have a certain signal-to-noise ratio. The change in the crime rate is an example of a very noisy performance measure. Performance measures with a lower signal-to-noise ratio decrease the power of incentives for two reasons. First, when the return to efforts to fight crime are highly uncertain, the police may decide it is not worth putting in high effort (assuming risk aversion, see Prendergast, 1999, Section 2). Second, noisy performance measures make it easier for the police forces to hide behind unfavorable external conditions – the effects of which are largely unknown to the government. The forces may ask the government to overrule the conditions for receiving the reward (if the definition of 'good performance' is specified *ex ante*) or may try to influence perceptions of evaluators (if performance is subjectively assessed *ex post)*. Clearly, greater possibilities of using their information advantage lower incentives of police forces.

Discretionary power

The performance assessment should leave some discretionary power to the police forces to set their own priorities and to choose the best methods of achieving these regional goals. Delegating responsibility to the police forces is in the government's interest for two reasons.

First, by purposefully leaving discretion to the forces, the government can make use of expertise at the level of the police force on how to best achieve the objective. The police need detailed knowledge of local circumstances to be effective since crime and public order problems differ between localities (see Skogan and Frydl, 2004, chapter 5 for a discussion). The police needs to know who is who, who lives where, who deviates from the norm, and who had which role in what happened. They need to build up social networks and they need to be aware of cultural patterns in the neighborhood. Knowledge of local conditions is also important to weigh conflicting demands for police services. Once taxes have been paid, the price for police services is zero for the public. Given limited police resources, there is an opportunity cost to reallocating resources from one activity to another activity. Based on their

daily work in the region, police forces have better information to weigh competing calls for service than the national government.

Second, incentives of policy makers are such that setting priorities is preferably left to police forces (Goldstein, 1963). Some laws are enacted as a symbolic act. Such laws either cannot effectively be enforced or are not meant to be fully enforced (especially in the face of more pressing concerns). A policy maker may enact such laws for several reasons. Maybe she does not want to be known as someone who does not support a law against reprehensible behavior. Maybe she also does not want to be known as someone who repeals a prohibition because it is not really meant to be fully enforced. She derives her comfort from enactment of the law, not from enforcement of the law. It falls to the police to live with the law without enforcing it.

Distortion of effort

Achieving high effort levels is one thing, getting the agent to work in the right direction is another. An incentive scheme invites distortion of effort (or 'gaming' of the evaluation) when it reflects only some of the many dimensions of police work. Focus on the measured and compensated outcomes often causes less effort to be directed towards unmeasured or uncompensated goals (Hart, Shleifer, and Vishny, 1997). Distorted effort results in higher *measured* performance, but is inefficient from the perspective of the government's objective for police work. Gaming could take place through shirking on quality of a particular task but also by substituting away from police tasks that do not produce measurable results to tasks that do.[10] Missing out on some dimensions of police work is likely since the police rely heavily on informal methods of enforcement, such as simply making their presence known to potential troublemakers and stopping and questioning them. Yet the police do not record all of these actions.

V. Objective performance assessment

Based on the three qualities of incentive schemes, we discuss the tradeoffs related to objective performance assessment.

Definition

Objective performance assessment is characterized by the use of predetermined rules that translate performance measures into a reward (or penalty). 'Good performance' is defined in advance. The assessment is mechanical, there is no subjective element to it. In this sense it is objective. The rule for reward can be simple. For example, a force receives a financial reward when it meets some performance targets. In that case, each measure is weighted equally. More complex methods have also been proposed, including Data Envelopment Analysis and

---

[10] Separating the (organization and) evaluation of different police tasks according to their measurability is not a solution to this problem. Simultaneous assessment is necessary since different police tasks are related and contribute to a common goal. Currently popular community policing and problem-oriented policing reforms stress the mutual dependence of a broad range of police tasks (Skogan and Frydl, 2004, chapter 5).

Stochastic Frontier Analysis (see for instance Spottiswoode, 2000 and Drake and Simper, 2001).

Discretionary power

The degree of discretionary power for the police forces depends on the type of performance measure used in the assessment. At the one extreme are outcome measures, which leave a great degree of discretion. At the other extreme are measures of internal processes, which leave very little discretion. The farther removed from outcome measures, the more performance measures intervene in the 'internal machinery' of police work. After all, by singling out process or output measures, the government prescribes the best way of achieving the outcomes. That comes at a cost in terms of not fully utilizing the knowledge present at the level of the police forces.

Discretion is not only static in nature – in the sense of achieving the objective given current regional conditions – but also dynamic. The police forces need to decide how to best respond to *changing* regional conditions given the objective of police work. Defining 'good performance' *ex ante* – with the actual performance reward being allocated some time later – may conflict with this dynamic optimization process. After all, changing circumstances may require different priorities and policing methods.

Noise vs. distortion of effort

Baker (2002) argues that there is a tradeoff between noise of a performance measure and the degree to which it distorts effort. A higher signal-to-noise ratio – and therefore stronger incentives – tends to go together with more distortion of effort, and the other way round. The intuition is straightforward. Imagine a hierarchy of performance measures with measures of internal processes at the bottom and outcome measures at the top. Figure 1 provides a stylized picture of the tradeoff between noise and distortion of effort when moving from one type of performance measure to another. For ease of exposition, we assume that incentive schemes are based on a single type of performance measures. We will discuss combinations of performance measures later. We use assessing and rewarding criminal investigations as an example. The three points in the graph depict the best available performance measures in each category.

Below on the right are incentive schemes that are exclusively based on process measures, which are at the bottom of the hierarchy of performance measures. Timely processing of criminal charges is an example. Relative to measures of output and outcome, such measures are affected by few outside actors and factors – they are part of the internal processes within the police after all. At the same time, such performance measures tend to distort effort. They only reflect one or a few specific dimensions of police work, inviting the police to neglect other dimensions of police work. The incentive scheme provides strong incentives but could result in an effort to 'meet the numbers' rather than to improve overall performance.

Figure 1   **Trading off noise against distortion of effort**
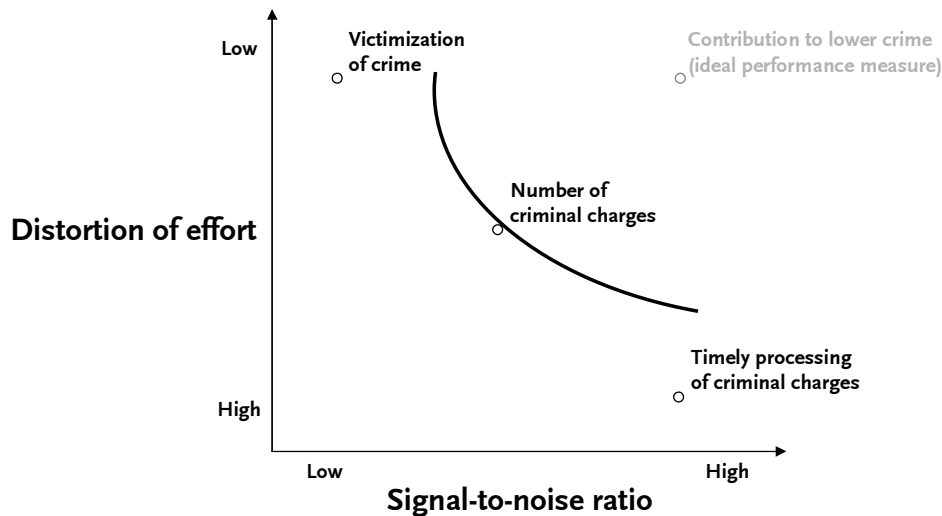
One step up are incentive schemes based on output measures. Outputs like citations, arrests and criminal charges are all within the exclusive domain of police work. They are more noisy than process measures since the interaction with the outside world is more important. As only some of the many police outputs are recorded and their quality is not easily observed, the scheme is likely to invite some distortion of effort. Police forces can substitute between measured and unmeasured actions and can shirk on (unmeasured) quality of output. The distortion of effort is lower than in the case of process measures, since the measures are closer to the ultimate goals of police work.

Above on the left are incentive schemes based on outcome measures. Outcome measures indicate the outcome of interest, such as the crime rate, but not the contribution of the police to it. These measures leave relatively little room for gaming since they encompass more of the dimensions of police work than output measures. For example, efforts aimed at bringing down victimization of crime are well in line with the government's objective. The incentives are relatively weak, however. The signal-to-noise ratio is low since these measures are no longer within the exclusive domain of police work.

Thus the higher the power of incentives, the more likely the scheme is to distort effort. The tradeoff between noise and distortion is one of the reasons to limit the power of incentives. The choice of which performance measure to use depends on how the amount of distortion and the power of incentives change as one moves from one performance measure to another. The government's preferences are depicted by the curved line. The curve includes the points at which the government is indifferent between some loss of incentives and some distortion of effort. Clearly, levels closer to the upper right corner are preferred. Given its current shape, the number of criminal charges provides the best option. When the government would put

higher value on preventing distortion of effort, the indifference curve would rotate counter clockwise. In that case, an outcome measure like the crime rate might be best.

Clearly, the government can choose any possible combination of outcome, output and process measures. A combination of measures may provide a better position on the tradeoff between noise and distortion of effort than measures within a particular category. A greater number of performance measures also reflect a greater number of dimensions of police work after all. With multiple performance measures, the government needs to weigh each measure according to its relative social value. Clearly, it is difficult to choose such weights (see Stone, 2002 for a discussion). The more so, since people in different police regions may value police activities differently. Thus, even when – in theory – the performance measures would reflect all essential quality dimensions, then some distortion of effort is still likely since the weights may not accurately reflect the social value of each performance measure.[11]

Figure 2   Comparing policing to welfare-to-work services
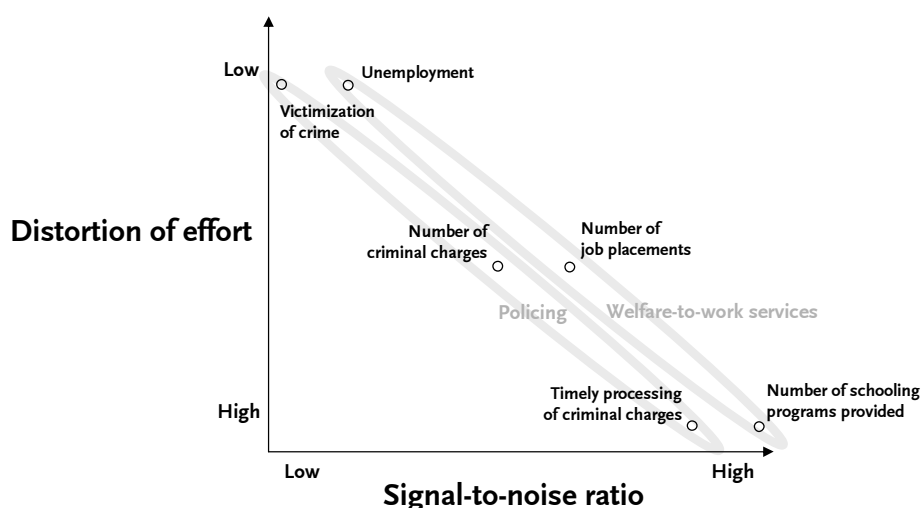
How does the police compare to other sectors?

Choosing performance measures that provide strong incentives without too much distortion of effort is particularly challenging in the case of policing. Not only is the police engaged in many activities ranging from assistance at traffic accidents to resolving a hostage stand off, each individual activity involves many different dimensions as well. The problem seems to be more complex than in the case of, say, welfare-to-work services. These services are aimed at an easily observable output: placement into jobs. In practice, budgeting of employment service providers tends to be based on this output indicator (current practice in Australia and the Netherlands), whereas we see no output-based budgeting in the case of policing. This difference in policy provides some indication that distortion of effort is greater in the case of

---

[11] Khalil and Lawarrée (2001) point out that while more information is typically better, the principal might not use all the available performance measures. The reason could be that these measures are costly to observe or that it is too costly to make all of them verifiable.

police forces than in the case of employment placement agencies – keeping the signal-to-noise ratio constant. Figure 2 illustrates the difference between policing and employment services. Compared to policing, the tradeoff between the power of incentives and distortion of effort is more favorable in the case of employment services.

Conclusion

Including performance measures with a high signal-to-noise ratio such as the number of fines in the definition of 'good performance' provides strong incentives, but tends to invite distortion of effort and to limit discretionary power of the police forces. Thus, when designing a rule for reward, the government faces a trade-off between power of the incentive scheme on the one hand, and distortion of effort and discretionary power on the other hand.

VI. Subjective performance assessment

Objective performance assessment may not provide for satisfactory choices in the case of policing. There may not be a set of performance measures that provides sufficiently strong incentives – without too much distortion of effort and with sufficient discretionary power. If this is the case, then subjective performance assessment may provide an alternative basis for the performance reward. In this section, we discuss the characteristics of subjective performance assessment and the tradeoff between the two approaches.

Definition

In the case of subjective performance assessment, the evaluation is not fully codified *ex ante*. The step from performance measures to an assessment of overall performance is not (fully) based on predetermined rules, but leaves room for subjective judgment. Examples are performance assessments by the police inspection, peer reviews between police forces, and benchmarks of police forces that leave it to the reader to make his or her own judgment of the relative performance of a police force.

Greater discretionary power

Subjective performance assessment leaves greater discretionary power to the police forces than objective performance assessment, because it demands *ex post* accountability rather than meeting *ex ante* rules how to improve performance. In a subjective evaluation, a police force's management needs to explain how they responded to the specific regional problems, and to what avail. Discretionary power is particularly important since regional conditions can also change during the evaluation period. Subjective assessment leaves room for police forces to respond to changing circumstances, whereas objective performance assessment tends to limit this flexibility, as we discussed in the previous section.

Possibly less distortion of effort

Subjective performance assessment allows a more complete picture of performance to be attained compared to objective performance assessment. Evaluators can include dimensions

of police work that cannot easily be measured and ask for further explanation when doing their work since the assessment is *ex post*. By including more of the dimensions of the work into the evaluation, subjective assessment may limit possibilities for gaming, common to objective performance assessment.

Since the assessment has a subjective element to it, it may invite new types of gaming by either the forces or the evaluator. Evaluators tend to be not only well-informed but also well-known in the field. They may include active police personnel (in the case of peer review) and may have repeated personal interaction with the forces (think of the police inspection). As a result of social pressure, evaluators may be reluctant to give bad ratings (the so-called 'leniency bias', see Prendergast, 1999). Moreover, the perspective of the evaluators may not fully coincide with those of the government. There may be inherent biases in the perceptions of evaluators, they may not take all dimensions of police work into account. Consequently, the forces may show inefficient behavior to please the evaluators.

Figure 3   Trading off objective vs. subjective performance assessment
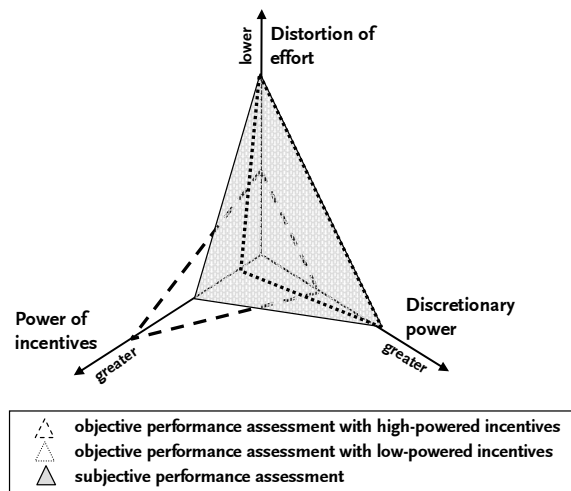
The tradeoff

When choosing between objective and subjective performance assessment, again the government faces a tradeoff between power of incentives on the one hand and distortion of effort and discretionary power on the other hand. Figure 3 illustrates the tradeoff in a three-dimensional graph, with a movement away from the origin being 'better'. The striped planes are two extreme types of objective performance assessment: one based on outcomes measures, the other on measures of internal performance. As discussed in the previous section, a gain in the power of incentives tends to go together with a loss in the distortion of effort and discretionary power.

The third plane depicts subjective performance assessment. As shown in the figure, subjective performance assessment may provide a superior alternative to the possible set of objective

performance assessments. On the one hand, the power of incentives of a subjective assessment is lower than possible with the objective approach. Ticking off whether a police force meets a number of performance targets is less noisy in nature than a subjective assessment of overall performance. On the other hand, subjective performance assessment may alleviate concerns about distortion of effort, while providing great discretionary power to the police forces at the same time. More dimensions of police work can be included in the evaluation and the evaluation leaves room for the management to respond to (changing) regional conditions.

Thus, the subjective approach may provide a way out of the unfavorable tradeoff related to the objective approach. Whether this is the case depends on the new possibilities for gaming related to the subjective nature of the assessment. The performance of the subjective approach in terms of distortion of effort may be worse than figure 3 suggests. *A priori* it is not obvious which method of performance assessment produces a behavioral response that is most in line with the government's objective for the police.

Table 1   **Subjective and objective performance assessment in police practice: key differences**

	Objective performance assessment	Subjective performance assessment
Type of evaluation	Checking whether force meets *ex ante* definition of 'good performance'	Critical evaluation of achievements *ex post*
Source of incentives	Performing well according to pre-determined rule how to receive a reward	Performing well in the eyes of evaluators
Role of government	Fixing definition of good performance in a contract with the police forces, oversee availability of reliable performance data	Organizing subjective assessment, oversee availability of reliable performance data
Discretionary power	Depends on how much room definition of 'good performance' leaves for own priorities and methods	Depends on room evaluators provide for regional response to specific regional circumstances during evaluation period
Determinants of success	Information to choose performance measures and rules that determine 'good performance', reliable, non-manipulated performance data	Unbiased evaluators, reliable, non-manipulated performance data
Practical form	Performance targets, quantitative methods of (relative) performance evaluation	Police inspection evaluations, peer review, benchmarks

Table 1 provides an overview of the key differences between subjective and objective performance assessment. As stated in the introduction, much of the same information could be used in both approaches, the crucial difference is whether the translation to a reward is based on a pre-determined rule or on subjective judgment *ex post*.

VII. The case of the Netherlands

Recent experience in the Netherlands illustrates two points of the analysis: first, the need for external incentives for the police forces and second, the tradeoff related to objective performance assessment of police work.

The need for external incentives

Until recently, the 25 regional police forces faced few external incentives from the national government. The government followed a hands-off approach: it distributed the police budget between the forces and left it to them to realize the government's objective in their specific regions. The police inspection was very cautious in its approach, did not clearly state when and how shortcomings within police forces needed to be addressed, and had no possibilities for intervention (Netherlands Court of Audit, 2002). The Interior Secretary called it 'the pussyfoot Inspection' (House of Parliament, Year of Session 1997–1998, TK 103–7244). A peer review between police forces every four years was initiated by the police as part of its system of quality management. Lacking management by the national government, the quality of the reviews differed greatly from force to force. The review was somewhat bureaucratic in nature and received limited attention from management and policy makers, including the Interior Department (Police Inspection, 2001). As part of its policy cycle, the government sets national priorities for police work, but these priorities had little practical implications beyond the paperwork that came with it (Terpstra, 2002). The government conducted a formal check whether national priorities were included in the regional policy plans. Given a complete lack of reliable information on policing activities, the government was not able to check actual compliance with the national priorities. Finding out how many FTEs the forces employed turned out to be an independent research effort, *reliable* and *comparable* figures on basic items like response times to emergency calls, allocation of police time to activities, and even offenses per crime category were simply not available (Netherlands Court of Audit, 2003a). Most of the priorities were put in such general terms that each force paid attention to these priorities anyway (In 't Veld et al., 2001 and Chief of police Meijboom in *Trouw*, 04-01-2003).[12] In practice, the police forces enjoyed a great level of discretion in setting priorities and choosing methods of achieving them.

Along the lines of the previous discussion, the lack of external incentives resulted in a lack of coordination *between* the forces and a weak drive to improve performance *within* the police forces. Police personnel only processed information that was of direct relevance to the own police force. A patchwork of information and communication technology (ICT) systems hampered efficient sharing of information between the police forces and other parties within the criminal justice system. Each police force developed its own computer applications. The forces operated as 'islands' ignoring the inter-regional character of many crime problems (Netherlands Court of Audit, 2003b). Even when developing ICT together at the urge of the Interior Department, there was no party with the authority to make final decisions on the way to go forward and keeping 'dissident' forces in line (Netherlands Court of Audit, 2003b). One

---

[12] The lack of external incentives does not seem to be filled at lower levels of government. Since the size of the police regions does not fit with either municipalities or counties ('provincies'), democratic control at this level is hampered. At the local level, the police had a greater handle on the municipality than the other way round, mainly because of the great information advantage of the police (Terpstra, 2002). Generally, mayors are not actively involved in drafting a local policing plan, they rather act like 'buyers of police services' (Struiksma and Winter, 2001).

consequence was that information on almost any police activity could not be compared between forces. The lack of coordination was also clear in policing methods. Several police forces independently developed their own approach to community policing (Boin et al., 2003). Generally, sharing of best practices between the forces was limited: there were few platforms to do so.

Within the forces, there did not seem to be a strong drive towards maximizing efficiency and effectiveness. Resources were not allocated to places and times where they were most needed. Generally, allocation of resources was not based on an analysis of regional problems (Netherlands Court of Audit, 2003a). There did not seem to be a culture of experimentation and evaluation either. Information on primary police activities was incomplete and unreliable because computer applications were old-fashioned and not user friendly (Netherlands Court of Audit, 2003b). Lacking systematic collection and analysis of performance data, the forces had difficulty identifying the most effective way of deploying resources. The lack of reliable information on any dimension of performance itself indicated that improving police performance was not high on the agenda. The goals of experimental projects (such as an alternative way of tackling burglary) tended to be redefined along the way, and not followed by an evaluation. If there was an evaluation, then it usually did not meet minimal scientific standards (see Bovenkerk (2001) for a discussion of small-scale evaluations within the police and Wittebrood and Van Beem (2004) for a review of the Dutch evaluation literature over the last twenty years).

To conclude, the Dutch experience over the period 1993-2002 supports the argument that external incentives from the national government are a necessary supplement to intrinsic motivation of police personnel within the police forces. Without such incentives from the government, there is a lack of coordination between the police forces and a limited drive to improve performance within the police forces.

The tradeoff related to objective performance assessment

In the late 1990s, the national government started to become more involved. In December 2000, the government asked the police to advise on the design of an incentive scheme. The police presented a performance reward that was based on a great number of mostly noisy outcome measures. The scheme included a complicated rule for awarding the financial bonus, partly based on individual achievements and partly on performance relative to other, similar forces. The financial bonus was equal to two percent of the police budget. Given the noisy nature of many of the performance measures, receiving the bonus was to some extent a 'chance-event'. Combined with the many ways in which a force could earn some part of the bonus, there was a reasonable chance that most forces would be better off. Payment of the performance reward was repeatedly put off because of the unreliability of performance data. The rewards were distributed for the first and last time in 2003. Again, the discussion focused on data issues: the Interior Secretary disqualified six police forces for receiving a reward

because of the unreliability of their performance data (Letter from the Interior Secretary to the Parliament (EA2003/87623) dated December 10, 2003). Given the low signal-to-noise ratio of many measures and these data issues, the financial reward did not do much to strengthen incentives. It can be seen as a first step towards more powerful external incentives from the national government.

In 2003, the government presented a performance contract that substituted the outcome measures in the above-mentioned scheme with a smaller set of less noisy performance measures. The aim was to create stronger incentives without introducing too much distortion of effort. The force-specific performance targets for the period 2002-2006 are based on output measures (fines, charges forwarded to the prosecution council), subjective performance measures (customer and citizen satisfaction with police work) and measures of internal performance (timely processing of criminal charges, overall efficiency, sickness absence, quality of telephone services). If a force meets all the targets, it receives a modest financial bonus and favorable publicity.

Stronger incentives

There are strong indications that the shift towards performance measures with a higher signal-to-noise ratio increased the power of incentives. In contrast to the first, largely outcome-based incentive scheme, the police forces are very vocal in expressing their feelings about the possible negative impact on their daily work, both at the management level and at the level of individual police officers. Apparently, in their experience, meeting the targets means doing police work in a different way. The police forces are very concerned about the practical implications of the shift towards stronger incentives.

The actual incentive to put in high effort depends on the level of ambition of the targets. In the absence of any firm knowledge about the means by which the targets can be achieved, setting the targets right is not easy. Clearly, incentives are weakened if targets are out of reach or too little ambitious. Given the information advantage of the police forces, it is most likely that most of the targets are 'workable', which is also the impression of several observers (Arie Kuijvenhoven, former police commissioner, *de Volkskrant*, 01-10-2005). For example, the target for the number of fines seems little ambitious. Producing 180,000 additional fines annually implies some 19 additional fines per active police officer, which is not a major achievement. Additional indication is that targets set by some forces go beyond the national requirement. The target for the number of criminal charges provides a more nuanced picture. The situation in base-year 2002 happens to affect the strength of incentives. For some forces it turns out to be relatively easy to meet the desired increase in the number of charges by about 20 percent. These forces had a low number of charges in base-year 2002 relative to earlier years or relative to other forces. Forces that produced a relatively high number of charges in 2002 will have a harder time meeting the target.

Less discretionary power

As argued previously, a higher signal-to-noise ratio of performance measures tends to limit the discretionary power of the forces. In fact, the government decides where and how the forces should improve their performance. The targets could limit their flexibility in responding to changing regional circumstances. For example, over the period 2002-2006, the forces should produce 40,000 additional criminal charges in total. As discussed above, the target number of criminal charges may be challenging for some forces. By 2006, these forces may be forced to look into not-so-serious criminal charges they otherwise would not have bothered with, simply because they are out of more serious cases by that time. It is questionable whether this possible side effect of the performance contracts contributes to police efficiency. Similarly, some Amsterdam police officers state that the yearly target number of fines does not track changes in safety in the city. Since the city is becoming safer, police officers have to fall back on frivolous fines to meet their yearly target (statement of the Amsterdam Police Union (APV), *de Telegraaf*, 08-17-2004).

Greater distortion of effort

A higher signal-to-noise ratio of performance measures also tends to result in greater distortion of effort. The police may redirect effort from one activity to another for the sake of meeting the targets. For example, there is no mention of preventative activities in the performance contracts (think of patrolling 'hot spots' and providing prevention advice to citizens and businesses). Some observers voice concerns that the police are being rewarded for substituting away from prevention to producing criminal charges. The discussion goes back to a conflict between some chiefs of police and the prosecution council, with the police arguing for more and better preventative strategies and the council pushing for more criminal investigations ('OM-baas kritiseert politie', *NRC Handelsblad*, 03-22-2005). At this point, it is unclear how serious the threat of undervaluation of preventative activities is.

Within activities that affect the measures, the contracts may invite the police forces to shirk on the quality of output. There are concerns that the target for the number of criminal investigations would invite the forces to concentrate on the easiest cases. When weighing each case with the severity of the crime, five police forces showed a substantial decline in their weighted output in 2003 and 2004.[13] The prosecution council keeps a close eye at the numbers, and voices concerns when necessary, which may explain that four forces after an initial drop in their weighted output rebounded in 2004. Overall, the severity of cases forwarded to the prosecution council did not decline relative to 2002 (Wiebrens, 2005). At this point, it is too early to conclude that some police forces gamed the evaluation. In some cases, the performance targets seem to lead to frustration, undermining intrinsic motivation. A district chief in police force Gelderland-Midden was replaced when he refused to remove a

---

[13] The severity of a crime is defined as the number of days someone is on average incarcerated for having committed that crime. Other sanctions such as financial penalties are translated into 'prison day equivalents', following the guidelines of the prosecution council. Thanks to Caspar Wiebrens for providing the data.

backlog in the number of fines necessary to meet the target in the performance contract ('Weigeren bonnenregen kost commissaris baan', *de Telegraaf*, 11-08-2004).

To conclude, experience in the Netherlands provides some support for the existence of a tradeoff between power of incentives on the one hand and distortion of effort and discretionary power on the other hand. A shift towards performance measures with a higher signal-to-noise ratio seems to strengthen incentives, but also to limit the discretion of the forces and to invite them to shift effort from unmeasured to measured activities and to lower the quality of the measured activities. To prevent such gaming, the government is closely monitoring the behavior of the forces, which seems to have paid off in the case of the number of criminal charges. Since the performance contracts concern the period until 2006, it is too early to draw hard conclusions about their effects on police behavior.

VIII. Conclusions

Devising incentive for police forces is necessary given a lack of rewards for effort within the police organization, environmental factors such as media attention that distort incentives, and coordination problems between the forces. Without incentives from the national government, the police forces tend to show a lack of coordination and a limited drive to improve their performance, as the Dutch experience illustrates.

When devising incentives, the government faces a choice between two alternative approaches. The first approach, objective performance assessment, defines 'good performance' in advance. There is a clear rule that defines when a force receives a reward. In the second approach, subjective performance assessment, subjective judgment closes the gap between performance data and a judgment of a force's contribution to the objective of police work. Given the multidimensional nature of police work and the prevalence of non-discrete outcomes, the objective approach proves to be difficult. There is a tradeoff between power of incentives on the one hand and distortion of effort and discretionary power on the other hand. Subjective performance assessment may provide a way out by alleviating concerns about distortion of effort, while leaving more room to respond to (changing) regional conditions. More dimensions of police work can be included in the evaluation. These benefits may be partly offset by new possibilities for gaming related to the subjective nature of the assessment. The case study of the Netherlands shows that objective performance measurement for police forces is indeed characterized by an unfavorable tradeoff, with stronger incentives inviting distortion of effort. It is however too early to draw hard conclusions about the net benefits of the performance contracts that were introduced in 2003.

Bibliography

Baker, George P., 2002, Distortion and risk in optimal incentive contracts, *Journal of Human Resources*, 37 (4), 728-751.

Boin, Arjen, Paul 't Hart and Edward van de Torre, 2003, Blauwe Bazen. Het leiderschap van korpschefs [Blue bosses, the leadership of chiefs of police], Politie en Wetenschap, Rapport No. 11, Apeldoorn.

Bovenkerk, Frank, 2001, Jaarcijfers over criminaliteit en haar bestrijding. Dwalen door het onbekende [Annual data on crime and law enforcement. Wandering through the unknown], Police force Amsterdam-Amstelland, Amsterdam.

Bratton, William with Peter Knobler, 1998, *Turnaround. How America's Top Cop Reversed the Crime Epidemic*, Random House, New York.

Corman, Hope, and H. Naci Mocan, 2000, A time-series analysis of crime, deterrence, and drug abuse in New York City, *American Economic Review*, 90 (3), 584-604.

Corman, Hope, and H. Naci Mocan, 2005, Carrots, sticks, and broken windows, *Journal of Law and Economics*, 48 (1), 235-266.

Di Tella, Rafael, and Ernesto Schargrodsky, 2004, Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack, *American Economic Review*, 94 (1), 115-133.

Drake, Leigh M., and Richard Simper, 2001, The economic evaluation of policing activity: an application of a hybrid methodology, *European Journal of Law and Economics,* 12 (3), 173-192.

Dryden Witte, Ann, and Robert Witt, 2001, What we spend and what we get: public and private provision of crime prevention and criminal justice, *Fiscal Studies*, 22 (1), 1-40.

Dubourg, Richard, Joe Hamed, and Jamie Thorns, 2005, The economic and social costs of crime against individuals and households 2003/04, Home Office, London.

Goldstein, Herman, 1963, Police discretion: the ideal versus the real, *Public Administration Review*, 23, 140-148.

Grogger, Jeffrey, and Gregory K. Ridgeway, 2004, Testing for racial profiling in traffic stops from behind a veil of darkness, *mimeo*, RAND Corporation, Santa Monica.

Hart, Oliver, Andrei Shleifer, and Robert W. Vishny, 1997, The proper scope of government: theory and application to prisons, *Quarterly Journal of Economics*, 112 (4), 1127-1161.

Hsiao, Cheng, 2003, *Analysis of Panel Data*, second edition, Cambridge University Press, Cambridge.

In 't Veld, Roel, Frank A. Beemer, Wouter de Haan, Ferdinand Mertens, Ed Romein, and Marijke van Roost, 2001, Vooruitgang of regendans? Evaluatie Beleids- en Beheerscyclus Politie [Progress or rain dance? Evaluation of the Policy and Management Cycle], Bestad Consultancy and Berenschot Procesmanagement, The Hague.

Khalil, Fahad, and Jacques Lawarree, 2001, Catching the agent on the wrong foot: ex post choice of monitoring, *Journal of Public Economics*, 82 (3), 327-347.

Klick, Jonathan, and Alexander Tabarrok, 2005, Using terror alert levels to estimate the effect of police on crime", *Journal of Law and Economics*, 48 (1), 267-279.

Kovandzic, Tomislav. V., and John J. Sloan, 2002, Police levels and crime rates revisited. A county level analysis from Florida (1980-1998), *Journal of Criminal Justice*, 30 (1), 65-76.

Levitt, Steven D., 1998, The relationship between crime reporting and police: implications for the use of Uniform Crime Reports, *Journal of Quantitative Criminology*, 14 (1), 61-81.

Levitt, Steven D., 2002, Using electoral cycles in police hiring to estimate the effect of police on crime: reply, *American Economic Review*, 92 (4), 1244-1250.

MacDonald, Ziggy, 2002, Official Crime Statistics: Their Use and Interpretation, *The Economic Journal*, 112 (477), F85-F106.

Manski, Charles F., 1993, Identification of Social Endogenous Effects: The Reflection Problem, *Review of Economic Studies*, 60 (3), 531-542.

Marvell, Thomas B., and Carlisle E. Moody, 1996, Specification problems, police levels, and crime rates, *Criminology*, 34 (4), 609-646.

Mundlak, Yair, 1978, On the Pooling of Time Series and Cross Section Data, *Econometrica*, 46 (1), 69-85.

Netherlands Court of Audit, 2002, Handhaving door Rijksinspecties [Enforcement by National Inspections], TK 28271 Nr. 2, The Hague.

Netherlands Court of Audit, 2003a, Zicht op taakuitvoering politie [Perspective on police performance], TK 28791 Nr. 2, The Hague.

Netherlands Court of Audit, 2003b, ICT bij de politie [ICT within the police], TK 29350 Nr. 1, The Hague.

Philipson, Tomas J., and Richard A. Posner, 1996, The economic epidemiology of crime, *Journal of Law and Economics*, 39 (2), 405-433.

Police Inspection, 2001, *De eerste cyclus zit er op. Een evaluatie van vier jaar kwaliteitszorg* [An evaluation of four years of quality management], The Hague.

Polinsky, A. Mitchell, and Steven Shavell, 2000, The economic theory of public enforcement of law, *Journal of Economic Literature*, 38 (1), 45-76.

Polinsky, A. Mitchell, and Steven Shavell, 2001, Corruption and optimal law enforcement, *Journal of Public Economics*, 81 (1), 1-24.

Prendergast, Canice, 2001, Selection and oversight in the public sector, with the Los Angeles Police Department as an example, NBER Working Paper 8864, Cambridge (MA).

Prendergast, Canice, 2002, Uncertainty and incentives, *Journal of Labor Economics*, 20 (2), S115-S137.

Rogerson, Pete, Rajan Batta, and Christopher M. Rump, 2003, Detection and prediction of geographical changes in crime rates, US Department of Justice, Washington D.C.

Skogan, Wesley, and Kathleen Frydl (eds.), 2004, *Fairness and effectiveness in policing. The evidence*, National Research Council, The National Academies Press, Washington D.C.

Spottiswoode, Claire, 2000, *Improving police performance. A new approach to measuring police efficiency*, Public Services Productivity Panel, UK Treasury, London.

Stone, Mervyn, 2002, How not to measure the efficiency of public services (and how one might), *Journal of the Royal Statistical Society A*, 165 (3), 405–34,

Struiksma, Niko, and Heinrich B. Winter, 2001, Bevoegdheden overd(r)acht. Een onderzoek naar delegatie en mandaat van beheersbevoegdheden in de politiepraktijk [Delegating management power within the police], Politie en Wetenschap, Apeldoorn.

Terpstra, Jan, 2002, Sturing van politie en politiewerk. Een verkennend onderzoek tegen de achtergrond van een veranderende sturingscontext and sturingsstijl [Managing police and police work], Politie en Wetenschap, Apeldoorn.

Tauchen, Helen, Ann Dryden Witte, and Harriet Griesinger, 1994, Criminal deterrence: revisiting the issue with a birth cohort, *Review of Economics and Statistics*, 76 (3), 399-412.

Vollaard, Ben A., 2004, Budgetverdeling leidt tot verschillen in politieprestaties [Police funding formula results in inequalities in police performance], *ESB*, 89 (4429), 135-137.

Weisburd, David, and John E. Eck, 2004, What can police do to reduce crime, disorder and fear?, *The Annals of the American Academy of Political and Social Science*, 59 (3), 42-65.

Wiebrens, Caspar, 2005, Verschillen, vergelijken, verbeteren [Differences, comparing, improving], in: Prosecution Council, *Goed Beschouwd 2005* [All things considered 2005], The Hague, 21-40.

Wittebrood, Karin, and Matthijs van Beem, 2004, Sociale veiligheid vergroten door gelegenheidsbeperking: wat werkt en wat niet? [Improving safety: what works and what doesn't?], Working Paper 2, Social and Cultural Planning Office, The Hague.

Wittebrood, Karin, and Marianne Junger, 2002, Trends in violent crime: a comparison between police statistics and victimization surveys, *Social Indicators Research*, 59 (2), 153-173.

Wooldridge, Jeffrey M., 2002, *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge.

About the author

Ben Vollaard (1975) graduated in Economics at Erasmus Universiteit Rotterdam in 1997. Between 1997 and 2000 he worked with the Netherlands Department of Economics and the CPB Netherlands Bureau for Economic Policy Analysis. His research focused on public sector reform. In 2000, he started to pursue the Ph.D. program in Policy Analysis at the Pardee RAND Graduate School in Santa Monica, California. At RAND, he was involved in many research projects ranging from modeling gas exploration in the Rocky Mountains to devising scenarios for electricity demand in the next twenty years. His dissertation work focuses on effectiveness of the police. He conducted most of his dissertation work at the CPB Netherlands Bureau for Economic Policy Analysis. In November and December 2005, he concluded his research efforts during an invited stay at the Institute of Criminology of the Hebrew University in Jerusalem.