Identifying Efficient Crime-Combating Policies by VAR Models: The Example of Switzerland

Patricia Funk and Peter Kugler*

Current research suggests that the crime-combating instrument sentence probability is more effective than sentence severity. However, the focus of the simultaneous (or very short-run) impact of the law enforcement policy on crime impedes a comparison of these two instruments with respect to their long-term effectiveness. With Swiss data, this article investigates the dynamic interrelationships between enforcement policy and crime and finds that overall, sentence severity is about the same effective as sentence probability. Furthermore, the authors show how the VAR-modeling technique can be exploited to conveniently distinguish between deterrence and incapacitation effects. More concrete recommendations toward a cost-effective crime-reduction policy can be derived. (JEL C32, K42)

I. Introduction

The seminal economic model of crime (Becker, 1968), claiming that crime can be reduced by increasing either the probability or the severity of punishment, has triggered a huge amount of empirical research. On one hand, the aim was to test the empirical relevance of Becker’s model of crime; on the other hand, information was desired to be gained concerning the relative effectiveness of higher sentence probability and harsher sentence severity. Obviously, in view of the growing amount of resources governments are spending on crime deterrence, knowledge about the crime-reducing power of these two policy instruments is crucial to achieve optimal resource allocation. More precisely, a hint as to whether crime should be combated by spending more on police (thereby increasing the probability of a sentence) or by convicting to longer prison sentences (thereby increasing the severity of a sentence) would be of greatest importance. Unfortunately, results published so far allow for no clear-cut policy recommendations.

Studies based on individual data suggest that the number of rearrests for released prisoners decreases with longer prison sentences experienced in the past as well as higher percentages of convictions (given arrest).1 However, the specific sample selection makes more general recommendations for crime policy difficult. More interesting for this purpose are studies based on aggregate data. Findings from cross-sectional analyses mostly indicate that sentence probability is more effective in reducing crime than sentence severity (i.e. the [absolute] elasticity of the former is larger).2 However, several flaws inherent in all these cross-sectional analyses (such as the impossibility of controlling for unobserved heterogeneity among the countries) fostered the emergence of panel data estimates with fixed effects. In panel data studies, differences between sentence probability and severity became even more apparent. Whereas higher arrest (and conviction) rates still showed a crime-inhibiting impact, the effect of longer

1. See, for example, Witte (1980), Trumbull (1989), Grogger (1991), and Tauchen et al. (1993); an exception is Myers (1983).
2. The earliest study is Ehrlich (1973). For an overview of later studies, see Cornwell and Trumbull (1994).
prison sentences turned out to be rather ambiguous. (Sentence severity was insignificant in Cornwell and Trumbull, 1994, and Mustard, forthcoming, negative in Viren, 1994, even positive for certain crimes in Marselli and Vannini, 1997, and not explicitly incorporated in Levitt, 1998a).

A primary goal of this article is to examine whether the seemingly greater effectiveness of sentence probability still holds in a dynamic context. For instance, although the above-mentioned studies suggest that higher sentence severity has a smaller simultaneous effect on crime compared to sentence probability, its impact may be longer-lasting in return. Obviously, a meaningful comparison of these two crime-reducing policies requires information about the overall (short- and long-run) effectiveness of sentence probability and sentence severity.

Second, this article presents a new method that enables the distinction between deterrence and incapacitation effects. Longer prison sentences as well as more frequent convictions (to prison) potentially reduce crime by two different channels: first, by deterring more offenders from committing crimes due to the higher expected penalty (deterrence effect); second, by having (more) convicted offenders (longer) imprisoned, who cannot commit any crimes during imprisonment (incapacitation effect). Because incapacitation is costly, a policy instrument that is able to reduce crime by deterrence rather than imprisonment is ceteris paribus preferable. Therefore, by separating deterrence from incapacitation effects, this article is able to give more precise directions toward a cost-effective crime-reduction policy than prior studies could.

Because the strength of law enforcement is possibly endogenous, the authors consider vector autoregressions (VARs) as a promising method for the stated purpose and apply them to Swiss property crime data (theft and robbery). The resulting impulse-response functions allow one not only to conveniently analyze the dynamic effects of a shock of sentence severity and probability on crime but also to distinguish deterrence from incapacitation effects. For instance, a shock on sentence severity cannot reduce crime by incapacitation of offenders before the (prior to the shock) customary prison sentence has been served. Therefore, any prior occurring effect must be attributed to deterrence. To give a concrete example, assume that a shock on sentence severity increased incarceration time from six months of prison to nine months. Therefore, if such a shock reduced crime immediately, it must have been due to deterrence, because the incapacitation effect of the longer prison sentence comes into force no earlier than after six months of prison.

This study refines current research in that it is the first one that considers the dynamic (deterrence and incapacitation) effects of sentence probability and severity.

The main findings from the estimations are the following:

1. In contrast to the latest studies, the authors find that both sentence probability and sentence severity are effective in reducing crime. Furthermore, accounting for the dynamic impact of these two crime-reducing measures, the article cannot conclude that sentence severity is less effective than sentence probability.

2. Whereas a shock in the sentence probability seems to reduce crime mainly through a higher number of incapacitated offenders, the benefits of a shock on sentence severity are deterrence as well as incapacitation. For theft offenses, substantial deterrent effects could even be found for an increase in the length of conditional prison sentences, which are similar to American probation sentences.

3. The standard procedure to distinguish between deterrence and incapacitation effects is to consider cross-crimes correlations: “As long as criminals are repeat offenders and do not specialize in one particular type of crime, an increase in the arrest rate for any one crime will lead to a reduction in all crimes due to the incapacitation effect. In contrast, if criminals are rational and different crimes are substituted for one another, the deterrence effect implies that increasing the arrest rate for one crime will lead to a decrease in that crime, but to an increase in other crimes as criminals substitute away from the first crime, the ‘price’ of which has risen” (Levitt, 1998b, p. 316). Because this procedure depends on the assumption that criminals are repeatedly committing different types of crimes, the authors consider the procedure, which exploits the time structure of the data and does not rely on any assumption about the “type of offender,” as a valuable complement.

4. There is only one VAR estimation focusing on crime and deterrence. However, sentence severity is not incorporated as a variable, nor are the different deterrence and incapacitation effects distinguished (see Corman et al., 1987). Similarly, governmental law enforcement is covered by the number of arrests and the police force but not by a measure of sentence severity in Corman and Mocan’s (2000) time-series analysis.
The rest of the article is structured as follows: section II describes and discusses the development of property crime as well as its punishment in Switzerland. The econometric model and the estimation results are presented in section III. Some robustness checks are discussed in section IV. The article concludes in section V.

II. DESCRIPTION OF THE DATA

While selecting variables, the authors kept in mind a Becker-type model of crime (Becker, 1968), in which an individual’s incentive to commit a crime depends on the probability of getting a sentence and the severity of the sentence. Therefore, the model assumes that a potential criminal fears a conviction as well as the related penalty, but not necessarily an arrest, if no conviction occurs. Although there might exist a certain fear of getting arrested even if later released, the main concern about committing a crime is the potential of being convicted and punished.

Therefore, the authors use a measure for the probability of getting convicted (or equivalently the probability of getting sentenced, because every conviction generally results in a sentence) rather than the probability of arrest. Although it seems plausible to approximate the probability of conviction by the number of convictions divided by the number of crimes, there are econometric problems associated with this measure. The reason is that the definition of this variable may induce an artificial negative correlation between crimes and the sentence probability estimate. Consider for illustrative purposes the simple case in which convictions are constant over time and have no causal impact on crime. If convictions per crimes are regressed on crimes, a negative correlation is observed even if there is no causal effect of convictions on crimes (for a discussion of this potential artificial negative correlation in the presence of measurement error, see Levitt, 1998b).

Therefore, the authors use the number of convictions as a measure for sentence probability. Because past crimes are controlled for in the regressions, changes in convictions given crimes can be interpreted as changes in sentence probabilities.

Next to the number of convictions, the authors include a measure for sentence severity in the regressions and also account for the labor market conditions by considering the number of unemployed people. Although the inclusion of a variable reflecting the economic situation is standard in the empirical literature on crime, the econometric advantage lies in the ability to account for the trend observed in the crime rates.

In this study Swiss quarterly time-series data covering the years 1984–98 for the following (two) vectors of \( n = 4 \) variables are analyzed

\[
(1) \quad (z_t^j) = [(ss_t^j), (cr_t^j), (co_t^j), (un_t^j)], \quad j = 1, 2,
\]

where \((y_t^j) = [(ss_t^j), (cr_t^j), (co_t^j)]\) represent the endogenous variables and \(x_t = (un_t)\) the exogenous variable.

The first vector \((j = 1)\) contains data for theft offenses. Thereby, \(ss\) abbreviates sentence severity, \(cr\) stands for (theft) crimes, \(co\) corresponds to the number of convictions occurring for theft, and the number of unemployed is given by \(un\). As a second crime, the authors consider robbery offenses. The second vector \((j = 2)\) therefore relates crime, sentence severity, and the number of convictions to robbery offenses.

To be precise about the definitions of the crimes, Swiss theft offenses include larceny-thefts as well as burglaries. Robbery offenses are defined as property crimes with force or threat of violence and are hence compatible with the robbery definition used in the Federal Bureau of Investigation’s uniform crime reports.

Figure 1 and 2 depict the time series of the crime data used in this study. As can be seen therefrom, the quarterly crime rates (per 100,000 inhabitants) were on average 740 for theft offenses and 7 for robbery offenses, which amounts to yearly values of roughly 3,000 theft offenses per 100,000 inhabitants and 28 robbery offenses per 100,000 inhabitants. Compared to the United States, Switzerland had similar burglary and larceny-theft rates in 1998 but a much lower robbery rate (in 1998, around 165 robbery offenses per 100,000 inhabitants were committed in the United States).

As far as punishment of criminal offenses is concerned, prison sentences were nearly exclusively used as sanctions for theft and robbery offenses. However, the Swiss legal system

5. In Switzerland, a person is registered as unemployed if unemployment benefits are applied for.
knows two different kind of prison sentences: conditional and unconditional. A conditional sentence means that an offender does not have to serve the prison sentence immediately, but only if he or she is reconvicted for an offense within a certain amount of time after the initial conviction. In the case of reconviction, the number of prior conditional prison days are converted to unconditional prison days and the sentence for the second crime is added to the sentence for the first crime. In contrast, offenders convicted to an unconditional prison sentence cannot escape being incarcerated and have to enter prison after conviction.

Figures 3 and 4 show that conditional prison sentences were predominantly used for sanctioning theft offenses and unconditional prison sentences for sanctioning robbery offenses. Due to this difference in the use of sentences, the authors measure sentence severity for theft offenses by the average number of conditional prison days theft offenders (convicted to a conditional prison sentence) received in a specific quarter. In contrast, the average number of unconditional prison days (for a person convicted to robbery and unconditional imprisonment) is used as a measure of sentence severity for robbery offenses. However, alternative specifications of sentence severity are discussed later.
Figures 5 and 6 show the graphs of the current measures of sentence severity. Although the average sentence severity for theft was 68 days (or roughly 2 months), sentence severity for robbery took on 1,244 days on average (or about 3.7 years); see also Table 1 for the summary statistics of variables.

Figures 7 and 8 depict the measure for sentence probability, that is, the total number of convictions for theft and robbery offenses. As can be seen from Figure 7, there is a relatively sharp decline in the number of convictions for theft offenses after 1994. The reason for this is a change in the law for punishment of minor theft offenses. Since the beginning of 1995, a revision of criminal law allowed minor theft offenses to be considered violations and not criminal offenses and therefore to be newly punished by fines. Because these convictions do not appear in the traditional conviction statistics, the authors introduce a dummy variable (which takes values of one after 1995) in the theft-VAR to account for this possible structural break.

Finally, a graph of the number of unemployed people is given in Figure 9 and some summary statistics of the variables can be found in Table 1.

### III. NONTECHNICAL DESCRIPTION OF THE ECONOMETRIC METHOD AND PRESENTATION OF THE RESULTS

VARs are used for estimating dynamic interrelationships between different potentially endogenous variables. If, for example, one is interested in the dynamic interrelationships between sentence severity and crime, VARs allow one to simulate how a shock in sentence severity affects crime over time and how a shock on crime affects sentence severity. To be able to identify these interrelationships, one needs to estimate the contemporaneous and lagged impact of each variable on the others (while controlling for the dependent variable’s own lagged effects). Unfortunately, it is not possible to distinguish the contemporaneous impact of

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theft per 100,000 inhabitants (per quarter)</td>
<td>740</td>
<td>638</td>
<td>925</td>
<td>55</td>
</tr>
<tr>
<td>Robberies per 100,000 inhabitants (per quarter)</td>
<td>7</td>
<td>4.6</td>
<td>10</td>
<td>1.56</td>
</tr>
<tr>
<td>Sentence Severity Theft (Average cond. prison days)</td>
<td>68</td>
<td>48</td>
<td>86</td>
<td>8.7</td>
</tr>
<tr>
<td>Sentence Severity Robbery (Average uncond. prison days)</td>
<td>1244</td>
<td>696</td>
<td>2343</td>
<td>358</td>
</tr>
<tr>
<td>Convictions for Theft (per quarter)</td>
<td>2110</td>
<td>1463</td>
<td>2752</td>
<td>308</td>
</tr>
<tr>
<td>Convictions for Robbery (per quarter)</td>
<td>95</td>
<td>54</td>
<td>133</td>
<td>18</td>
</tr>
<tr>
<td>Number of unemployed persons</td>
<td>85,957</td>
<td>15,602</td>
<td>204,686</td>
<td>67,880</td>
</tr>
</tbody>
</table>
sentence severity on crime from the opposite effect without any further identifying assumptions. Therefore, it is common to estimate a reduced-form VAR as a first step (which ignores the contemporaneous effects) to get unbiased coefficient estimates of the different lagged variables. The reduced form regressions can be represented as follows (for simplicity, the authors suppress the distinction between the two offenses):

\[
y_t = \mu + \sum_{l=1}^{k} A_l y_{t-l} + \sum_{l=0}^{k} C_l x_{t-l} + \epsilon_t,
\]

with intercept vector \(\mu\), VAR coefficient matrices \(A\) and \(y\) and \(x\) as defined in section II. Due to the omission of contemporary effects, the vector of error terms is expected to be orthogonal to the regressors, which allows one to estimate unbiased coefficients. However, because \(\epsilon_t\) has a nondiagonal covariance matrix \(\Sigma\), there is generally no structural interpretation of the error vector.

To get a structural interpretation of the residuals, the authors impose certain restrictions on the contemporaneous relationships between crime rates, convictions, and sentence severity as a second step. With these restrictions, the structural VAR (accounting for contemporaneous effects) was recovered from its reduced form and inverted it to its (very conveniently interpretable) impulse response form. In the impulse response form, each endogenous variable is represented by a weighted sum of the (contemporaneous and lagged) shocks of all endogenous variables as well as the level of the exogenous variables, which makes it straightforward to simulate the dynamic effects of a shock in one variable on the others.

Because the interested reader may get the details about the identification of VARs and estimation procedure from Appendix A, the authors would just like to mention that they identify the structural VAR by using the Choleski decomposition and by posing three restrictions on the contemporaneous relationships between crime, sentence severity, and convictions. One restriction seems very plausible because of the use of quarterly data: It is
very unlikely that sentence severity is contemporaneously affected by an unexpected change in crimes. Although crime may have an influence on sentence severity, the authors think that its impact occurs with a considerable lag: Statistics on offenses are published with a substantial delay. Therefore, it may take some time (at least more than one quarter) before courts eventually adjust sentence severity to a changing degree of criminal activity. Furthermore, sentence severity seems to be contemporaneously unaffected by the number of convictions (restriction 2). This can be seen from the correlation matrix of the reduced form residuals (see Appendix B), in which no significant correlation between convictions and sentence severity is found. Finally, the authors assume as a third restriction that the number of offenses is contemporaneously unaffected by the number of convictions (rather than the opposite). However, we will check the robustness of the results with respect to changing this identifying assumption.

The impulse response functions under the above-mentioned identifying assumptions are now presented. Note further that the authors control for seasonality by adding three seasonal dummies and set the lag length to 4, which is optimal according to the Akaike criterion considering lags up to 5. Due to the relatively small sample size, the authors further use the standard errors of the impulse response coefficients based on Monte Carlo simulations because the ones derived from asymptotic distributions might be downward biased.

A. The Determinants of Probability and Severity of Punishment

Although the authors are primarily interested in explaining the amount of theft and robbery crime, they start with analyzing the determinants of the two policy variables, sentence severity and convictions. The goal is to gain some information concerning the exogeneity of the policy instruments with respect to the level of crime. Furthermore, the responses of the policy instruments show the magnitude of the policy shocks, whose effects on crime are simulated.

The impulse response functions depicted in Figure 10 indicate the responses of the variables sentence severity and convictions to unexpected shocks in sentence severity, convictions and crimes. The magnitude of the shock of each variable amounts to one standard deviation of its own past (unexpected) variation (i.e., the residual in the respective variable’s VAR equation). Therefore, the impulse response functions are simulations of unexpected changes in one variable, say, sentence severity, on the other variables, whereas the magnitude of the shock equals the sentence severity’s variation in past shocks. As such, one can see from the impulse-response functions which effect unexpected changes in, for example, past sentence severity had on crime.

On the horizontal axis in Figure 10, the number one denotes quarter one (or the contemporaneous effect), the number two quarter two, and so on. Obviously, because this identifying assumption is that sentence severity is contemporaneously unaffected by crimes and convictions, there is no response of sentence severity to shocks in crimes and convictions in period one. The interpretation of the responses is straightforward. Because the VAR is estimated in linear form, the vertical axis indicates the change in the absolute value of the considered variable (as described in Table 1) after a shock occurred. For instance, the middle picture of Convictions of Theft shows that a one-standard-deviation innovation in thefts per 100,000 inhabitants (which corresponds to a value of 20) lead to roughly 20 more convictions about four quarters later. Because the magnitude of the (one-standard-deviation) shock of a certain variable (say, theft) is equal to the response of the variable to its own contemporaneous shock (Response of Theft to Theft), a simple inspection of the responses of one variable to its own shock at period one gives us the magnitude of the one standard-deviation shock (which equals 20 in the case of theft; see Figure 11, first row, middle picture).

Overall, the two upper impulse-response functions in Figure 10 show that the severity of punishment does hardly depend on the amount of offenses committed (middle column). Therefore, the severity of governmental sentencing seems to be exogenous with respect to the level of crime. In contrast, the amount of convictions reacts to an increase in the crime level at least for theft offenses, with a peak response four quarters later. In spite of this positive reaction, it remains unclear whether more convictions result because more offenders facilitate catching one of them or because more resources are invested in police.
Next to the level of crime, deterrence might be driven by the amount of governmental enforcement in earlier periods. However, the impulse-responses show no longer-lasting autocorrelation either for sentence severity or convictions. As such, unexpected changes in deterrence are likely to happen quite rapidly.
B. The Dynamic Impact of Enforcement on Crime

The authors start with a detailed investigation of the impact of a shock in sentence severity on the offenses committed (first column, Figure 11). For theft offenses, an increase in the sentence severity exerts a negative effect on crime, which remains strong until seven quarters later. Obviously, because sentence severity refers to conditional prison sentences for this type of offense, the negative impact on crime most likely reflects deterrence.6 As such, an increase in the sentence length of a one-standard-deviation shock (i.e., roughly three-days-longer prison sentences) has reduced the number of theft offenses per 100,000 inhabitants by roughly five over several quarters. Although this number seems to be small at first glance, one has to bear in mind that the authors are simulating the response of a temporary and rather small shock in sentence severity. The effect of a permanent increase in sentence severity would undoubtedly be larger. However, this article sticks to the simulation of temporary shocks because it facilitates the distinction between deterrence and incapacitation effects. Furthermore, the primary interest here is in comparing the relative effectiveness of sentence severity and sentence probability.

For robbery offenses, a shock on sentence severity may reduce crime by deterrence as well as incapacitation. However, because the average incapacitation time for robbery offenses comes to 15 quarters, the incapacitation effect of a shock on sentence severity can occur at the earliest after the average prison sentence of about 15 quarters has been served. Therefore, the whole impact of a shock on sentence severity on the offenses committed in the next 12 quarters can be explained solely by deterrence. As such, an increase in the sentence severity by

---

6. An incapacitation effect of longer conditional prison sentences might only occur for the recidivists, who are caught and whose (longer) conditional prison sentences are turned into longer unconditional prison sentences. A deterrent effect occurs through different channels. Though an increase in the length of unconditional prison sentences undoubtedly increases the costs of the first crime, an increase in the sentence length of conditional prison days affects the expected costs of the second crime. Hence, longer conditional prison days are expected to reduce crime through a lower degree of recidivism. Also, criminals who consider committing multiple crimes might be deterred by longer conditional prison sentences.
roughly 250 days (see Figure 10 for the magnitude of the shock) has been found to exert an immediate and persistent deterrent effect. However, the impulse responses are not always significantly different from zero. Furthermore, in contrast to theft offenses, one cannot reject the null hypothesis of no Granger causality of sentence severity on robbery crimes.7

In contrast to an increase in sentence severity, a shock on the amount of convictions does not show an immediate impact (neither for theft nor for robbery). Therefore, it seems that an increase in sentence probability does not deter that much. Rather, the dynamic response of the offenses to a shock on convictions seems to reflect the effect caused by incapacitation of offenders. Considering that in Switzerland, it often takes a couple of months until a convicted person gets imprisoned,8 the drop in the impulse-response function after some quarters fits the incapacitation hypothesis quite well (remember that the authors use the total number of convictions). The subsequent pattern of the offense responses also supports the suspicion that the driving force behind the sentence probability’s crime-reducing impact is incapacitation. For theft, the average incapacitation time lasts three quarters. Accounting for a lag of one to three quarters between conviction and imprisonment, one would expect the incapacitation effect to be strongest between one/three and four/six quarters after the shock on convictions occurred—a pattern that can be exactly observed.

The incapacitation hypothesis seems to be equally confirmed for robbery offenses. A lack of immediate response to increased convictions casts doubt on potential deterrence effects. Rather, the lagged drop in the offenses is likely to coincide with the gap between conviction and imprisonment.

IV. ROBUSTNESS OF THE RESULTS WITH RESPECT TO ALTERNATIVE SPECIFICATIONS

A. Alternative Identifying Assumptions

As a first robustness check, the authors change the previously taken identifying assumption that crimes are contemporaneously unaffected by convictions but not vice versa. Figure 12 depicts the impulse-response functions for the crime equation under the assumption that convictions are contemporaneously unaffected by crimes rather than the reverse.

As can be seen, the impulse-response functions look very similar to the ones in Figure 11. Generally, it can be shown that changing the identifying restrictions does not have a big impact on the shape of the impulse responses (all the different estimations are available from the authors on request).

B. Alternative Measures of Sentence Severity

In the preceding analysis, the authors measured sentence severity for theft offenses by conditional prison days and sentence severity for robbery offenses by unconditional prison days. However, Figures 3 and 4 illustrate that roughly 30% of convictions occurring for theft were punished by unconditional prison sentences and about 40% to 50% of the robberies by conditional imprisonment.

Therefore, it is interesting to see whether a change in the sentence length for unconditional prison days had any impact on theft offenses or a change in the sentence length of conditional prison days on robbery offenses. As can be seen from Figure 13, an increase in the sentence length of unconditional imprisonment is less effective in deterring theft offenses than an increase in conditional prison sentences and about 40% to 50% of the robberies by conditional imprisonment.

For robbery offenses, Figure 14 shows that even an increase in the sentence length of conditional imprisonment had a certain crime-reducing effect. As such, the Swiss system, which gives first-time offenders a chance to

534 CONTEMPORARY ECONOMIC POLICY
FIGURE 12
Determinants of Crime: Impulse-Responses under Different Identifying Assumptions

Theft
Response of THEFT to SSTHEFT
Response of THEFT to COTHEFT
Response of THEFT to THEFT

Robbery
Response of ROBBERY to SSROBBERY
Response of ROBBERY to COROBBERY
Response of ROBBERY to ROBBERY

FIGURE 13
The Impact of Unconditional Imprisonment on Theft

Response of THEFT to SSunconditional
Response of THEFT to THEFT
Response of THEFT to COTHEFT

FIGURE 14
The Impact of Conditional Imprisonment on Robbery

Response of ROBBERY to SSconditional
Response of ROBBERY to ROBBERY
Response of ROBBERY to COROBBERY
escape prison as long as they are not reconvicted, seems to deter offenses, thefts as well as robberies.

C. Testing the Effect of Unemployment on Crime

As a last specification test, the authors simulate the responses of the crimes to a shock in unemployment. Although it is theoretically not straightforward that an increase in the number of unemployed persons results in an increase in crime, the opposite would be against the economic intuition all the same. Figure 15 depicts the response in theft and robbery offenses to a shock in unemployment. Although the effect is more distinct for robbery offenses, property crime seems to increase after a shock in unemployment, with a peak response occurring between two and five quarters after the shock.

V. CONCLUSIONS

An analysis of the dynamic crime-reducing impact of harsher law enforcement (via the instruments sentence probability and sentence severity) provided very interesting results. First, in contrast to the findings from the latest panel data studies, the authors observed a substantial crime-reducing impact of harsher sentence severity. Comparing the overall dynamic effectiveness of harsher sentence severity and higher sentence probability, this article questions the so far postulated superiority of sentence probability over sentence severity.

Second, the authors exploited the VAR estimation technique to distinguish between incapacitation and deterrence effects. The ability to differentiate between incapacitation and deterrence effects is crucial for judging an instrument’s cost-effectiveness and for deriving concrete policy recommendations. Because incapacitation is costly, an instrument that is able to exert a large deterrent effect next to incapacitating the offenders is ceteris paribus preferable.

The present estimation results indicate that sentence probability mainly reduces crime by incapacitation of offenders, whereas sentence severity exerts an incapacitation as well as a deterrent effect. Therefore, an increase in sentence severity rather than sentence probability bears the advantage that crime is reduced not only by incapacitating offenders but also by deterring potential offenders from crime due to the higher expected penalty. However, to be able to give concrete policy recommendations, more information about the costs of increasing sentence probability and sentence severity would be needed.

Summing up, these authors think that the study shows that VAR models are an interesting tool in the economic research on crime. The major advantage of VAR estimations, compared to traditional panel data or cross-sectional analyses, lies in the better understanding of the crime-reducing impact of harsher governmental enforcement. Especially the convenient distinction between (dynamic)

FIGURE 15

The Impact of Unemployment on Crime
deterrence and incapacitation effects allows one to get a step further toward concrete policy recommendations.

APPENDIX A: TECHNICAL APPENDIX

Before starting with the estimations, the authors addressed the issue of a possible nonstationarity of the times series. To this end the authors did a unit root testing exercise for the four series collected in z, which gave the following results. Two variables, sentence severity (only for robbery) and convictions, appear to be stationary or mean reverting—note that there is a structural break in the convictions for theft crimes in 1995 caused by a change in the practice to substitute prison sentences by fines (which do not yet appear in the traditional conviction statistics); therefore, the relevant unit root test statistics are the ones in Perron (1989). The two remaining series, crimes and unemployment, seem to be trending, but it is not clear whether it is a unit root (stochastic) trend or a deterministic trend. Under these circumstances it appears reasonable to analyze the level variables and account for the trend in the crimes by the exogenous unemployment variable.

As a first step, the following reduced-form VAR model is estimated

\[
y_t = \mu + \sum_{j=1}^{k} A_j y_{t-1} + \sum_{j=1}^{k} C_j x_{t-j} + \epsilon_t.
\]

Because the contemporaneous effects are omitted, the error vector \(\epsilon\) has no structural interpretation if the variables are contemporaneously related (see Appendix B for the correlation matrix of the VAR residuals). To identify structural shocks the authors have to formulate additional restrictions on the VAR system. One approach consists in postulating a linear relationship between the reduced form errors and structural errors,

\[
\epsilon_t = B u_t,
\]

where \(B\) is a (restricted) \((n \times n)\) matrix and the \(u\) vector is a white noise vector with identity covariance matrix. Thus, the covariance matrix of the reduced form errors \(\epsilon\) is given by

\[
\sum = B' B.
\]

This equation allows one to recover \(n \cdot (n+1)/2\) coefficients of the matrix \(B\). Thus, least \(n \cdot (n - 1)/2\) restrictions are needed to identify \(B\). Sims (1980) restricted the matrix \(B\) to be lower triangular to obtain exact identification, an approach, which since then has been used routinely in VAR analyses (also referred to as Choleski decomposition). However, such a recursive structure is often not convincing from the viewpoint of economic theory. Thus, the authors check whether such a recursive structure of the contemporaneous relationships of the variables considered can be justified by a priori reasoning. Otherwise they have to use a nontriangular \(B\) matrix in the framework of the so-called structural VAR model pioneered, among others, by Bernanke (1986).

In this case, a triangular matrix \(B\) with appropriately ordered variables is justifiable because sentence severity is plausibly not contemporaneously affected by an unexpected change in crimes and convictions (see section II for a more detailed argumentation). Furthermore, by assuming that either convictions are unaffected by crimes (or the opposite), the authors are able to get a triangular matrix. As such, based on theoretical arguments and an inspection of the correlation matrix of the VAR residuals \(\epsilon\), a lower triangular structure of the \(B\) matrix seems to be justified in the VAR application.

Finally, the authors would like to mention that they estimated this VAR in logarithmic as well as in linear form. Because the explanatory power of the two different models is similar and the impulse-response functions do not differ substantially, the article only depicts the easier interpretable estimations from the linear version.

\[
\text{APPENDIX B: TABLE 2}
\]

| Correlation Matrix of the VAR Residuals \(\epsilon\) |
|---------------------------------|---------------|---------------|
|                               | SS-Theft      | Theft         | Co-Theft      |
| SS-Theft                      | 1             | -0.24         | 0.002         |
| Theft                         | -0.24         | 1             | 0.29          |
| Co-Theft                      | 0.002         | 0.29          | 1             |
| SS-Robbery                    | 1             | -0.17         | -0.06         |
| Robbery                       | -0.17         | 1             | -0.12         |
| Co-Robbery                    | -0.06         | -0.12         | 1             |

REFERENCES


———. “Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or


