Challenges in Collecting Police-Reported Crime Data

Colin Babyak
Statistics Canada

Abstract

The Uniform Crime Reporting Survey (UCR) collects administrative data on all police-reported criminal incidents in Canada. In an ideal world, extraction programs are integrated with police record management systems to automate this collection and ensure timely analysis and dissemination. Unfortunately, in the real world of crime data, missing data occurs as police services change or upgrade their systems, new legislation is introduced, data needs evolve over time and crime never takes a break to let all the UCR players become synchronised. This paper describes the structure of the UCR and how we treat methodological issues such as multiple collection instruments, data validation and quality assurance. Particular emphasis is placed on recent changes to the UCR and how these changes and issues differ from typical business surveys due to the institutional nature of the respondents.

Keywords: Administrative Data, Institutional Survey, Crime Statistics

1. Introduction

The UCR is the main vehicle in Canada for collecting data on police-reported crime. Although many police services release their own statistics on crime, the UCR is the only national program that allows comparisons across jurisdictions using consistent definitions and scoring rules. The UCR has been in production since 1962 and has evolved significantly since that date, particularly over the last two decades. Originally an aggregate-only survey that was used to tabulate some basic counts, the UCR has evolved into a large microdata survey with extensive cross-Canada coverage that is able to address various research and policy-related issues.

2. Overview of the UCR

The UCR is an administrative survey which collects information that is extracted from the records management system (RMS) of police services across the country. Unlike many administrative data sources, however, Statistics Canada has had some input into the development of the RMSs and the extraction procedures that take place to convert the data into a format usable by Statistics Canada. Originally, the UCR was a pencil and paper survey used to tabulate some basic counts. The current version of the survey, UCR 2.2, is a microdata survey and will be the focus of this paper. There are in fact three versions of the microdata survey in production: UCR 2.0, UCR 2.1 and UCR 2.2. These three can be referred to collectively as UCR 2, as their basic designs are very similar. To build the original microdata version of the survey (also referred to as UCR 2.0), national data requirements (NDRs) were created in the late 1980s in consultation with the policing community. These NDRs outlined a variety of data elements to be captured at the incident level. Respondents to the UCR are all police services in Canada and all data collected refer to criminal incidents that are reported to these police services. For the reference year 2006, 90% of all criminal incidents were collected via the microdata version of the survey. This coverage is expected to extend to close to 100% within a year or two.

As mentioned previously, the UCR collects information on incidents, which includes such data elements as the type of crime that was committed (violation), the date, time, location, presence of a weapon, etc. Also collected is information on both the accused (if one has been identified) and the victim (in the case of violent crime). Data elements for these latter two include, for example, age, sex, encrypted name, and relationship between the victim and the accused. This information comes to Statistics Canada in three distinct electronic files for each respondent for each month of the year. These files are subjected to a stringent set of edits to ensure data quality. These edits can result in each incident (and associated accused and victim records) being classified as accepted, accepted with warning(s) or rejected. In the latter two cases, respondents are informed of the data quality problems via a monthly edit report which is sent to each respondent generally within 24 hours of receiving their data.

1 An incident refers to a “set of connected events which usually constitute an occurrence report”. For more details on UCR terminology, see Statistics Canada (2006).
At the end of the year, all data are compiled and a set of verification tables are sent out to all respondents for the purpose of verification. These tables contain counts by offence and persons charged. Once all of these statistics have been verified, the UCR data is released to the general public via The Daily. Data is released at two levels: aggregate data for each individual respondent and more detailed microdata-derived information for all respondents pooled together. The main statistics released include the crime rate (number of incidents per hundred thousand population), the clearance rate (effectively the proportion of incidents that are "solved") and the breakdown by age (adult/youth) and sex of the accused. Below is an excerpt from the publication of the 2003 Crime Statistic Tables (Statistics Canada 2004) showing a breakdown of weapon type by offence for all UCR respondents.

Table 3.4
Violations against the person involving firearms by type of firearm present*, 2003

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Handgun</th>
<th>Fully Automatic</th>
<th>Sawed-off Rifle</th>
<th>Rifle</th>
<th>Shotgun</th>
<th>Other Firearm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>number</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>percentage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homicide</td>
<td>80</td>
<td>70.6</td>
<td>0.0</td>
<td>3.8</td>
<td>13.3</td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td>Criminal Negligence/Other Violations Causing Death</td>
<td>4</td>
<td>20.0</td>
<td>0.0</td>
<td>0.0</td>
<td>75.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Attempt/Conspire Murder</td>
<td>111</td>
<td>72.1</td>
<td>7.2</td>
<td>3.8</td>
<td>11.7</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Sexual Assault - Total</td>
<td>43</td>
<td>28.1</td>
<td>14.0</td>
<td>4.7</td>
<td>14.0</td>
<td>9.3</td>
<td></td>
</tr>
<tr>
<td>Aggravated Sexual Assault</td>
<td>2</td>
<td>50.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>Sexual Assault with Weapon</td>
<td>21</td>
<td>70.2</td>
<td>4.8</td>
<td>4.8</td>
<td>14.3</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Sexual Assault</td>
<td>20</td>
<td>40.0</td>
<td>25.0</td>
<td>5.0</td>
<td>15.0</td>
<td>16.0</td>
<td></td>
</tr>
<tr>
<td>Non-Sexual Assault - Total</td>
<td>1,102</td>
<td>42.4</td>
<td>3.0</td>
<td>3.7</td>
<td>12.3</td>
<td>38.6</td>
<td></td>
</tr>
<tr>
<td>Aggravated Assault</td>
<td>51</td>
<td>84.3</td>
<td>3.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>Assault with Weapon/ Causing Bodily Harm</td>
<td>812</td>
<td>39.6</td>
<td>1.6</td>
<td>3.9</td>
<td>10.0</td>
<td>44.7</td>
<td></td>
</tr>
<tr>
<td>Assault</td>
<td>136</td>
<td>37.6</td>
<td>11.8</td>
<td>4.4</td>
<td>24.3</td>
<td>22.1</td>
<td></td>
</tr>
<tr>
<td>Discharge Firearm with Intent</td>
<td>101</td>
<td>47.5</td>
<td>2.0</td>
<td>2.0</td>
<td>18.5</td>
<td>29.7</td>
<td></td>
</tr>
<tr>
<td>Assault People OffFloor</td>
<td>11</td>
<td>83.6</td>
<td>0.1</td>
<td>0.0</td>
<td>18.2</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Other Assaults</td>
<td>3</td>
<td>33.3</td>
<td>0.0</td>
<td>11.1</td>
<td>22.2</td>
<td>33.3</td>
<td></td>
</tr>
<tr>
<td>Assault - Total</td>
<td>1,103</td>
<td>43.0</td>
<td>3.4</td>
<td>3.7</td>
<td>12.4</td>
<td>37.5</td>
<td></td>
</tr>
<tr>
<td>Other Sexual Offences</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Kidnapping/Hostage Taking</td>
<td>232</td>
<td>70.7</td>
<td>4.3</td>
<td>4.7</td>
<td>8.0</td>
<td>5.2</td>
<td></td>
</tr>
<tr>
<td>Abduction</td>
<td>1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Robbery</td>
<td>3,106</td>
<td>81.3</td>
<td>1.4</td>
<td>5.2</td>
<td>5.2</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td>Extortion</td>
<td>28</td>
<td>0.7</td>
<td>3.0</td>
<td>14.3</td>
<td>10.7</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>Criminal Harassment</td>
<td>27</td>
<td>37.0</td>
<td>29.6</td>
<td>14.8</td>
<td>11.1</td>
<td>7.4</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>321</td>
<td>40.5</td>
<td>4.7</td>
<td>8.4</td>
<td>20.2</td>
<td>15.3</td>
<td></td>
</tr>
</tbody>
</table>

Total Violations Against the Person | 5,076 | 69.6    | 2.5             | 5.1             | 8.6   | 14.3    |              |

Source: Incident-based UCRS, Policing Services Program, Canadian Centre for Justice Statistics.

Note: Data are provided from a non-representative subset of 122 police departments accounting for approximately 41% of the national volume of crime.

1. Where there is more than one weapon present in an incident, the most serious weapon is captured.
2. Other firearms include all firearm-like weapons (e.g., starter pistols, flare gun, pepper gun) as well as unknown types of firearms.

2 The Daily is the main release vehicle for Statistics Canada publications and is an on-line publication available at www.statcan.ca.
3. Comparing the UCR and a “Typical” Business Survey

3.1 Similarities

The UCR is considered an institutional survey and thus is more similar to a business survey than it is to a household survey. For example, like a typical business survey, the UCR respondent population is highly skewed, with a small number of respondents representing a large share of the variables of interest (in this case the number of incidents taking place). In the UCR, there are approximately 1200 respondents (also called police services). The six largest (according to number of incidents) are the Toronto, Montreal, Edmonton, Winnipeg, Vancouver and Calgary Police Services and together, they account for approximately one quarter of all police-reported crime in Canada. It is worth noting that the population policed by these 6 services is just slightly less than one quarter of the total Canadian population, showing the high correlation between population policed and the number of criminal incidents.

In addition to having a skewed population, another similarity between the UCR and a typical business survey involves the frame. For the UCR, it is of very high quality and there exists rich auxiliary information (i.e., population policed), which is similar to that of a business register. As well, police services do not change significantly in size from year to year, especially so for those services which are already large to begin with. Another analogy that can be drawn between the UCR and business surveys is that the people responding are representing their organizations; supplying data to Statistics Canada is often part of their job requirements. It should be noted, however, that also like a business, reporting data to Statistics Canada is not the highest priority for police services. This makes it especially important to maintain positive relationships with the data providers.

Another analogy that can be drawn between the two involves the users of the data. Like a business survey, UCR respondents are among the key users of the data, enabling them to compare their own trends with those of the rest of the population. Similarly, when all the statistics are aggregated, there is considerable interest from the general population on basic statistics such as the overall crime rate.

The UCR is essentially an administrative survey which uses each police service’s RMS with minimal added requirements on the part of the respondent. This is similar to the use of tax data by business surveys for replacing or imputing respondent data. Statistics Canada is simply collecting data that the respondent already needs to capture for another administrative purpose.

A final similarity between the UCR and business surveys is the relationship that Statistics Canada has with the largest respondents. In the case of large businesses, since they have such a large impact on their industry and/or province, it is important to maintain an ongoing personal relationship to ensure maximum cooperation. The situation with the UCR is much the same in that close contact is maintained with many of the large municipal respondents as well as provincial and federal police services. Statistics Canada staff communicate regularly with the police services to answer questions, work through data issues and to provide various levels of support.

3.2 Differences

While many of the similarities derive from the nature of the respondent population, one does not have to examine these similarities too closely to see that the situations are far from identical. Although the composition of the frame for the UCR is similar to a business register, there are also significant differences. The frame for the UCR is more up-to-date than a typical business register. Births and deaths become known to Statistics Canada staff almost immediately, the frame coverage is virtually 100% at all times and the geographic boundaries are known for every respondent right down to the census subdivision (CSD). The population for the UCR is also much simpler than a business population. With the exception of the Royal Canadian Mounted Police (RCMP), there is no respondent that operates in more than one province. For every other respondent, the area and population that they police is clearly defined with little or no overlap. Additionally, Statistics Canada receives data at the individual police service level so that no disaggregating of data need be done, as is commonly the case with a large enterprise reporting on behalf of many establishments.

Another basic difference between the UCR and a typical business survey is that the UCR is a census as opposed to a sample survey. As a mandatory survey, the UCR also has high compliance with 100% of all respondents supplying data in most years. There are exceptions to this as will be discussed later, but it is a

---

3 A census subdivision (CSD) is a census term that is roughly equivalent to a municipality. In Canada, an “average” CSD contains approximately 6,000 residents, although this varies considerably from province to province and from CSD to CSD.
rarity indeed for any police service to “refuse” to supply data. Along with the high quality frame mentioned previously and the census design, these differences give the UCR an advantage over a typical business survey in that estimation is greatly simplified, there is no sampling error and treatment of random nonresponse is usually not required.

As mentioned earlier, extracting data from the RMS of a police service is akin to the use of tax data. The key difference here, however, is that Statistics Canada has played a significant role in the development of both the administrative database and the extraction software. While many administrative sources have the shortcoming of definitional differences with survey concepts, this is not the case with the UCR.

Not all differences however are advantageous. For example, unlike a business survey, when data is released at year-end, it is released not only at the national, provincial and local levels, but also at the individual respondent level. This can be problematic for a variety of reasons. First, any errors at the aggregate level that make it through the production process are available for all to see. Data analysis must be very thorough to avoid this type of potential embarrassment. Second, while a business survey can revise data at an aggregate level regardless of where a mistake might have occurred, any adjustment to UCR data can be seen at the respondent level. Third, the treatment of missing or erroneous data is not straightforward. When outliers are discovered in the UCR data or when a respondent is unable to supply some or all of its data, Statistics Canada representatives will attempt to contact the police service to determine the nature of the problem. This is done on an ongoing basis throughout the year and more vigorously at the end of the year. Despite all attempts to assure high quality data and to correct any detected errors, it may not be possible to resolve all of these issues before the data are released. As a result, Statistics Canada is left with one of two options: modify the data or leave it as is. In the case of the latter, it is not feasible to omit respondents or to permit obviously erroneous data to appear at the respondent level in the final publication. However, if the data is modified or imputed, then the individual police service can argue that Statistics Canada is “making up” data for their particular service and then publishing it as if it was its own. Neither solution is particularly attractive, but this is an issue that Statistics Canada faces relatively frequently with the UCR. In the end, a solution is usually agreed upon after consultation with the individual police service.

Another significant difference between the UCR and a typical business survey is the speed with which the collection instrument can be changed. The UCR is currently in the field under four separate versions: an aggregate version that has been in place since 1962, the first microdata version called UCR 2.0 (1988), the revised microdata version called UCR 2.1 (1995) and the most recent microdata version called UCR 2.2 (2004) which collects data on new crimes (hate crime, organized crime and cyber crime). As a result, updating the survey can have serious repercussions for respondents that have not yet converted to the newest version. Typically, some small modifications can be made to the current version to incorporate new offence codes or new edit rules. In the case of the new crimes which are collected by the UCR 2.2, the changes that were required were so significant that a new version was required to incorporate them all. However, simply because a new version of the survey exists, there is no requirement for respondents to start using it. Typically, when a respondent is ready to update their RMS (usually due to their own operational requirements), they will also update to the most recent version of the UCR. Unfortunately, this means that some respondents will lag significantly behind others in the version of the survey they are using. Currently, there are a handful of smaller, mostly rural police services that are reporting to the aggregate version of the survey and plans are in the works to convert these respondents to the newest version of the survey. Additionally, there remains one respondent reporting to UCR 2.0. The important point to note here is that these two versions were essentially “replaced” 19 and 12 years ago respectively, but we still receive data in these formats. It is likely that such a lag will always be present with the UCR and some sort of conversion software will always be necessary to maintain consistency among respondents.

One of the key operational measurements of any Statistics Canada survey is the response rate, usually defined as the number of sampling units that responded divided by the number of sampling units that were asked to respond. For most business surveys, there are a number of well-defined ways to calculate the response rate depending on the objective (i.e., weighted, un-weighted, at the enterprise level, etc.) For the UCR, however, this is not as obvious. At first glance, we typically have a 100% compliance rate with the UCR, meaning that virtually every respondent sends us some data in every year. However, since each respondent submits a variable number of records in each time period, whenever they do not respond to the survey, it is not known how much data is missing. The most likely cause of a police service not responding is due to technical reasons resulting from switching
RMSs. This situation has many implications including the fact that it becomes very difficult to calculate a traditional response rate. Clearly, the impact of a large city like Toronto failing to respond is much more significant than having a small, rural respondent unable to submit their data. The question remains as to how a response rate should be calculated that is both accurate and representative in both of these cases.

4. Data Quality

There are a variety of measures in place to ensure that data problems are minimized in the UCR. Every month, every microdata record is subjected to a series of stringent edits. By way of example, the type of offence must be consistent with the presence of a victim, the level of an injury, the presence of a weapon, the type of property stolen, etc. There exist edits to ensure chronological consistency, appropriate ages for persons involved and more. Each month, each respondent receives an edit report detailing the results of these verifications. Every incident is classified as accepted, accepted with a warning, or rejected. Respondents are asked to correct their data and re-submit the revised microdata incidents and most comply with this request to a certain extent. The likelihood of updating is dependent on the size of the police service and, more importantly, the availability of internal resources for this task.

In addition to the edit process, there are a variety of other methods used to ensure data quality. At the incident level, we have developed a program which analyzes the distribution of a number of key variables. For example, in incidents where a weapon was present, we look at the distribution of the types of weapons that were present for each respondent. The typical problem we find with this type of analysis is that police services over-use the “unknown” value. In cases such as this, Statistics Canada analysts follow up with the police service to determine if there is a definitional or systems issue. Some of the other variables that are examined using this process include the aboriginal status of the accused, property stolen, relationship between victim and accused, age of accused, and location of incident. The reports created through this process have been used to correct a number of systematic problems with a variety of respondents that had previously gone unnoticed. These typically involve situations where individual records for a particular police service will pass the edits, but when analyzed as a whole; their distributions are quite different from other police services. Included in this process is a score function which provides a relative measure of the outlyingness of each respondent’s distribution for each variable. This score is used as a means of prioritizing follow-up with respondents.

Data quality is not only examined at the incident level—there are numerous verifications at the aggregate level as well. There is a process at Statistics Canada which converts microdata into its equivalent aggregate format (known as the Derived Aggregate Offence or DAO) to allow for the publishing of crime counts with consistent historical definitions. It also allows us to examine respondents’ data in a uniform and consistent manner going all the way back to 1962. In addition to the respondent verification process mentioned previously, there are two processes, called the Outlier Detection Routine (ODR) and the Tolerance Edit Program (TEP) which are used to pinpoint potential data problems. In the ODR, we examine the ratios of the number of people charged versus the number of incidents that took place. While there is some natural fluctuation, these ratios have proven to be very consistent over time, particularly for larger respondents. The ODR produces a report by offence code that lists the most outlying respondents—again based on a score function. Follow-up can again concentrate on the respondents with the highest scores. The TEP looks at individual respondents’ changes over time, examining the number of offences that took place this year compared to last, by offence code. The results from the two programs are then merged, so that respondents who show up on both are classified as being the highest priority for follow-up.

5. Recent Methodological Developments

The UCR has undergone a number of developments in recent years. Of course, the addition of new variables on hate crime, organized crime, cyber crime and geocoding have had a significant impact on the survey. As mentioned above, we perform an analysis on the distribution of variables, recognizing that passing the edits is not sufficient to ensure quality. We have also recently developed a methodology for analyzing correction rates. By this, we are asking the question, “Are respondents acting on the edit reports?” Early results show that there is considerable variation in the answer to this question, depending on the respondent. In some cases, almost all errors and warnings are fixed within a month or two. In other cases, little effort appears to go into correcting the data. While this process is not yet formalized as part of the regular production stream, it is planned that this will take place in the near future.

On a more analytical note, recent developments have taken place in the area of spatial modeling. Statistics Canada analysts are interested in understanding the
characteristics that are correlated with the presence of crime at the neighbourhood level. This analysis was initially performed using a standard linear regression model for the city of Winnipeg. (Fitzgerald 2004) Further research, however, revealed that there existed a more efficient and accurate method to model this type of data. The reason for this is that crime does not occur randomly in a city; there tend to be concentrations or “hot spots” of crime. As a result, observations are not independent and spatial modeling techniques are preferred to accurately estimate model parameters and test hypotheses. In the presence of this situation, termed “spatial autocorrelation”, standard regression techniques are inefficient and can lead to inaccurate conclusions. Much of this work is summarized in Collins (2007).

In another area of analysis at Statistics Canada, there is increasing interest in recidivism and the paths of people through the justice system. To perform this analysis, record linkage activities take place involving data from policing, courts and corrections areas. Unfortunately, a unique identifier does not exist for this purpose. Instead, police provide a Soundex code which historically has been a 4-byte encryption of the name of an accused. We have developed a quality code which is attached to each soundex based on the commonness of the name in the Canadian population. This was done using data from the provincial electronic telephone directories to simulate the distribution of names across Canada. As a result of this process, each soundex receives a code indicating the relative likelihood of obtaining false positive matches. This work is continuing and we have recently received approval to use Soundex codes obtained from work by the 2006 Census to more accurately simulate the name distribution of the Canadian population.

Finally, we have addressed an issue that has long been outstanding regarding missing data in the UCR. In the past, when a respondent was unable to submit data due to a problem with their RMS, we would “copy forward” the data for the same months from the previous year. There are a variety of reasons for this, including the fact that respondents do not wish to have Statistics Canada “make up” their data and that these problems are often not detected until very late in the production cycle. Recently, however, Statistics Canada has used time series modeling to more accurately impute missing data. An internal study of several Canadian cities showed that long term trends in crime are missed by the “copy-forward” methodology but easily picked up by time series models. As a result, time series imputation is far more efficient and accurate (Collins 2006). Data for some respondents were imputed using this method in 2005 after extensive consultations with the police services implicated. An example is given below for theft in Edmonton (Figure 1). As can be seen, both the copy-forward and time-series methods show similar shapes, but the carry-forward method misses a longer-term upward trend which the time series model easily detected. The result is that for the entire 6-month time period, the time series imputation is much closer to the total than is the copy-forward methodology.

Figure 1 – Thefts in Edmonton, 2003

6. Future Work

As was mentioned earlier, we currently do not have a reliable method to calculate a response rate for the UCR. We are able to provide an accurate measure of coverage at the respondent level. To address this issue, the UCR is represented on a committee that is revising the guidelines for reporting on nonresponse. It is hoped that the revised guidelines will enable us to calculate a meaningful response rate for the survey.

Another project currently underway at Statistics Canada is the development of a crime index as an additional measurement of overall crime in the country. The crime index is analogous to the Consumer Price Index (CPI) which is a tool used to measure price changes. Currently, Statistics Canada uses the crime rate per hundred thousand (number of incidents divided by population multiplied by one-hundred thousand) as the key measurement of crime in the country. One of the drawbacks with this measure is that high-volume, low-seriousness crimes tend to

---

4 For a detailed description of the Soundex code and algorithm, see Armstrong (2000)
5 For more information on an application of this process, see Carrington (2004).
6 For more details on the CPI, see Statistics Canada (1995).
dominate the crime rate and drive the annual trend. The crime index uses a weighted approach which takes into account the relative seriousness of each offence by using average sentence length from courts data as a proxy for seriousness. Development is still ongoing, but early results show that the crime index is driven by more serious crimes (such as robbery and break and enter) and the index is much more resistant to fluctuations in less serious crimes\(^7\) (such as mischief and disturbing the peace).

As UCR 2 coverage approaches 100%, the desire to have a complete microdata file for analysis has increased. Work has started to be able to impute microdata records (using donor imputation) for situations where only aggregate data is available. The end result would be a complete microdata file for analysis that would allow for the examination of trends over an extended period of time. This work serves an additional purpose as well. Even when the UCR 2 reaches 100% coverage, there will always be times when a respondent is unable to supply data due to technical problems. In the case of trend analysis, such a situation would require the removal of this respondent from the analysis. With microdata imputation, we could compensate for these temporary issues and retain as many respondents as possible in the trend analysis. Of course, such an analytical file would not be usable for record linkage purposes, since we could simulate the characteristics of the incident, but not the identity of the accused.

There are also two new processes within the UCR that we are hoping to formalize in the near future. Time series imputation of aggregate data and the production of correction rates are both in development. The ideal situation would be to have both of these functions integrated into the production system. Methodologists will be working with subject matter and technical staff to incorporate these processes in the near future.

7. Summary

The UCR and a typical business survey share many similarities, including frame quality, distribution of the respondent population and the use of administrative data sources. The UCR is unique in a variety of other ways, such as dissemination strategy, content modification and respondent cooperation. It is the institutional nature of the survey which drives many of these differences. The evolution of user needs regarding the content of what data are collected also plays an important role in directing the UCR. In addition to data users requesting new data elements to be collected, the nature of crime also tends to change. Almost every year, new crimes are added to the Canadian Criminal Code, often due to recent technologies (e.g. computer-related) or to address specific policy concerns (e.g. organized crime, terrorism). For respondents, being able to update their systems to accommodate these changes are of high importance, since operationally, they must be capable of charging people with these new crimes when they come into effect. The net effect is that it is incumbent on Statistics Canada to ensure that data quality is being maintained. Collection of data for statistical purposes is still a by-product of the police services' operational requirements and as such will never hold the highest priority for individual police forces.

Acknowledgements

The author would like to acknowledge the hard work of the methodologists with whom he has worked on the UCR over the years, including Zak Pritchard, Fred Hazelton, Brad Fetter, Krista Collins, Karelyn Davis and Asma Alavi. Additionally, none of this work would have been possible without the cooperation of the Canadian Centre for Justice Statistics and past and present UCR Managers, Tim Leonard and Anthony Matarazzo.

References


\(^7\) For more details on the development of the crime index, see Davis (2007).


