Evaluation of the Multijurisdictional Task Forces (MJTFs), Phase II

MJTF PERFORMANCE MONITORING GUIDE

Final

Contract #2005-DD-BX-002

February 27, 2009

Prepared for

Brett Chapman
Office of Research and Evaluation
National Institute of Justice
810 Seventh St., N.W.
Washington, DC 20531

Prepared by

William Rhodes, Project Director
Christina Dyous, Deputy Project Director
Meg Chapman
Michael Shively
Dana Hunt
Kristen Wheeler

Abt Associates Inc.
55 Wheeler St.
Cambridge, MA 02138
www.abtassoc.com

Contents

Introduction

State and local law enforcement is the first line of defense in reducing the availability of illegal drugs on American streets. Recognizing the critical role of local agencies, federal and state agencies augment local enforcement funding with grants and transfers. However, with increased frequency those federal and state agencies are requiring local accountability of the use of funds—hard evidence that grants and transfers are worthwhile. As a result, programs that are unable to provide solid evidence of their activities and effectiveness are more vulnerable to having their resources cut or diverted to programs that can better demonstrate success.

Multijurisdictional Task Forces (MJTFs) are not immune to this new accountability in government, but like many other law enforcement programs, MJTFs lack a history of collecting performance measures and using them to justify program operations. Consequently, MJTFs are disadvantaged in the new competition for federal and state funding. Overcoming that disadvantage requires developing a *performance monitoring system*.

Performance measurement is a system for gathering information about how programs operate and what they accomplish. *Performance monitoring* refers to the periodic analysis and use of the data collected through performance measurement to track program implementation and execution, making changes as necessary based on subjective and objective assessments of the information. *Performance evaluation* is a subset of performance monitoring that requires application of rigorous research protocols to make conclusions about program performance. A *performance monitoring system* is a set of procedures for integrating performance measurement and performance monitoring into a framework that supports public decision making. A performance monitoring system, including its performance measurement and performance monitoring components, is the focus of this guide.

Why Develop a Guide?

Multijurisdictional task forces exist throughout the United States. The National Institute of Justice and the Bureau of Justice Assistance commissioned Abt Associates to develop a methodology for evaluating those task forces based on a demonstration evaluation performed across four states. The goal was to transfer that methodology both to state agencies that had oversight authority and to task force commanders who were interested in program evaluation. NIJ and BJA envisioned a workbook format to facilitate the technology transfer.

After extensive investigation by Abt researchers, NIJ and BJA concluded that developing a workbook was impractical. Not only were data insufficient to estimate what task forces accomplished, data were inadequate to even tell what the task forces did as routine work. The problem was that state and federal grants were awarded without even the rudiments of performance monitoring.[1] The revised

[1] The absence of a performance monitoring system is not inherent to the nature of multijurisdictional task force operations. Although High Intensity Drug Trafficking Area (HIDTA) programs tend to target higher level drug trafficking organization than do MJTFs, the HIDTA and MJTF programs have much in common. Yet the HIDTA has a fully functioning and improving performance management process. See http://www.whitehousedrugpolicy.gov/hidta/perfinfo.html, downloaded on February 18, 2009.

goal was to provide a template that BJA and state agencies might follow to develop performance monitoring systems. This report provides that guide.

A reader might infer that developing a data reporting tool for a monitoring system is a simple exercise that requires a straightforward approach: Ask experts to determine what needs to be known and ask task force commanders to provide the data. Why is a guide necessary? Any researcher who has ever designed a survey knows that this simple approach is a formula for disaster. Certainly experts who understand MJTFs should help frame survey questions, but this is hardly sufficient to develop a performance monitoring system. First, questions need to be posed so that they quantify the salient aspects of program activities and outcomes. Second, the questions need to be posed so that they mean the same thing to all respondents. Third, the questions must be answerable without undue effort by MJTF commanders or their designees. Finally, even the best experts are unlikely to know if the monitoring system works without preliminary and final testing. This guide explains the requisite approach using a step-by-step demonstration.

What Is This Guide?

This guide is designed to assist BJA, state administrative agencies (SAAs) and others in developing a performance monitoring system. This guide provides:

- a generic logic model that can be tailored for each state and MJTF to provide a structure for understanding MJTF funding, activities, outputs, outcomes and impacts; and
- a model data reporting template that forms the basis of the performance measurement component; and
- a model analysis plan that forms the basis of the performance monitoring component.

For Whom Is This Guide Intended?

The primary audiences for this guide are BJA and SAAs, although MJTF commanders and all others with an interest in how to measure how MJTF programs operate and perform may benefit from this guide. Because SAAs are the organizations most likely to have the need, interest, and the resources to monitor or evaluate MJTFs, this guide is specifically designed to assist them in developing performance monitoring systems. Most of the performance data must be gathered from individual MJTFs, so the guide is also designed to assist SAAs in developing systems that are user-friendly for MJTFs.

How Was This Guide Developed?

The template for a performance monitoring system presented in this guide was the result of collaboration between BJA, the National Institute of Justice, and Abt Associates Inc., with the assistance and advice of SAAs from Colorado, Georgia, Illinois, and Tennessee, and in consultation with task force commanders. The objective of this project has been to produce a set of practical tools to help SAAs, BJA, and MJTF commanders measure the implementation of MJTFs and to describe and monitor the work they do. To pursue this objective, the study team examined the range and quality of existing data regarding MJTFs by surveying 56 SAAs and 757 Byrne-supported MJTFs,

conducting site visits to 18 MJTFs in six states, and gathering extensive retrospective data from four states.

Specifically, the study team examined the monitoring and reporting needs of BJA and the SAAs, compared them to the information that was currently available, and used the available data to assess their usefulness for performance measurement, monitoring and evaluation. As reported, investigations showed that existing data are inadequate for these purposes. Most MJTFs and SAAs report only basic information, and examination found that data were often incomplete and/or the quality too poor to support effective program monitoring, comparisons among programs, or the tracking of changes over time.

Why Should Programs Be Monitored?

To Promote Accountability

There was a time when performance measurement, monitoring and evaluation of government-funded law enforcement programs were the exception rather than the rule. However, in recent years federal accountability initiatives and competition for funding has intensified, and the federal and state governments are demanding evidence that programs are implemented as intended and accomplish what they promise. Increasingly, agencies receiving federal funds are required to report how funds are spent and the results of those expenditures. For example, the Government Performance and Results Act of 1993 (or GPRA) put the burden on programs to provide evidence of their implementation and effectiveness. This has been reinforced by the White House's issuance of the Program Assessment Rating Tool (PART) to rate all government agencies by examining their individual program initiatives.

Performance measurement is not simply a reporting burden. Performance measurement and monitoring help SAAs assess the fit between the MJTFs' designs and implementation, and the consistency of implementation over time and across task forces with an eye to improvement. An effective monitoring system allows states to track and compare the operations and performance of:

- Individual task forces.
- Task forces of distinct types, such as
  - those focusing on specific drug types.
  - those covering urban versus rural areas.
  - those with a focus on street gangs versus higher-level, organized crime networks.
- The entire set of MJTFs within a state over time.

Accountability in government is an unavoidable part of today's world. Government agencies must provide evidence that public resources are being deployed as promised and produce benefits commensurate with their costs. The structure provided by the performance monitoring system explicated in this guide will help to expand the capacity of BJA and states to account for (a) the resources deployed among MJTFs, (b) the activities those resources supported, and (c) whether those

activities produced results. The performance monitoring system will also allow BJA to better demonstrate the overall value of the hundreds of MJTFs supported by Byrne/JAG funds.

To Support Decisions about MJTF Sustainment

With tight budgets and many competing demands for law enforcement resources, programs with hard evidence of effectiveness are more likely to be retained, while those without such evidence are more vulnerable to elimination. Performance measures can provide information attesting to the effectiveness of MJTFs, thereby justifying public investments. Conversely, evidence may demonstrate poor performance, which can guide decisions to modify under-performing MJTFs or to replace them with something more promising.

To Inform Program Improvement

Feedback on program implementation and results can also be used to make adjustments that improve task force performance. For example, several task forces could fail to show progress in conviction rates for task force cases, leaving the SAA and separate task force leadership teams to determine how to respond. Perhaps these relatively unsuccessful MJTFs fail to work as closely with prosecutor, while successful MJTFs engage prosecutors as soon as investigations open. Armed with this information, SAAs may begin providing guidance to some of the state's task force on effective strategies for early coordination with local and federal prosecutors.

Using Performance Measures To Monitor MJTFs

Performance measurement systems require that programs report periodically (usually quarterly or semi-annually) in a standard format to continuously update program records. For example, the data reporting template provided in this guide asks SAAs to collect data on any training funded for MJTF investigators. Whenever MJTF staff receive training, the MJTF should update its internal records of who provided the training, the number of officers or other staff trained, how many hours of training each received, any certification on training done, etc. If data input into a performance measurement system are routinely kept current, up-to-date information on training or any other measure of MJTF operations are then readily available for periodic reports requested by program sponsors.

A wide range of questions can be addressed with the information provided by MJTF performance monitoring systems. For example:

- What are the MJTFs' goals and objectives?
  - Are these goals and objectives consistent with the intentions of the agencies providing funding?
  - Are the stated goals and objectives up to date?
  - Have they changed over time in response to new needs or concerns?
- What are MJTFs doing (what activities are being performed)? For example,
  - Investigating.
  - Training task force investigators.
  - Public education and outreach.

  - Holding regular meetings of MJTF personnel.
  - Collaborating with law enforcement in other jurisdictions.
  - Sharing intelligence.
  - Seizing drugs and assets.
- Are MJTF activities well-suited to achieving stated objectives?
- What resources are being mobilized for MJTFs?
  - Are these resources adequate to support MJTF activities?
  - What is the total annual cost of MJTFs?
- Are MJTFs producing their intended results?
  - Have there been increases in drug seizures, arrests, etc.?
  - What programs or activities worked particularly well?
  - Do the benefits justify the costs?

This guide and the template provided in Appendix 2 offer a base upon which BJA and SAAs can model and build MJTFs performance monitoring systems. The development of these systems can be pursued in a three-step process. The following four chapters are each one step in developing and using a performance monitoring system:

- Chapter 1 discusses the first step when developing an ongoing performance monitoring system—deciding what should be measured.
- The second step is to develop an ongoing system for defining measurements and collecting data. This is the subject of Chapter 2.
- As data accumulate, developing a performance monitoring system requires data analysis, and Chapter 3 turns to basic data analysis.
- Some SAAs will seek to go beyond basic data analysis, and Chapter 4 provides some guidance for SAAs that are so inclined, though advanced data analysis may require the assistance of consultants trained in social science statistics.

Chapter 1: Performance Measurement: What to Measure

MJTF goals, structures, resources, and operations vary across states and within states. Nevertheless, diverse MJTFs will have similar long-term goals (reduce drug availability, reduce drug crime); structure (task force commanders, policy board); resources (grant funds, staffing, partnerships); and activities (covert and overt activity, community prevention). Performance measurement should reflect these core elements, capturing unique task force components through supplementary data collection efforts. This chapter describes a process for identifying these core elements, offering a logic model that BJA and SAAs can modify to accommodate special features of MJTFs within the state.

Logic Models

Logic models are commonly used to link a program's goals with its strategies and tactics. Logic models are effective tools for studying any type of program, but are particularly useful when trying to break down a program as complex as MJTFs.

All programs have goals and objectives. To pursue these goals, programs use resources that support activities intended to produce targeted results. Logic models show the logical links from program goals, to resources, to activities, to outputs (the direct representation of activities), to program outcomes (indication of the change that activities are seeking to accomplish), and finally to impacts (the indications that the program's broader goals have been realized).

Graphically:

[1] goals → [2] resources → [3] activities → [4] outputs → [5] outcomes → [6] impacts

Each component of the logic model is described in more detail below:

1. Goals

Goals reflect what the program is trying to accomplish. Every MJTF has goals. These might be stated broadly: Reduce the availability of illegal drugs. They may be more targeted: Reduce the availability of methamphetamine. Without identifying program goals, it is not possible to determine whether the MJTF is structured to achieve those stated goals. For example, does it have the required resources to match the plans? Does it engage in other related activities?

2. Resources/Inputs

Resources include all funding, space, equipment, personnel, in-kind contributions, etc. that support program activities. For task forces this might include funding from federal and other sources; state and local partners; forfeiture proceeds; access to intelligence support systems; in-kind contribution of equipment and space; and task force and on-loan staff from partner agencies.

3. Activities

This is what the program "does" on a regular basis, including all activity directed at meeting program objectives. For MJTFs, these activities may include launching investigations, training agents, conducting community prevention and public education programs, sharing intelligence, and coordinating investigations with prosecution.

4. Outputs

Outputs are the direct or immediate results of program activities, i.e. numbers of people trained, numbers of seizures made. The outputs are representations of activities, or measures of the implementation of the program activity. The outputs of operational activities should link directly to the specific activity and "make countable" what occurred. For example, if the activity is drug eradication, outputs could be counts of marijuana plants or hectares destroyed, or drug labs eradicated. If the activity is undercover operations, outputs would include the number of active or new investigations.

5. Outcomes

The outcome of each activity is what is expected to happen when the activity has been fully implemented. Rather than simply a count of the number of officers trained, for example, the outcome of the training would be an indication of increased skill level or knowledge gained. Similarly, an outcome of mobilizing new prosecutorial teams might logically result in increased numbers of prosecutions of drug cases. Outcomes of operational activities reflect the result of putting into place those activities.

6. Impact

The impacts mirror the goals. The impacts are what one would reasonably expect from the successful execution of strategy and tactics. For MJTFs impacts may include reduced drug availability, drug use, and drug-related crime. An impact is typically a distant result of program activities, acting through a causal linkage: activities → output → outcome → impact. Successful operation of a MJTF is only one of many factors that affect goal achievement. However, the presumption is that the MJTF can have some meaningful effect on local drugs markets (or whatever else it set as a goal), or else there would seem to be little basis for spending public funds on task force activities.[2]

To summarize, there should be a logical connection across each component of the logic model, and the logic model should be a comprehensive statement of how and why an MJTF works. Understanding this logic is a prerequisite for developing a performance monitoring system, because why would one want to monitor something that was unimportant to program performance, and how

[2] Task force commanders and enforcement personnel in general avoid being held accountable for impacts such as the prevalence of drug abuse. From a social science perspective, their argument is that so many diverse factors affect drug abuse that it is difficult to distinguish the effect of task force activity from the effect of those other factors. Social scientists think of this as the problem of distinguishing signal from noise. The point here is that impacts are part of the logic model even if statistical analysis can only imprecisely estimate the links across elements of that logic model.

could one justify failing to monitor an essential component of the logical structure of an MJTF program? Of course, it may be impractical or impossible to quantify parts of the logic model, but that is a problem discussed later in this guide.

Developing a Logic Model

Working with the MJTFs, BJA or the SAAs should develop a logic model reflective of every MJTF under its jurisdiction. The purpose of this exercise is to identify MJTF core elements to ensure that the performance monitoring system is tracking measures appropriate for assessing the performance of all BJA-funded task forces in the state.

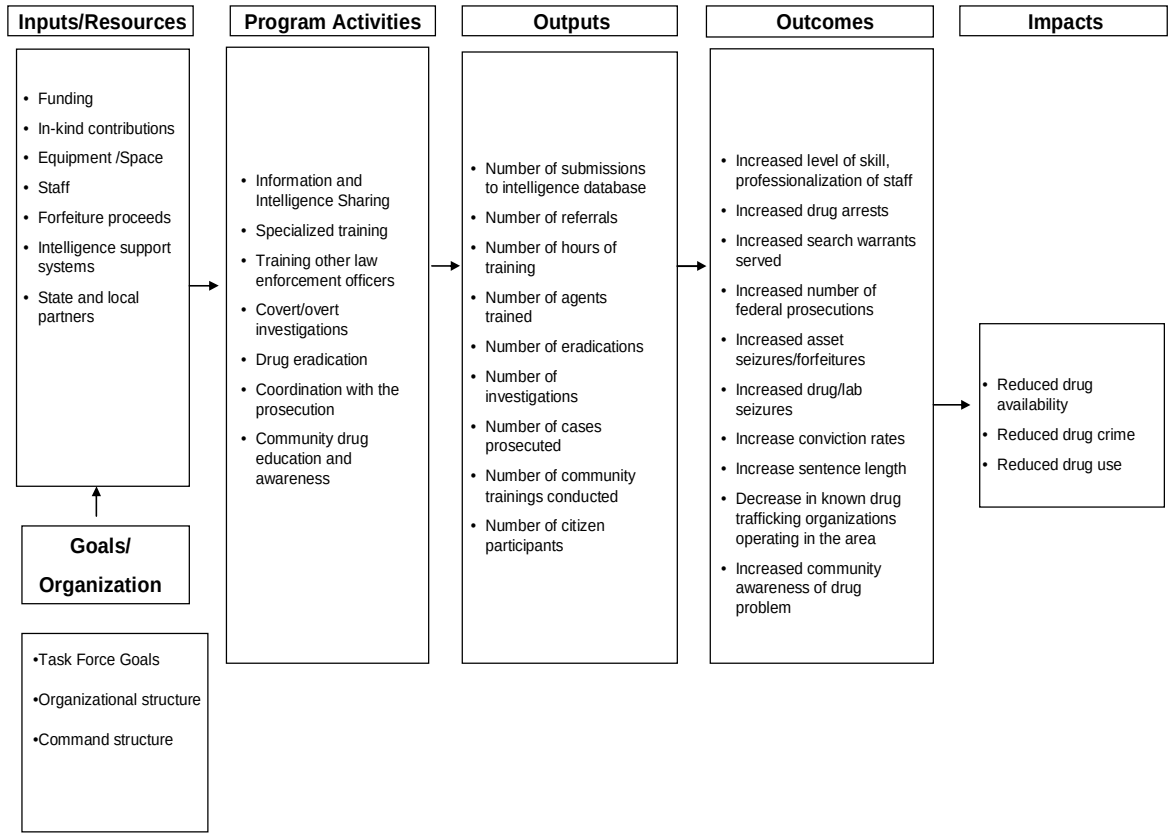
To assist this process, the study team developed a generic logic model as a prototype. The prototype was developed in coordination with NIJ and BJA after working with SAAs in four states and reviewing the goals, structure, resources and inputs for task forces in each state. The figure below displays this model.

The study team recommends starting the process of building each logic model with the SAA program staff. The assumption is that SAA staff has been working with MJTFs in the state for a number of years, reviewing grant applications and monitoring program implementation, and should have some knowledge about how these task forces operate. Assembling this group to review the prototype logic model and to discuss how the logic model differs across task forces in the state is a useful way to begin tailoring the logic model. It will also help facilitate discussions during the next step, which should be to engage MJTF commanders in the process. Soliciting input from MJTF commanders is best done in a group setting with a representative from the SAA or an external party facilitating the conversation. If this is not possible, feedback may be solicited individually from MJTF commanders, though this may be less efficient when opinions differ. The importance of this second step is to ensure the logic model has been validated by the field and is comprehensive of all the task forces in the state.

Tailoring the logic model to the state's MJTFs requires that the SAA and task force commanders discuss each component of the logic model. The following includes some guidance for these discussions.

Figure 1.1

Logic Model Prototype for Multijurisdictional Drug Task Forces



Defining Goals

What is the mission? Answering this question should focus on identifying both the long- and short-term program goals. Note that outputs and outcomes are not goals, but rather, producing outputs and outcomes are instrumental to achieving goals. Some questions that might help to identify goals are:

- What are the task forces in the state trying to accomplish?
  - Reduce drug availability?
  - Reduce drug use?
  - Reduce drug related crime?
- Are there other things the task forces in the state are trying to accomplish?
  - What goals are included in the task force mission statement?
  - How do task forces in the state describe their purpose?

Threat assessments are useful when setting goals. Given a threat assessment, the goals might include targeting specific types of drugs, distribution networks, or drug-related crime. While there should be a consensus about goals for task force operations, every task force is likely to target a subset of those goals as consistent with its threat assessment. At the end of this process, there should be consensus on task force goals, even though the identified goals will likely differ across task forces.

For example, the Abt investigators found that in rural areas, MJTFs often become the principal law enforcement agency for combating drug sales. In those settings, goals are likely oriented toward street sales and related behaviors. In urban areas, MJTFs are more likely to augment the resources of local enforcement agencies. In those settings, local police focus on street sales; the MJTF focus on distribution networks. Goals would be set accordingly.

Identifying All Inputs/Resources

The logic model should reflect all resources directed at the drug problem by MJTFs. Therefore, it should be inclusive of all resources, not just financial. Below are some questions that may help drive this discussion.

- What kinds of resources are listed in task force grant applications? For example,
  - Personnel
  - Equipment
  - Technology
  - Office lease
  - Travel
  - Confidential funds
  - Supplies
- What other kinds of resources are used to support task force activities? For example,
  - In-kind contributions of personnel or equipment
  - Forfeiture proceeds
  - Task force revenue (e.g., proceeds from fines levied on seized drugs)
  - Intelligence systems
  - State and local partnerships
- Consider each of the activities engaged in by task forces in the state and identify any other resources used to engage in those activities.

After identifying what is believed to be a comprehensive list, it is helpful to revisit the list to confirm consistency between task force activities (see the next subsection) and task force resources.

Obviously, when evaluating program operations, the operational cost is a fundamental to a benefit-cost calculation, and without at least a crude benefit-cost calculation, one cannot assess whether a program is worthwhile. Without knowing the inputs into a program, one cannot determine what accounts for failure or success; one cannot make a reasoned determination about how to improve program operations.

Yet the Abt researchers found that BJA and SSAs had no way to determine program inputs. Grant applications identified how Byrne grant funds were to be spent, and of course audits could determine the integrity of that process. But there was no way to reliably determine the inputs into that process because many, if not most, of those inputs were in-kind contributions and some stemmed from undocumented revenue sources including seizures and fines. The logic model should identify all inputs into MJTF operations because inputs must be a principal component of a performance monitoring system.

Listing Program Activities

This discussion should focus on generating a comprehensive list of what task forces do. Because of the variation across task forces, we have included some guidance below to ensure activities are specified sufficiently for the next steps of the logic model.

- Identify all of the activities engaged in by task forces in your state.
  - You may use the following three categories to organize feedback:
    - Enforcement activities, which might include:
      - » Multi-agency investigations
      - » Interdiction activities
      - » Working with prosecutors
  - Intelligence and information sharing activities, which might include:
    - » Intelligence analysis
    - » Information collection and sharing in regional database
    - » Deconfliction
  - Training and education, which may include:
    - » Specialized training of agents
    - » Community education and awareness
- Establish a definition for each activity
  - For example, does intelligence and information sharing refer to hosting regional workshops to discuss the local drug problem, accessing and applying information in regional intelligence system to local threat assessments, or submitting intelligence/operations into a shared system to prevent conflicting investigations and operations?
  - If a single activity has multiple definitions (e.g., types of investigations), consideration should be given to partitioning that activity into more than one activity (e.g., money laundering, pharmaceutical fraud, street-level drug dealing).

The goal is to generate a comprehensive list, even if some task forces fail to engage in some activities. Definitions of activities should be sufficiently refined to capture what an MJTF does to achieve its goals.

It may be evident that a performance monitoring system should identify MJTF activities, but the Abt investigators were unable to identify the specifics of MJTF activities, and in fact, sometimes the

investigators could not determine if an individual MJTF did anything. This is not a criticism of MJTFs. It is a simple assertion that neither BJA nor SAAs can readily determine how Byrne grant funds affect law enforcement operations with sufficient detail to determine if the expenditure of public funds is appropriately targeted.

Identifying and Developing Output Measures

The discussion should address how activities can be quantified. Below are some questions to facilitate the discussion.

- How can each activity be measured?
  - What would indicate whether or not an activity is occurring?
    - For example, how can one measure information sharing? Through the number of times a database is accessed or the number of multi-agency meetings? How "information sharing" was defined in the earlier exercise will help make this distinction.
  - What does that measure reflect?
    - For example, enforcement activities. Should the measure reflect the quantity through a measure of the number of active investigations or the scope through a measure of the type of organization targeted?

At the end of the discussion, a final check should be made to ensure every activity has at least one output measure.

Working with task forces, the Abt researchers understand the difficulty of quantification. For example, intelligence centers engage in a variety of activities. At one extreme, they perform comparatively simple functions: They track individuals and groups who are under investigation, thereby informing two or more agents that they are unwittingly working on the same case, and they may track operations, informing agents when two operations may conflict and endanger participating agents. At another extreme, they may perform wire taps, an extremely time intensive operation. A simple count will not work to untangle these diverse tasks and to quantify operations. Perhaps more useful would be a count of the number of wiretap operations, a count of the number of case file preparations (organizational charts and time lines), and a count of the number of trap and trace devices and pen registers.

Identifying and Developing Outcome Measures

This discussion should focus on what success looks like, which may mean different things to different task forces. Below are some questions to help guide this potentially difficult discussion.

- What is success for each activity?
  - For example, increased number of arrests may be considered success by some MJTF commanders who focus on lower level drug distribution, but not by task forces focusing on higher level drug trafficking organizations where cases take longer to build and fewer people are arrested.

- How do you measure that success?
  - In the above example, both task forces would probably be satisfied if the measures included both increased number of drug arrests, increased number of federal prosecutions, and increased number of dismantled drug trafficking organizations. The point is that no single outcome tells the story for how task forces are performing in the state.

At the close of the discussion, each activity listed in the logic model should have at least one outcome.

Quantification is difficult. Consider the definition of a drug trafficking organization (DTO). While the US Sentencing Commission does not identify a DTO, it identifies roles played in a DTO. For example, the guidelines recognize a role:

> If the defendant was an organizer or leader of a criminal activity that involved five or more participants or was otherwise extensive…

Based on this role, an outcome measure might define a DTO as a drug trafficking organization that had five or more participants and had an organizer or leader. This definition would greatly improve how MJTFs count their outcomes, but the Abt researchers' experience is that this improvement is still insufficient. Over time agents develop investigations – what they identify as cases – and large scale investigations can involve more than one DTO who buy and sell drugs or otherwise act cooperatively but as separate organization. Without clear instructions, the agent will identify all parties in a large scale investigation as comprising a DTO. When that happens, the number of DTOs will be understated. However, agents who work street investigations may take an expansive view of this definition, reasoning that a street-level drug dealer must buy from at least one source and must sell to at least three buyers and, hence, the street-level dealer must represent a DTO. When that happens, the number of DTOs will be overstated. When defining and quantifying the term DTO, care is required to avoid assuming that a definition will be applied consistently, else statistics generated by the performance measurement component will be misleading and potentially uninterruptable.

Thinking Through Impacts

The discussion should focus on long-term impacts of activities that are consistent with program outcomes and goals. For many states, the generic logic model will require only minor modifications. Some discussion questions are included below to help confirm this.

- What are the long-term impacts of task force activities?
  - For example, what would be the outcome if enforcement activities were successful over a long period of time? Reduced drug-related crime? Reduced drug availability?
- Is there a logical relationship between the impacts and task force goals?
  - For example, if the primary task force goal is to target major transportation routes for illegal drugs, then reducing local marijuana and methamphetamine production would be excluded from the logic model.

- Is there a logical relationship between impacts and outcomes?
  - For example, if all activities are focused on enforcement and outcomes are focused on the prosecution of street-level drug trafficking organizations, then one might expect to have a significant impact on drug availability in a region, but not necessarily money laundering.

During this discussion, the SAA should be mindful that MJTF commanders will have the most concern with the measurement of impacts, principally because MJTF activity is only one of many causes of drug availability and drug use. No one likes to be held accountable for what they cannot control, and during the study team's investigations, it was clear that MJTF commanders objected to being held accountable for impacts. To address this concern, remind MJTF commanders that it may be impractical to measure all the linkages in the logic model, but measurement issues should be distinguished from the logic of task force objectives, strategy and tactics.

Moreover, the problem of estimating the linkage between impacts and earlier components of the logic model is common in social science enquiry and surely not unique to monitoring and evaluating task forces. For example, early childhood interventions are often done to ultimately improve the outcomes from secondary education and to improve earning prospects. Few social science studies would actually track program participants through the decades required to quantify the relationship between early interventions and long-term impacts. Instead, one relies on a logic model: early interventions lead to higher grade school performance; improved grade school performance leads to improved high school performance; better high school performance leads to higher earning prospects. If the logic model is compelling, then demonstrating that the early intervention improves grade school performance argues for the utility of early interventions. The logic model must be compelling, however, a point that emphasizes the importance of spending time developing the logic model even if some of its linkages cannot be estimated with precision.

Summary: Developing Logic Models

After this process is complete, the resulting logic model should reflect a shared understanding of the core elements for logic models operating in the state. As mentioned previously, task forces differ from each other and a useful logic model must reflect that diversity. For example, rural and urban jurisdictions operate differently and, therefore, may have quite different outcomes. Consequently, it is unlikely that a single logic model will fit both, but the generic model can be tailored to represent each.

This two-step process of engaging SAA program staff and then MJTF commanders in a review of a logic model should only be required initially. After that, SAAs can ensure the logic model(s) remain current in a number of ways. For example, grant applications may be organized around the logic model so the SAA can easily crosswalk the information in the applications to the current logic model. States with annual meetings of MJTFs might also include a review of the logic model on its meeting agenda.

Using the Logic Model to Move to a Performance Monitoring System

A logic model provides the framework for ensuring that the monitoring system accounts for how a task force carries out its work. Ideally, the system should capture this information from multiple sources, e.g., task forces, SAAs, other state agencies. The best source is the one that can provide the information most efficiently.

A performance monitoring system should be both reasonable and not overly burdensome. It should measure information generated by or otherwise accessible to task forces. In other words, if information can be obtained from other reliable sources (e.g., state comptroller) it places unnecessary burden on the task forces to also track that information. Collection of these data should be a state responsibility. This will be clear for data that are already being collected at the state or local level, but less obvious for information not previously collected, which may require adding to extant data collection efforts or generating new data collection efforts.

The next chapter reviews a data collection template that has been created based on the prototype logic model. This tool will facilitate the process of transitioning from logic models to performance monitoring systems.

Chapter 2: Performance Measurement: Collecting Data

As discussed in the previous chapter, the logic model identifies what should be measured to reflect MJTF performance:

- Goals
- Inputs/resources
- Program activities
- Outputs
- Outcomes
- Impacts

Often these data elements are either unavailable in existing sources such as program documentation or are measured inconsistently so they are not useful for explaining program operations and accomplishments. This chapter describes a template for collecting measurement data. While this template should provide a good starting point for most states, it may need modifications to accommodate special features of MJTFs within the state, as well as adjustments to address issues of special policy relevant to individual states. This chapter walks readers through the process of reviewing and tailoring the template.

The process remains valuable even in states with existing performance measurement systems. In these cases, existing measures can be included in the process of reviewing and tailoring the template to ensure everything that should be measured is being measured. In other words, should any shortcomings exist between the existing system and what needs to be measured, the template can provide measures that will improve the quality of existing systems.

Building from the logic model presented in the previous chapter, the study team worked with BJA and NIJ to develop a prospective, generic data collection form, a field-tested template for gathering data about performance monitoring. The template is prospective in the sense that it is intended for future use to measure MJTF operations. Consistent with the discussion in the previous chapter, the tool is generic in the sense that a SAA may modify the prototypical logic model and those modifications will lead to changes in the generic data collection template. We present the entire template in Appendix 2, and will discuss each item separately in this chapter.

To develop the data collection template, the study team began with the prototype logic model presented in Chapter 2. Recall that the logic model anticipated data from two sources: MJTFs and other sources accessible by the SAA. Using the prototypical logic model as a guide, the study team drafted a tool to collect core data from MJTFs. The template was refined through a focus group with task forces in one state and a field test with task forces in four other states. Those MJTFs provided feedback that allowed the study team to clarify questions that were ambiguous and to eliminate questions that participating MJTFs could not readily answer. In other words, the tool has been validated by the field, which is worth remembering when tailoring the tool to fit local needs.

When tailoring or otherwise fielding the BJA template, the SAA should consult with MJTF commanders. The SAA and MJTF commanders should review the data collection template and adjust it as necessary to deal with departures from the prototypical logic model(s) developed for task forces in the state. Below are some suggestions for how to start discussions within the state regarding how to collect performance data from the MJTFs:

- How will the items in this data collection template supplement or replace items in an existing performance monitoring systems?
- How will the information be collected? For example, through an on-line system or will task forces submit periodic reports?
- How will the information be maintained? For example, will data be entered into a database after it is submitted to the SAA?
- How often will this information be collected?
- How will the information be used?

The next step is to review each section of the data collection template, crosswalking its data elements back to the logic model. Below are some suggested questions to address during the review:

- Is this information relevant for the state's unique logic model?
- Is the question clear?
- Is the information available from another source? If not, what is the burden associated with collecting this information?
- Is other information needed to monitor MJTF performance?
- Should an additional subset of questions be added for certain MJTFs?

The result should be a new or revised data collection template. Revisions should not be done in the abstract by experts who decide that some parts of the template merit change, because such a simple revision process risks producing a meaningless template. Structured focus groups are useful for judging that the instrument provides sufficient definitions that every task force commander or his or her delegate will interpret questions consistently and will be able to provide valid and reliable answers. This is difficult to judge in the abstract. Furthermore, focus groups can provide insight, but there is no substitute for field testing. There is a science to instrumentation, and bypassing that science by taking shortcuts can lead to worthless measures and a dysfunctional performance monitoring system.

To assist in this exercise, the rest of the chapter will review each section of the data collection template, demonstrating how a user might tailor each section to match the state's generic logic model.

A Template for a Multijurisdictional Task Force Performance Measurement System

This template has two parts. Part I, comprises sections A, B, and C and covers the first two components of the logic model—goals/ organization and inputs/resources. The information in this section tends to be static, requiring periodic updates by MJTFs, as necessary. Part II comprises sections D and E and covers outputs and outcomes.[3] These sections cover information that is dynamic and should, therefore be collected on a regular basis (e.g., quarterly, semi-annually).

As each section is reviewed below, the discussion defines terms and provides guidance about how BJA and SAAs might tailor the tool. Tips on how MJTFs should respond to specific items are also provided.

Definitions

The first page of the template provides definitions. Read these definitions before reviewing the tool, as they will clarify questions that otherwise may be ambiguous. Note and correct any inconsistencies with local definitions, for example, definitions of what constitutes a gang or drug trafficking organizations, and revise as necessary, and always document changes. While working through the rest of the template, refer to these definitions, modifying them as necessary.

Definitions may also be added for existing or new items in the data collection tool. As necessary, provide additional definitions to increase the clarity of the tool and to ensure consistent data collection across task forces. Keep definitions consistent over time, because a consistently maintained time-series will increase the value of these data for a performance monitoring system.

Part I

Before reviewing the first three sections of the template, users should review the elements included in the following two components of the state's logic model: ***Organization/Mission*** and ***Inputs/Resources***.

Section A: Task Force Characteristics
This section gathers information on basic characteristics of each task force. Those characteristics are primarily, but not entirely, administrative:

- task force name;
- jurisdiction;
- formal partners;
- staffing; and
- task force objectives.

[3] Impacts were not covered in the template, as this is not information that would be generated or provided by a task force

The reporting cycle is also noted, suggesting that this information be reported initially and updated as needed, but at a minimum annually.

A user might deem some of this information as unnecessary, given existing record keeping systems at the state level, or a user might decide to add a request for information on, for example, informal partners (those who aren't included in the MOU) or the geographic size of the task force jurisdiction.

Field testing of the data collection template suggests being specific about objectives (short-term goals). Although it may seem unnecessary, given that most BJA-supported MJTFs share the broad goal of reducing drug availability and drug-related crime, modifications should capture changes in specific task force objectives (such as focusing on investigating trafficking or manufacturing one drug type versus another), as these are likely to vary across MJTFs and over time as circumstances change in a given jurisdiction. Current information on more specific objectives is also critical to any examination of changes in task force performance over time, as outputs and outcomes should reflect these shifts in priorities. Therefore, review stated objectives for accuracy in terms of whether they are comprehensive of the objectives of task forces in the state, as stated in the tailored logic model.

Section B: Task Force Annual Budget

This section requests information on the task force's annual budget. It includes a breakdown of personnel and non-personnel costs by federal funds; state, local, and/or tribal funds; and task force revenue. It also requests information on in-kind contributions.

This section is designed to capture estimates of task force budgets for the upcoming fiscal year. Although one would anticipate accurate reports of federal, state, local, and/or tribal funds, estimating revenue might be difficult for some task forces. The assumption is that the task forces responding to this request are sufficiently established that they can anticipate what might be expected from other sources (e.g., forfeiture proceeds, contributions from other agencies). Of course, this assumption may be false for some states.

A user might consider an alternate approach, for example, collecting budget and revenue information separately or modifying the request to reflect annual expenditures. When talking to focus group participants, the study team determined that while budget information would be easier for task forces to provide, expenditures are a more accurate reflection of actual resources invested. Deciding which to collect should be based on where information resides, whether it be at the state, task force, or agency level, as this impacts the burden associated with collection. Decisions on which to collect should reflect the reporting cycle, since budget information is better collected at the beginning of the fiscal year and expenditures at the end of the fiscal year.

This section also requests information on in-kind contributions by category. Depending on the extent to which task forces in each state rely on in-kind contributions, users might consider supplementing the question by asking for specification on what is being received, particularly with respect to personnel contributions. Users might also consider assisting task forces with estimating the value of these contributions (use of a cruiser on weekends, assignment of an officer part-time to the task force). Recording non-cash contributions are an important part of understanding the full set of resources at a task force disposal, both in light of fluctuations in activity over time and when comparing outputs across task forces.

Section C: Task Force Structure and Operations

This section collects information on a number of elements of a task force's structure and operations. The items include:

- MOU or Interagency Agreements
- Policy Boards
- Written protocols, policies, procedures
- Training resources
- Physical space
- Collaboration with prosecutors
- Access to intelligence support systems (state-level, regional, national)
- Use of threat assessments to set priorities
- Deconfliction of events

Some users may view one or more of these items as irrelevant for task force operations in their states and more detailed than dictated by the logic model. However, most of these items are elements BJA considers important to successful task forces, and BJA intends to target technical assistance to task forces weak in any of these areas. Therefore, we recommend that users retain the above items (unless inapplicable to all task forces in your state), and we encourage reviewing and supplementing the section with additional items that users deem appropriate.

Part II

Before reviewing the next two sections of the template, users should note the elements listed under task force Activities, Outputs and Outcomes in the state's logic model.

Section D: Task Force Activities

This section is focused on capturing the output from the following task force activities:

- Information and intelligence sharing;
- Education and awareness;
- Investigation activity;
- Drug seizures;
- Firearm seizures; and
- Property seizures and forfeitures.

Outputs are measures that count the activities of the task force, such as the number of submissions to an intelligence database, number of agents trained, and number of new investigations opened.

After reviewing its logic model, an SAA may identify additional activities that require measurement, or the SSA may disagree on how an activity is quantified, requiring some modifications. For example, an SAA might feel that the number of queries to an intelligence database is a better indicator of use than is the number of submissions.

This section also includes information on seizures and forfeitures. Review this section. For example, it may not make sense to collect information on the number of hazardous sites cleaned if another agency in the state is responsible for hazardous waste cleanup and cleanup is, therefore, not an activity occupying task forces in that state.

Ensuring that this section accurately reflects activities performed by task forces in the state, as depicted in the logic model, is probably one of the most important steps in developing the performance monitoring system. It might also be controversial, as MJTF commanders may disagree on how to quantify activities or they may have concerns that data may be misconstrued. MJTF commanders may also disagree on relevance because task forces engage in different activities. Again, the information collected should be comprehensive of all task forces, even if that means some task forces will report "not applicable" for some items. Similarly, the information collected should be as detailed as possible, within reason. In other words, avoid collapsing information into broad categories in an effort to reduce burden. For example, tracking arrest data by severity of charges and charge type might seem burdensome, but detail will be useful in the future and the information can be collapsed later to simplify reporting. The key is to be thoughtful now because it will be difficult to collect information retrospectively.

Information should be reported periodically (e.g., monthly, quarterly, annually). The decision on how frequently this information will be collected will depend on how much the relative value of frequent reporting offsets the burden placed on the task force for collecting and reporting the information.

Section E: Task Force Outcomes

This section collects information on the outcomes or results of MJTF activities. Information on whether individuals arrested by task forces are prosecuted, convicted, and sentenced is the focus of the data collected in this section. This information is an important indicator of the quality of the cases generated by the task force and the level of cooperation with prosecution.

However, as field testing confirmed, there are some difficulties in tracking this information. The first difficulty is tagging cases as task forces cases as they move though the criminal justice system. If this is not already being done, some tagging mechanism would need to be established before collecting this information. The second difficulty is the time it takes for a case to move through the criminal justice system and what happens to the case during that time. Task forces often lose track of a case once it is accepted for prosecution, because prosecutors combine and break-up cases that may go to trial or be pled. Therefore, the case the task force built may look very different a year or two later.

For these reasons, consider the best source for this information. If you decide to obtain the information from the District Attorneys or through court records, then time must be invested in determining how task forces' cases will be identified and tracked. You must also decide how often to collect this information. The suggestion is that the information in this section be reported periodically. If official sources are used, a longer reporting period (e.g., annually) may be desirable to increase the likelihood that a case has progressed through multiple points in the criminal justice system during the same reporting cycle.

If the task forces can provide some or all of the information, then you should provide specific instructions for reporting cases as they move through the criminal justice system (for example, report

on the number of convictions that took place during the reporting period, regardless of the number of cases accepted for prosecution during that same reporting period). Consider the burden of this decision if the task forces were not already collecting this information; avoid undue burden to encourage compliance.

Making decisions on how to collect and track information will be difficult. However, without this information, it would be impractical to measure task force outcomes and, therefore, assess task force performance over time.

After each section has been reviewed, making revisions and additions as needed, users should go back and review the tailored logic model, the definitions page, as well as the overall burden of the data collection on local task forces and whether the burden associated with collecting performance data has been minimized without compromising the purpose of the performance monitoring system. Some of this will depend on the existing performance monitoring system and how making these improvements will change that system. For example, task forces in a state that had an on-line performance monitoring system that included most of the measures in their final data collection tool will unlikely feel an increased burden with the addition of a few items. However, a state where task forces are unused to submitting anything beyond a quarterly or annual report on task force outputs may experience resistance from task forces. This is another reason for engaging MJTF commanders early in the development process.

The next two chapters will demonstrate how performance data can be used to monitor task force performance.

Chapter 3: Performance Monitoring: Analyzing MJTF Performance

This chapter discusses how to use data gathered by a performance measurement system to answer questions about the operations and effectiveness of MJTF programs:

- What are the specific ***goals*** and ***objectives*** of MJTFs?
- What ***resources*** are being mobilized to pursue these objectives?
- What ***activities*** are engaged in by task force personnel?
- Are the program's resources and activities combining to produce the intended ***results***?

Providing answers to these questions requires understanding the elements of MJTF programs, collecting data about program operations, and analyzing those data to draw inferences about program operations and achievements. Chapter 1 introduced a generic logic model and Chapter 2 introduced a template for collecting program data consistent with that logic model.

This chapter explores how to use the state's logic model and accumulated data to monitor MJTF program performance. The discussion is framed around the logic model presented earlier and describes how it can guide performance monitoring, illustrating the linkages between program goals, resources, activities, and outcomes, using hypothetical examples to illustrate key points.

Using MJTF Logic Models to Guide Assessments of Performance

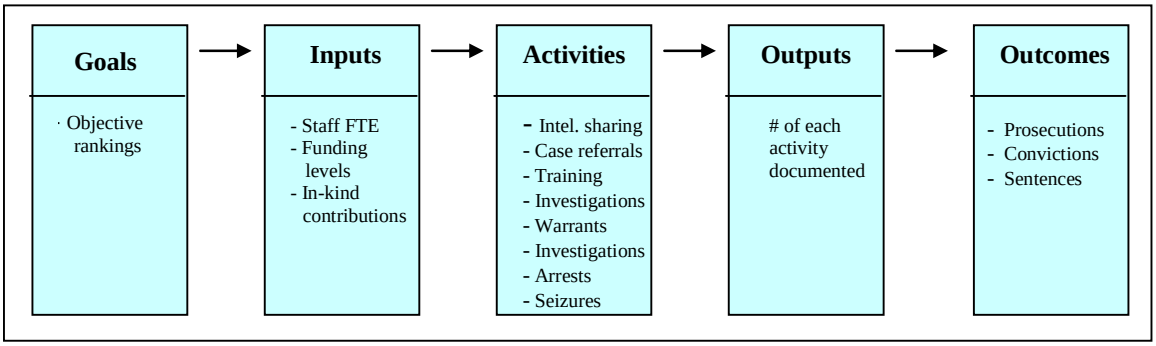
A well conceived and executed performance measurement system provides evidence of connections between a program's goals, the resources that are mobilized and the activities supported by those resources, and the results of those activities. The job of those who monitor programs is to use these data to quantify or at least describe the connections between the model's components. Recall the earlier discussion: At present, one cannot identify task force resource; nor can one tell what a task force does as its routine activities. Even a rudimentary quantification would advance an understanding of MJTFs.

The template has identified a number of performance measures for each component of generic MJTF logic models. Figure 3.1 summarizes some measures that are associated with each logic model component and the linkages between them.

A performance monitoring system should describe MJTF goals, inputs, activities, and outputs using simple tabulations, descriptive statistics, and graphics. What follows is a discussion and illustration of some basic methods for compiling and presenting task force performance data. Subsequent sections of this chapter discuss tracking measures over time and examining the relationships between them to make inferences about linkages in MJTF logic models.

Figure 3.1

Summary of Measures of MJTF Logic Model Components



I. Using the System to Describe MJTF Programs

The section provides examples of how information maintained in a performance monitoring system might be used to describe MJTF programs in the state.

Describing Resources Committed to MJTF Programs

A performance measurement system should document MJTF resources. Most questions about programs, including task forces, are tied in some way to level of effort. For example, in assessing whether the programs have produced tangible public safety benefits, a monitor might ask whether the results are proportional to the level of investment. Answering many questions about MJTF programs requires understanding how the task forces are structured, how many sworn law enforcement staff they deploy, and their total expenditures.

The study team's preliminary research found that many MJTF programs were unable to provide data about program resources. Much of the data routinely compiled by SAAs and MJTFs addressed only the resources associated with grant funding, so compilations often omitted local and state agency staff and in-kind contributions. In many cases, Byrne/JAG grants provided a relatively small portion of all MJTF inputs (such as funding two positions on a task force with 12 full-time sworn staff), so available data were a meaningless summation and dramatically understated program resources.

The performance measurement template described in Chapter 2 requires reporting on the total full-time equivalent (FTE) staff (e.g., task force agents, intelligence analysts) committed to each task force. Staffing levels can be presented and analyzed in a number of ways, and the discussion will return to them when talking about tracking performance measures over time and analyzing linkages in the logic model. For descriptive purposes, one simple way to report on staffing is to summarize the number of full-time equivalents to each MJTF. The example in Table 3.1 presents total FTEs for each task force (although staffing can also be broken down according to level of government), the number of local, state, and federal agencies that formally participate in each MJTF (i.e., those named in MOUs), and the MJTF's total operating budget.

Table 3.1

Summary of Key Resources Committed to MJTFs

Task Force	Number of Participating Agencies			Total FTEs	Total Operating Budget
	Local	State	Federal		
MJTF 1	3	1	0	8	$ 227,000
MJTF 2	12	1	0	19	1,440,000
MJTF 3	5	1	1	7	702,000
MJTF 4	8	2	0	12	1,122,000
MJTF 5	16	3	1	28	2,150,000
MJTF 6	10	1	0	12	958,000
MJTF 7	7	1	0	6	495,000
MJTF 8	9	1	0	10	729,000
MJTF 9	14	1	0	22	1,950,000
MJTF 10	4	1	1	13	1,250,000
MJTF 11	8	2	1	7	644,000
MJTF 12	3	1	0	9	875,000
MJTF Totals	99	16	4	153	$12,542,000
Averages Per MJTF	8.25	1.33	.33	12.75	$1,045,166

Summarizing the budget, number of FTEs, and agencies participating is a starting point for illustrating the size and scope of MJTF programs within a state, and points to areas for subsequent analysis. For example, the basic information presented in Table 3.1 for one reporting period can be presented graphically, such as with bar charts. Figure 3.2 is a bar chart presenting information on in-kind contributions, which the template measures by means of a checklist. This figure displays the percentage of MJTFs that report receiving each kind of contribution. For some analyses and reports, SAAs may find it useful to present the information by MJTF or partition the results according to region, urban versus rural areas, task force focus, or some other basis of comparison.

The performance measurement template also supports describing task force structure and coverage areas. For example, SAAs can describe their MJTF programs using:

Agency Participation

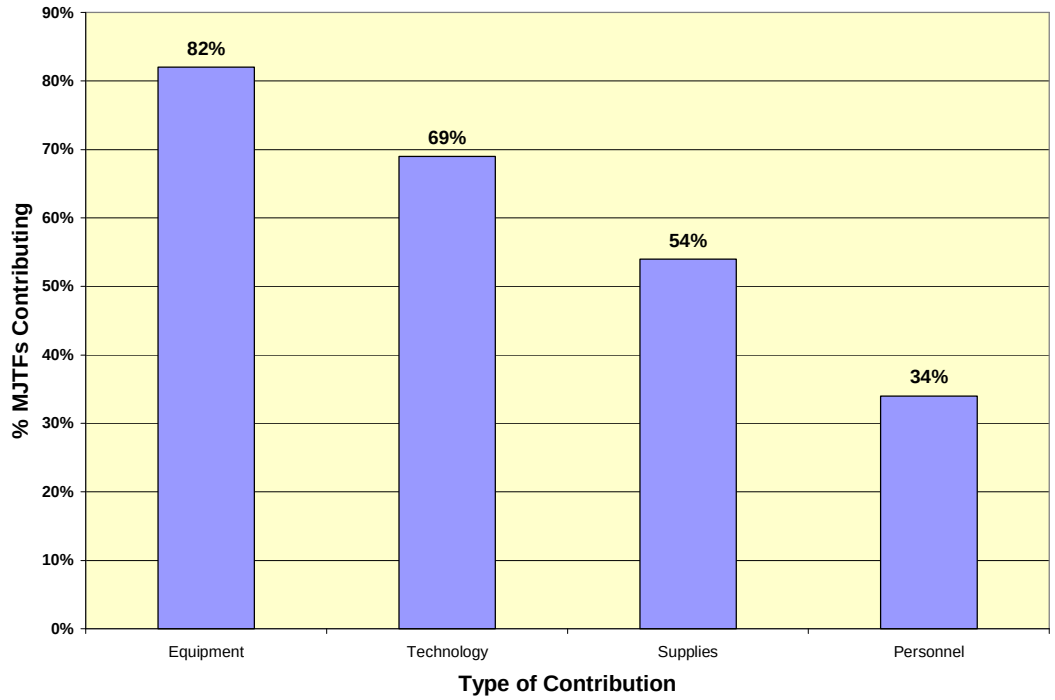
- % of MJTFs with state law enforcement agency participation
- % of MJTFs with federal law enforcement agency participation
- % of MJTFs with tribal law enforcement agency participation
- % of MJTFs with prosecutor/court participation
- % of agencies participating in MJTFs from each category: federal, state, local, tribal
- % of sworn staff from each of federal, state, and local level

Coverage Areas

- % of counties within the state that are covered by MJTFs
- % of the population within the state (population would have to be acquired from census data) covered by an MJTF

Figure 3.2

Percent of MJTFs Receiving Each Type of In-Kind Contribution, Statewide Levels



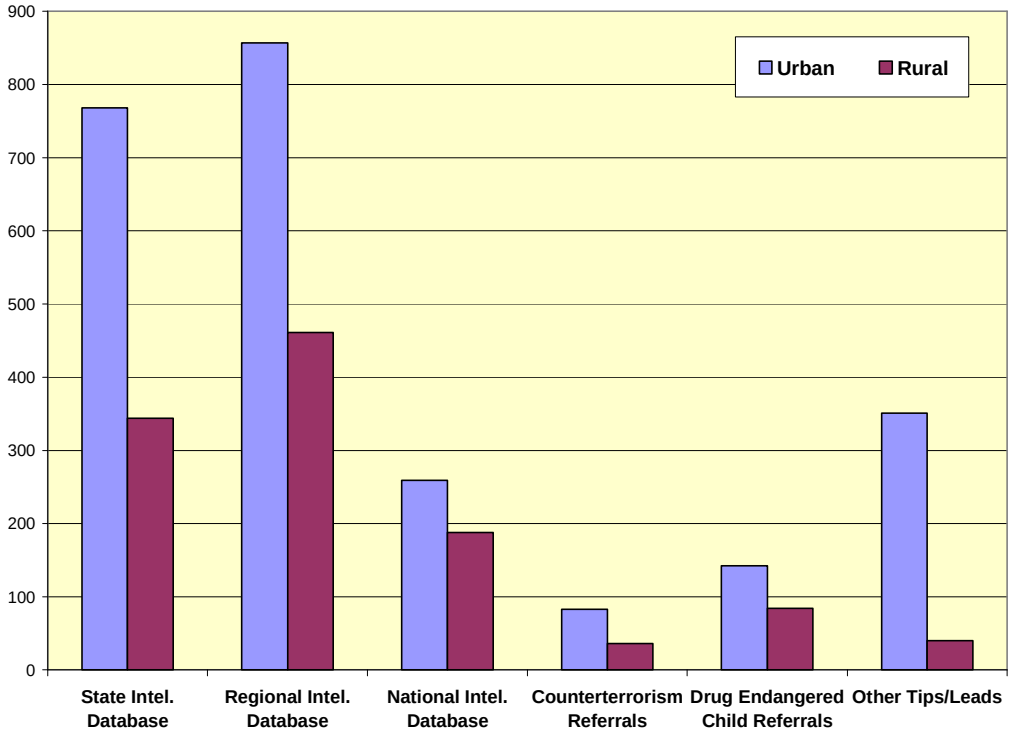
MJTF Structure and Policies

- % of MJTFs with an MOU or IAA
- % of MJTFs with a policy board
- % of MJTFs with a policy board that meets:
  - Monthly
  - Quarterly
  - Annually
  - As needed
- % of MJTFs with written protocols that include:
  - Chain of command
  - Mission statement
  - Goals and objectives
  - Policies and procedures
    - FOP for agents
    - Targeting and approving cases
    - Information referrals

Exhibits and summaries similar to those in Table 3.1 and Figure 3.3 can be made for these and other measures of task force structure and operations.

Figure 3.3

MJTF Intelligence Sharing, Statewide Levels



Here and elsewhere we encourage users to be creative and thoughtful when adapting the template. Measuring the number of FTEs dedicated to a task force is important and may be sufficient. Many MJTFs in rural areas essentially provide the only narcotics enforcement for the area, so knowing the number of sworn personnel is informative. Other MJTFs have complicated structures and missions. For example it might be helpful to partition the FTEs across enforcement, intelligence and training activities. The current template does not allow that detail, but the template could be modified if task force reporting requires that detail.

Describing MJTF Activities and Outputs

Ultimately, program success or failure depends on whether activities are correctly targeted and of sufficient frequency and magnitude to produce the desired effects. MJTF activities can be described in many different ways, including how the programs are organized and structured; the rules and policies that govern them; and the steps that exist to investigate drug-related crime, seize drugs, make arrests, and prosecute those arrested. Performance measures provide a means of documenting that these activities have occurred as intended. When evidence of program activity is collected consistently, it provides a foundation for (a) examining whether the resources are in place to adequately support the necessary activities, (b) tracking over time to assess whether activities are keeping pace with changing crime problems, task force objectives, and available resources, and (c) explaining outcomes.

As discussed in previous chapters on logic models and measurement issues, program monitors should distinguish between activities and outputs. Outputs are the *immediate results* of the program's activities. A task force, for example, may have as an objective providing public outreach and education to enlist support for its anti-drug efforts. To further that objective, the MJTF may conduct a series of public presentations and "training sessions" in communities within the task force coverage area. "Training" would be the activity; the number of sessions held and the number of persons trained would be considered "outputs."[4]

Intelligence Sharing

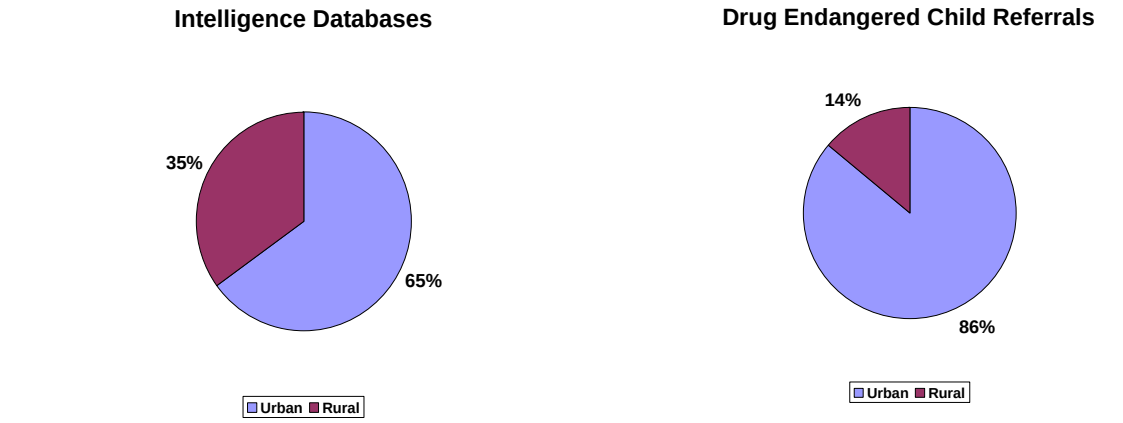
One of the fundamental reasons for the development of MJTFs is to enhance the flow of information across agency and jurisdictional boundaries, and the new performance measurement template provides several measures of this exchange of information (Items D1 to D4). To illustrate how these measures can be presented, Figure 3.3 displays the number of referrals to state, regional, and national intelligence databases, and referrals to other agencies about drug endangered child and counter-terrorism cases. For this display, the MJTFs are divided into those covering urban versus rural areas. This bar graph makes it readily apparent that in this State, most of the referrals are generated from MJTFs covering urban areas. This would be expected in a state where most of the population is concentrated in cities, and thus more criminal and MJTF law enforcement activity would be urban. States that are more heavily rural, with only one or two urban centers, might exhibit different patterns.

The proportions vary across each kind of database or referral. If it is important to highlight the rural versus urban differences across the kinds of intelligence databases and referrals, one can post the rural/urban split in percentages at the top of each column. Another option is to select contrasts of interest and produce pie charts, which effectively convey percentage splits across categories (Figure 3.4).

[4] Among program evaluators there is no consensus on how to determine what should be classified as measures of activity versus output. Law enforcement programs present special challenges, since arrests can be regarded as a program activity (engaging in the act of making arrests), an output (the number of arrest that occurred as a measure of the "arresting" activity), or as a program outcome (arrests serving as a proxy measure of crime). Given that there is no single "right" way to classify arrests, for the purposes of this guide we have elected to regard arrests as a program activity, and counts of arrests as program outputs. For the purpose of MJTF performance measurement, the most important issue is ensuring that the data are collected. How SAAs choose to classify measures as representing different components of the logic model is a secondary concern.

Figure 3.4

Comparing Rural/Urban Difference in Intelligence Database Entries Child Endangerment Referrals for all MJTFs



As before, we encourage creativity. An intelligence center may perform different activities that vary in intensity for supporting anti-drug operations. Intelligence analysts sometimes assist agents in the preparation of documents supporting prosecution – for example, organizational charts, time lines, and so on. They might also assist agents by supporting investigations by analyzing information gained using trap and trace devices and pen registers. The same staff members may perform or assist with labor intensive wire intercepts. Still other staff members might assist investigators with activities as diverse as mapping urban areas, tracing license plate numbers, and providing deconfliction services. When a task force supports an intelligence center, the performance monitoring system might usefully differentiate how labor is allocated across activities, and track trends in usage as a way to anticipate how future labor needs might differ from current labor needs.

Describing MJTF Investigations, Arrests, and Warrants Served

Investigating crimes and making arrests are basic police activities, and drug task force operations are often defined as disrupting and dismantling trafficking organizations and dealers. Non-drug task forces have their own specific definitions. Law enforcement agencies, program sponsors, and oversight bodies are experienced in collecting and presenting data on arrests. The Uniform Crime Reporting system has been in place for over 30 years, and reporting arrests at the MJTF level is a Byrne/JAG funding requirement.

While the concept of reporting arrests is not new, the purpose of the new requirements is to establish uniformity and consistency when reporting across MJTFs and states, and over time. This subdivision of arrests may represent a change in reporting for some task forces. Items in the template (items D10a to D10d) ask for task forces to report separately arrests for possession and for delivery of controlled substances—on the felony and misdemeanor levels—and to report on how many of the arrestees were gang members. A summary table is a good starting point for presenting and assessing this information. Table 3.2 presents one option for compiling descriptive data on investigations, warrants, and arrests. Counts for each task force can be presented, along with state totals and average numbers for each MJTF. Tabular data such as these can also be presented graphically.

Table 3.2

Summary of MJTF Investigations, Arrests, and Warrants Served

Task Force	Investigations				Warrants served	Arrests	
	New Opened	Pending	Dropped	Closed		Possession	Delivery
MJTF 1							
MJTF 2							
MJTF 3							
MJTF 4							
MJTF 5							
MJTF Totals							
Averages Per MJTF							

Again, we encourage creativity when adapting the template. Counts of investigations, arrests and warrants may be sufficient for a rural task force that has a street-level enforcement mission, but it can be misleading for a task force that has a mission to disrupt and dismantle higher-level trafficking organizations. For the latter organizations, an additional and perhaps preferable metric is the number of DTOs that were disrupted or dismantled. Of course a meaningful count requires valid and reliable definitions and measures of "DTO," "disrupted," and "dismantled".

II. Using the Performance Monitoring System to Track MJTF Programs Over Time

All performance measures can be tracked over time to document trends in program focus, structure, resources, activities, and indicators of effectiveness. Among the most useful tools for presenting trend data is a line chart, which can be easily produced using standard spreadsheet programs, and simple summary tables with year or other reporting period along one axis and the measure of interest along the other.

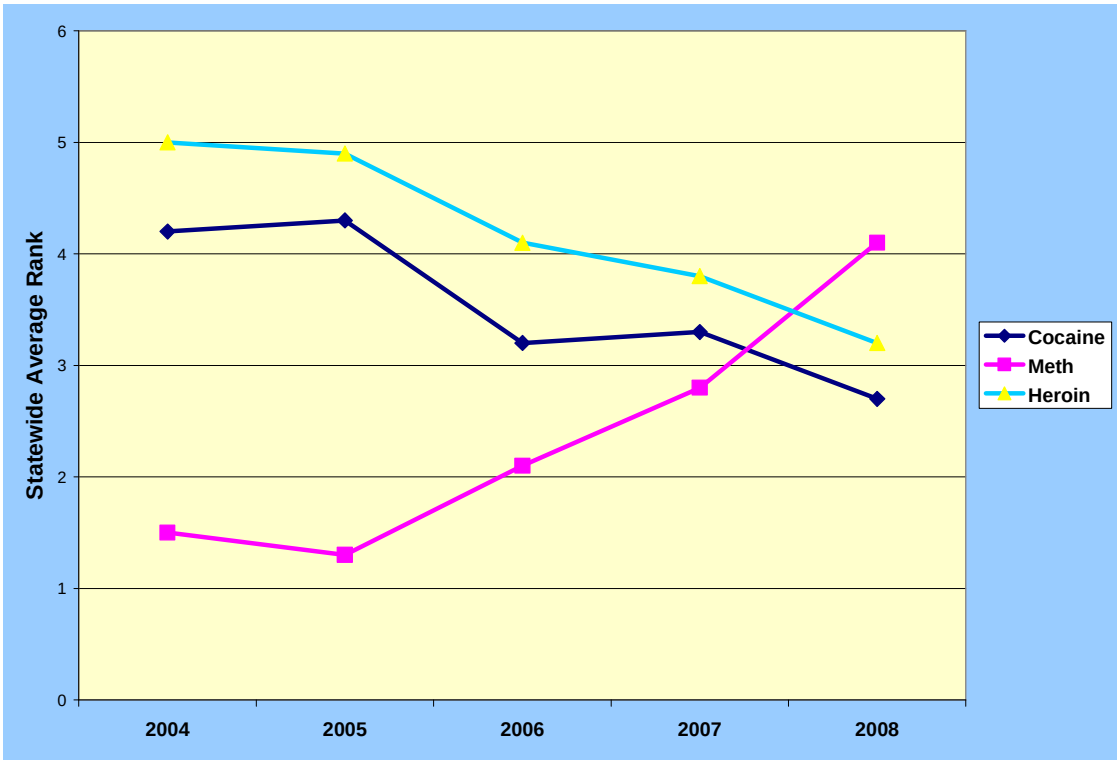
Tracking Task Force Objectives

Judging whether MJTFs are successful requires knowing what they are trying to accomplish. Goals define program objectives, and clearly-stated objectives are necessary for establishing measures of program effectiveness. Objectives can be thought of as "goalposts" for a program, and they need to be defined precisely so that progress reaching them can be measured unambiguously. All BJA-supported MJTFs share the broad goals of reducing drug availability and drug-related crime, and these goals have not changed over time. However, specific objectives vary across MJTFs due to different local crime problems and the fact that they are operating under different conditions. Moreover, even within a specific task force, circumstances change with time, so objectives do as well. As patterns in drug availability and drug-related crime change, MJTFs may shift their focus to adapt to those and other changing conditions.

The new performance measurement template quantifies MJTF goals by ranking the priority of a set of common task force objectives. The objectives of MJTFs within a state can be presented and tracked using a simple line chart of average priority rankings of the various categories of drugs addressed by task forces. In a hypothetical example presented in Figure 3.5, one can see that the task force objective rankings changed little from 2004 to 2005, then combating methamphetamine rose on the priority list, while combating marijuana and heroin fell.

Figure 3.5

Ranking Scores for MJTF Objectives, 2004-2008



Since objectives provide the foundation for the deployment of resources and prioritizing different activities, a change in objective will affect subsequent elements of the rest of the logic model. For example, an area experiencing an increase in the manufacturing and trafficking of methamphetamine might expect local MJTFs to shift focus from other types of drug enforcement and investigation. The MJTF's resource needs would also change (e.g., staff might need specialized equipment and training to meet meth-related challenges); as would its activities (investigations, meth lab seizures and eradications); and program outputs (e.g., the amount of meth seized, the number of lab eradications).

The same approach can be used to monitor task forces within groups. Task Forces fall naturally into certain categories, and it is a simple matter to group them for analytic purposes. For example, SAAs might find it useful to track and compare MJTFs across:

- Rural vs. urban areas
- Different regions of the state

- Areas with certain geographic features, e.g.
  - Interstate highway
  - Sharing a border with another state or country
  - Ports, coastal areas
  - Recreational areas, such as lakes and state parks

Tracking MJTF Staffing Levels

Our research on MJTFs found that most expenditure is for staff. Since staff levels drive the task force activity, SAAs will want to track the numbers of people deployed in task force programs. Figures 3.6 and 3.7 present two options for displaying such data. The first of these figures presents tracking data on FTEs for a subset of seven individual MJTFs. The graph illustrates that FTEs do not vary substantially from year to year for most MJTFs, while a few show significant trends indicating expansion or contraction.

When there are a small number of task forces, a line chart can effectively present their individual trends. Even with small numbers of programs, overall trends can become obscured by plots presenting individual programs. Many states have in excess of 25 MJTFs, and line charts for this number of task force can become unworkable. Therefore, it is often helpful to add a trend line for statewide totals, or to present a graph showing only the statewide trend line. As Figure 3.6 illustrates, the separate lines show individual MJTF trends but reveal little about overall trends, which are made apparent by the combined trend line (Figure 3.7). In this hypothetical example, there was a gradual increase in FTEs from 2002 to 2006, and a drop in the final year.

This graph raises a number of questions. For example: What drove the increase? Did an increase in drug-related crime lead to the increased commitment of staff? Was there a corresponding increase in MJTF activity, and if so, was it proportional to the increase in staffing? Answers to many questions like these can be addressed by examining other performance measures. For example, later in this chapter we discuss how to examine relationships between measures such as staffing levels and arrests.

Figure 3.6

Full-Time Equivalents Deployed to Sample of MJTFs, 2000-2007

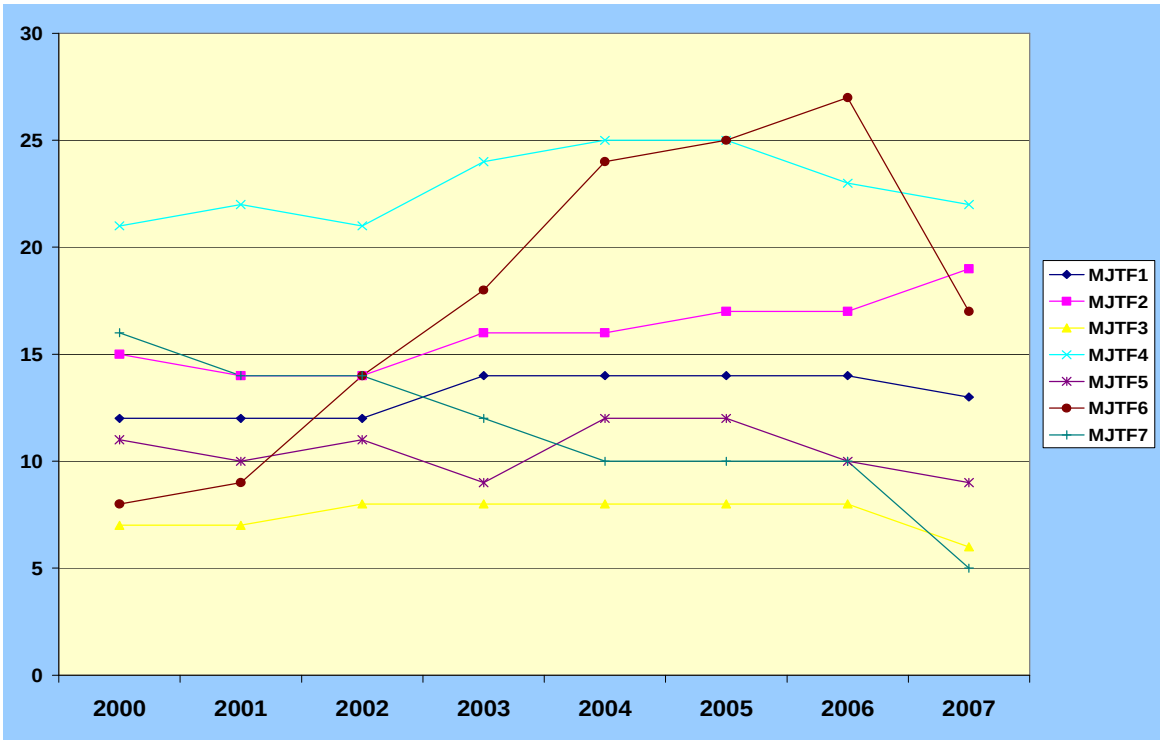
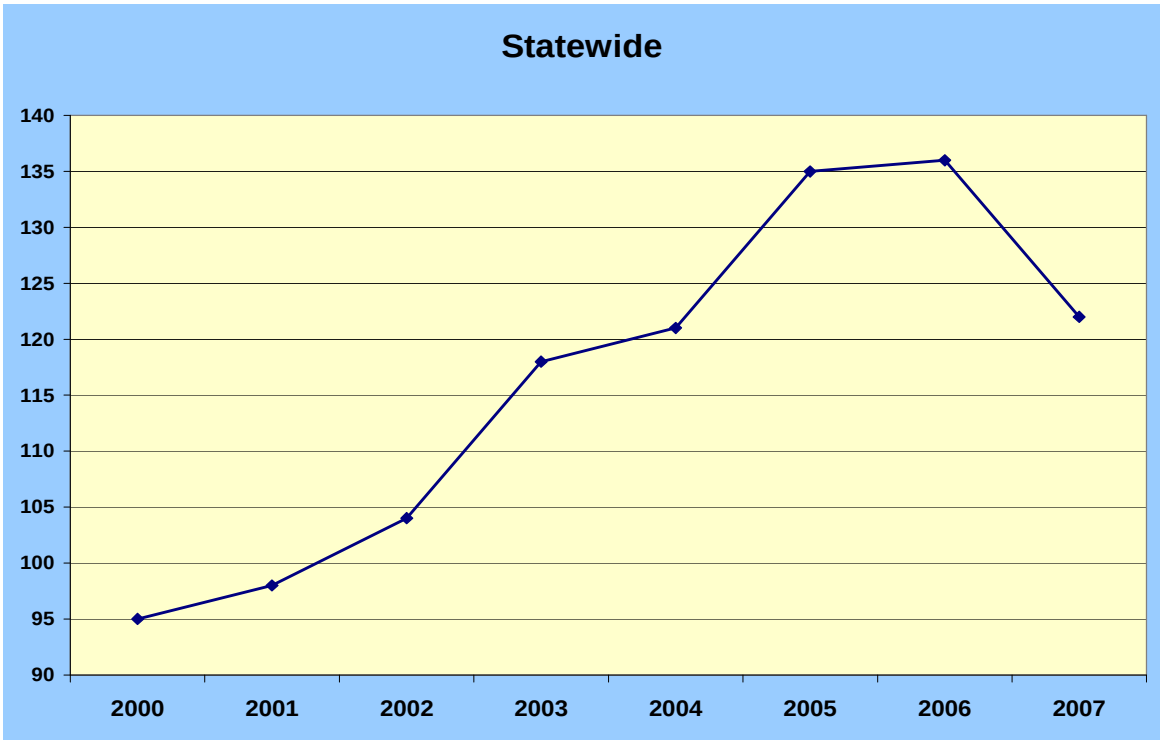


Figure 3.7

Full-Time Equivalents Deployed to MJTFs, Statewide Levels, 2000-2007



Tracking the Portion of MJTF Expenditures Offset by Revenue

To examine how program costs have been offset (or budgets supplemented) by asset forfeiture and fine revenue generated by MJTFs, one can begin by compiling data on total budget, total forfeiture and fine revenue for each year. Revenue can be divided by total MJTF cost, and the ratio interpreted as the portion of the MJTF cost offset by revenue. The hypothetical example presented in Table 3.3 provides an illustration. As can be seen here, MJTF revenue fluctuates substantially from year to year (although this is a hypothetical example, it reflects the variation said to occur by SAAs and task force commanders in our study of four states). When such variation occurs, an effective way to communicate the extent to which MJTFs recover their costs is to compute longer-term averages, rather than yearly rates. Totaling across years, one can see that the hypothetical task force generated $902,000 in revenue, and had operating costs of $11.65 million over this period. The MJTF recovered nearly 8% of its total operating costs through forfeiture and fine revenue.

Table 3.3

Costs of Operating a Hypothetical MJTF Offset by Forfeiture and Fine Revenue

	2000	2001	2002	2003	2004	2005	2006
MJTF Revenue	$175,000	$10,000	0	$475,000	$130,000	$90,000	$22,000
MJTF Costs	$1,450,000	$1,450,000	$1,450,000	1,600,000	2,100,000	2,100,000	1,500,000
% of MJTF Costs Offset by Revenue	12 %	1 %	0 %	30 %	6 %	4 %	2 %

When interpreting cost offsets for individual years, be mindful that there is often a substantial lag between the time the assets are seized and when the forfeiture amounts are made available. For each case there is a lag between seizure and a forfeiture ruling by the courts; then (in most cases) the non-cash assets must be liquidated, and the proceeds must be processed. Forfeiture decisions can be appealed and delayed for years in litigation. Thus, the forfeiture revenue realized by MJTF seizures in any particular reporting period may be the result of cases that occurred months or years earlier.

Although "earnings" is an intuitive metric for judging task force performance, the metric is misleading by attributing all earnings – some of which would have accrued without task force funding – to the task force. One really wants to know the marginal gain in earnings due to the task force, but that cannot be inferred from Table 3.3.

Tracking MJTF Investigations, Arrests, and Drug Seizures

As discussed above, MJTFs investigate crimes, make arrests, and seize drugs and assets; law enforcement agencies and oversight bodies are experienced in collecting and presenting data on related trends. The level of detail required in reporting on these measures has increased with the new performance measurement template, so the following discussion reviews basic methods and discusses a few less obvious means of presenting these data.

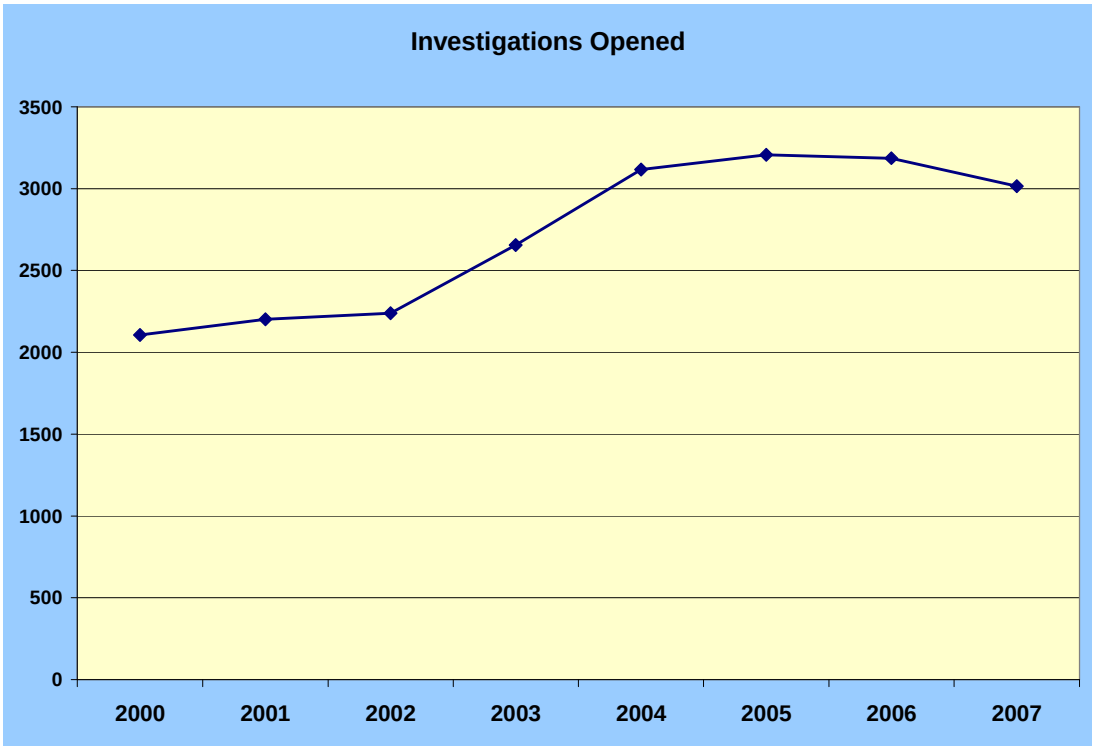
Investigations

If the data have been collected consistently, presenting the number of new investigations opened by MJTFs is relatively simple, and interpreting total counts over time is straightforward. The new template also requires MJTFs to report on (a) the number of investigations pending as of the end of the reporting period, (b) the number closed, and (c) the number dropped, which opens several possibilities for reporting and analysis.

The substantive interests of SAAs will drive the selection of presentation methods, but many SAAs are likely to find value in certain presentations. For example, one of the key assumptions underlying MJTFs is that the quantity and flow of intelligence across agencies and jurisdictions (along with increasing the number and training levels of sworn law enforcement staff) will lead to better cases. If so, one might expect that over years of investment in MJTFs, the proportion of cases opened should increase (Figure 3.8). One might also expect the number of cases dropped to fall, and the proportion of cases that are closed by making arrests to increase over time. To examine this, one can calculate the percentage of investigations closed, dropped, and pending each year.

Figure 3.8

Number of Investigations Opened by MJTFs, Statewide Totals, 2000-2007

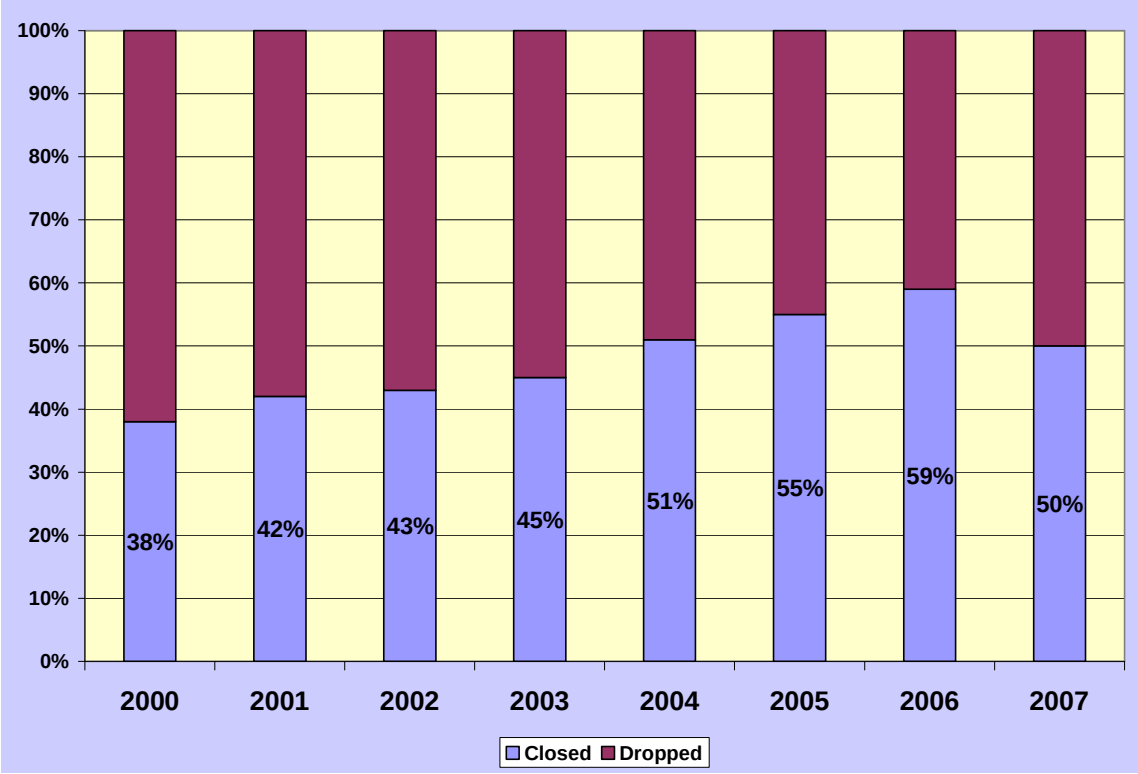


A divided bar graph such as that provided in Figure 3.9 can effectively present this information. As this graph shows, from 2000 to 2006 the portion of investigations ending in arrest ("closed") rose in relation to those that did not end in arrest ("dropped"). Whether the drop in 2007 is the beginning of a downward movement or "correction" back to lower levels, or merely a one or two year hitch in a larger upward trend, one cannot tell until data have been collected for subsequent years.

The trend through 2006 is positive and consistent with the intentions of MJTF programs. When faced with this information, an SAA might be tempted to attribute the trend to MJTF performance, and this is certainly a possibility. However, as in any analysis of performance measures, one must exercise caution when attributing cause and effect when examining whether MJTF activities have produced trends in program outputs and outcomes. In this example, there are many other influences on the ratio of closed versus dropped investigations besides task force intelligence gathering and sharing. For example, it may be true that the relationships among law enforcement agencies and the intelligence they share have continually improved, but other factors such as an influx of less sophisticated drug trafficking organizations, an increase in law enforcement staffing, or improvements in analyzing forensic evidence may have produced the increase in "successful" outcomes of investigations.

Figure 3.9

Percentage of Investigations Dropped versus Closed, Statewide Totals, 2000-2007



The message here is that one should not attribute any particular trend or result to the program without exercising due diligence in pursuing alternative explanations. One can and should take reasonable steps to examine whether an attribution of a result to the program is the strongest explanation for the result. To do this, one can generate a list of potential explanations for the trends, and then use whatever information is at one's disposal to help weigh their relative merits. The difficulty is determining what would have occurred absent the MJTF activity, or what occurred due to some other activity or event unrelated to the task force. (Examining relationships among variables is discussed in the next section of this chapter and in Chapter 4, but we will touch upon it here.) One can brainstorm

about all feasible explanations for these trends, and pare in to a "short list" of more likely explanations such as:

1. Shift in drug markets, from drugs and trafficking networks that are more difficult to investigate and prosecute to those that are less difficult.
2. Improved training of law enforcement personnel, resulting in better investigative and evidence-handling techniques that in turn produce stronger cases on which to base arrests.
3. The trend may not be a sign that any significant change of improvement in law enforcement; it could simply be a natural level of fluctuation that is hinted at by the decline in 2007, and would become apparent when longer trend lines can be observed.

Answering these and similar questions about opening and closing investigations is fundamental to performance monitoring, but answering these questions suggests elaborating on the graphs. Not all investigations are the same. Some target major distribution networks, and at some point, those investigations can consume most of a task force's resources, meaning that the task force necessarily delays culmination of other open investigations and may not initiate new ones. Lumping investigations together into a generic type can be misleading for task forces that work both large and small cases.

For a performance monitoring system to be useful when task forces work diverse cases, the performance measurement system requires some means for distinguishing scale across investigations. A meaningful definition of a drug trafficking organization is a starting point. See the earlier discussion on this topic that reported the US Sentencing Commissions role in the offense definition: "If the defendant was an organizer or leader of a criminal activity that involved five or more participants or was otherwise extensive…" A drug trafficking organization might be defined as a group that has such an organizer or leader. A monitoring system might further refine this definition. The DTO might be local or it might have intra-state, inter-state or international employees.

Our experience working with task forces shows that such terms need to be carefully defined. A DTO has an organizational hierarchy. It may be ephemeral because DTOs form and dissolve readily, but an organizational hierarchy is distinct from a buyer-seller relationship. For example, a local DTO is not international just because its cocaine is provided by Mexican cartels. Most investigations by MJTFs are not investigations of DTOs according to the Sentencing Commission definition. This does not preclude adopting definitions that differ from those provided by the Commission, of course; the observation nevertheless suggests the need for some meaningful and transparent definition that will usefully distinguish investigations by type. Although a count of investigations sometimes overstates the number of DTOs under investigation, we have observed that large scale investigations can involve multiple DTOs. This happens because of the way that investigations develop – over time, an extensive investigation will link multiple DTOs. It may disrupt or dismantle more than one as part of the same investigation.

Performance measurement can provide evidence or how MJTFs affect drug markets, drug manufacturing, and drug trafficking. Performance monitoring may draw on useful tabulations and cross-tabulations. But a performance monitoring system without good measures is useless and it cannot substitute for judgment.

Arrests

As discussed above, SAAs and law enforcement agencies have long been required to track and report on the number of arrests, so this guide does not discuss how to present simple arrest trends. However, the new performance measurement template subdivides arrests in ways that may not be as familiar to some MJTFs and states. Options for presentation of these data are discussed here.

The new template asks MJTFs to report on possession and trafficking of drugs, and further subdivides arrests as felonies or misdemeanors. There are a number of options for presenting this data, and the interests of SAAs will determine which methods to pursue. For example, a hypothetical set of MJTFs may have focused on higher-level drug trafficking, pursuing a smaller number of longer-term investigations aimed at arresting key operatives. If there has been a shift from a higher number of street-level arrests, one might expect to see a decrease in the overall number of drug arrests, but a larger proportion of drug arrests would be for felonies. Another state—or set of MJTFs within a state—may be focusing on cracking down on lower-level street dealers and end users (with arrest serving as the gateway to local drug courts and treatment programs), and may expect an increase in overall arrest numbers and the proportion that are for misdemeanors.

To examine this, one could collapse into a single category the possession and trafficking charges, and calculate the proportion of all arrests that are felonies and misdemeanors. The graph in Figure 3.10 presents trends in the percentages of felony and misdemeanor drug arrests made by MJTF staff. As this type of area graph makes readily apparent, the proportion of arrests for felonies varied from year to year but show a general downward trend. Such a trend would not be welcome news for MJTFs intending to make more felony arrests of traffickers, but would be expected for task forces that had increased their street level enforcement and intended to route drug users into treatment. To further examine the latter, one could also track the proportion of arrests for possession versus trafficking, and would expect to see an increase in arrests for misdemeanor (as well as felony) possession.

Other types of arrest trends that may be of interest to SAAs include percentages of trafficking versus possession arrests, and felony arrests made by urban versus rural MJTFs. The previous subsection explained why arrests may be a poor metric when the task force emphasizes the investigation of drug trafficking organizations.

Drug Seizures

The previously discussed methods of tracking, displaying, and interpreting data on arrests and investigations are readily applied to data on drug seizures. In monitoring a state's MJTF programs, SAA will want to begin with simple data tracking presentations, such as line charts of drug seizures by drug type across reporting periods. One consideration important to presenting seizure data is the need to ensure that when presenting together seizures of different drug types, that the units of measurement are standardized or, at least, made very clear and considered when interpreting trends and differences among drug types. For example, many drugs are commonly recorded in grams, but marijuana is often tracked in kilograms or pounds. If a single graph presents trends in seizures of cocaine, methamphetamine, and heroin, it is important to convert them all to the same unit of measurement.

Figure 3.10

Percentage of Arrests for Misdemeanor versus Felony Offenses, Statewide Totals, 2000-2007

In practice, it is often helpful to track marijuana separately when making tracking graphs. The amounts of marijuana (measured in grams) often are far larger than for seizures of other drugs, and a graph with a scale able to accommodate large amounts of marijuana can obscure trends in drugs seized in far smaller amounts. This is illustrated in the following figures. Figure 3.11 presents a line chart of five major drug types, including marijuana. As can be seen here, marijuana seizures occur in far greater amounts, and the other trend lines are all packed closely together at the bottom of the scale. Figure 3.12 presents the same trend data for the other four drug types, but with marijuana removed the scale is reduced, revealing the nature of the variation over time in seizures of the other drugs that was not apparent when marijuana was included.

Tracking MJTF Outcomes: Convictions and Sentences

While investigating and arresting offenders are the key law enforcement activities, they will have little impact on crime if the charges are not prosecuted successfully and do not result in substantial criminal penalties. If MJTFs operate as intended, one would expect increases in the proportion of cases referred for prosecution that are accepted, an increase in conviction rates, and longer prison sentences. These expectations are grounded in the assumption that MJTFs produce better intelligence, communication, and collaboration among police agencies, and more well-trained

Figure 3.11

Trends in Drug Seizures, by Drug Type, Measured in Grams

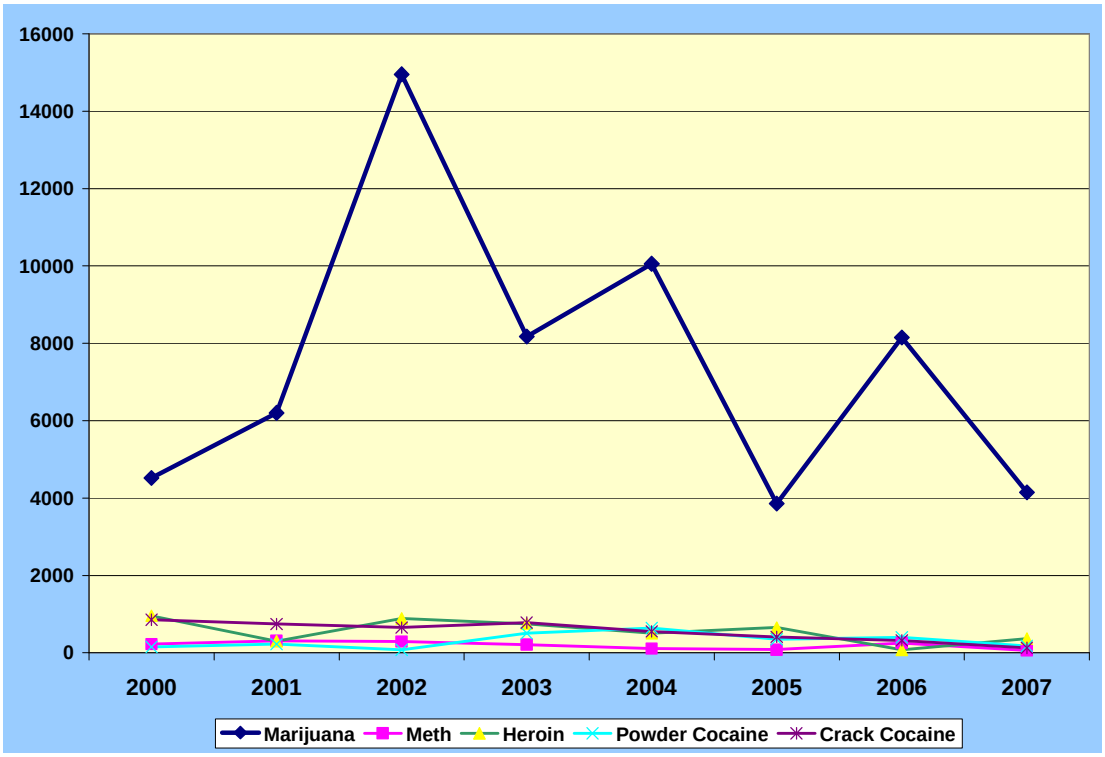
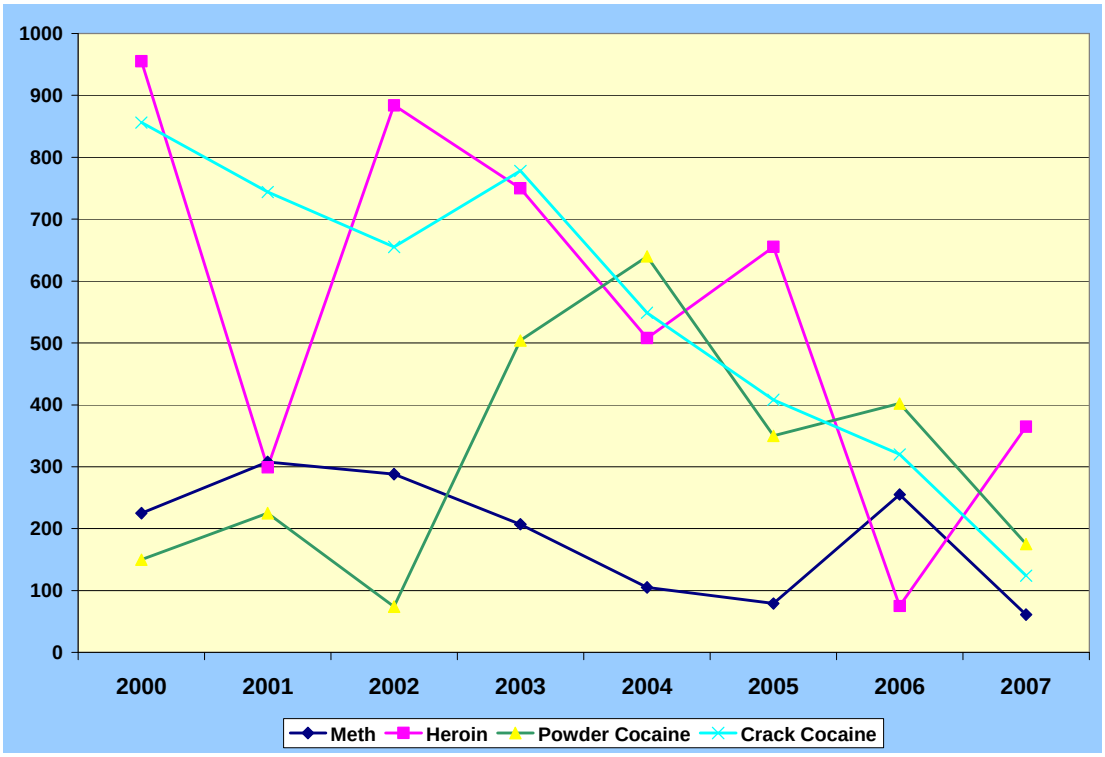


Figure 3.12

Trends in Drug Seizures, by Drug Type, Measured in Grams—With Marijuana Removed

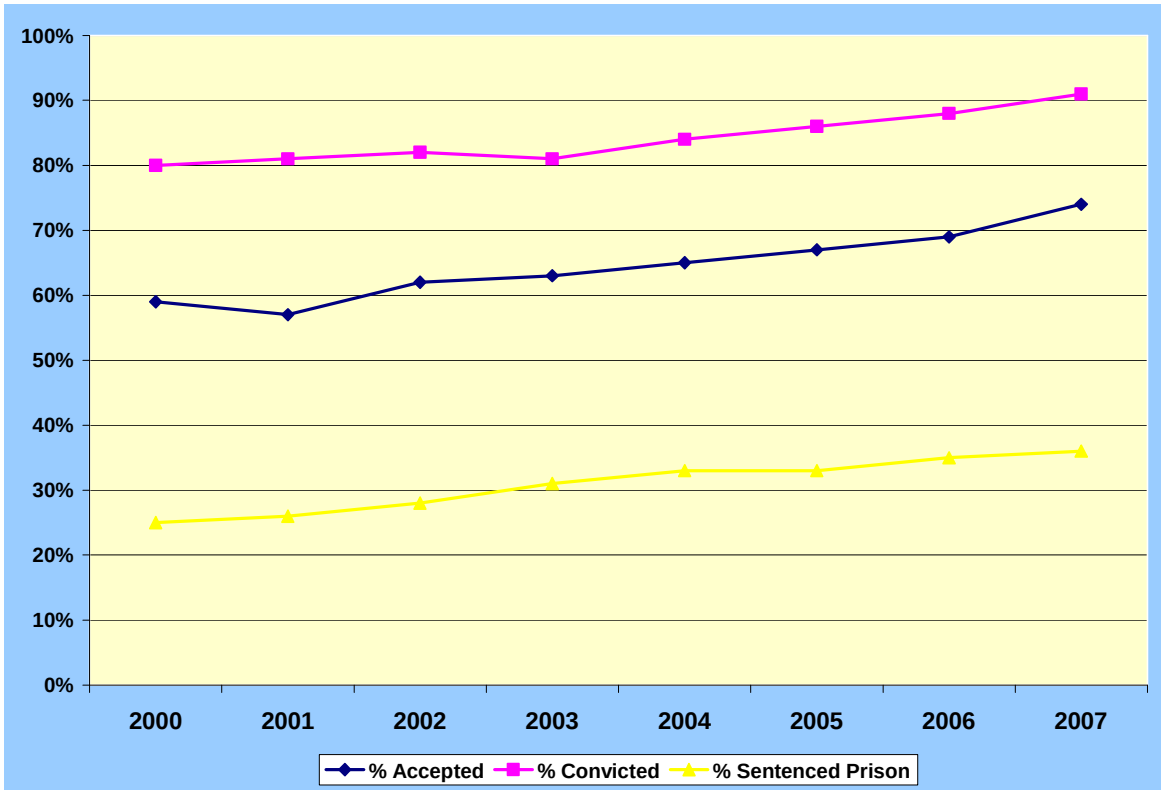


investigators who can better assemble, preserve, and present evidence. Thus, MJTF should be presenting prosecutors with stronger cases that can withstand the efforts of defense attorneys, put prosecutors in a stronger position for plea bargaining, and support longer sentences.

Figure 3.13 presents one option for presenting performance data on MJTF outcomes for a hypothetical state. In this example, trends in performance indicators are clearly going in the desired direction, with the proportion of referred cases accepted by prosecutors, the conviction rate, and proportion receiving prison sentences all increasing.

Figure 3.13

Percentage of Cases Referred to Prosecution That Are Accepted, Result in Conviction, and Result in Prison Sentences, Statewide Totals, 2000-2007



Using convictions and sentences as a measure of performance may be misleading for some task forces. For a task force whose mission is low-level drug-law enforcement, the goal may be diversion through drug court. Increasing convictions would contraindicate improved performance, which would be better judged by the number of drug court graduates. Diversity across MJTFs requires diversity across measures.

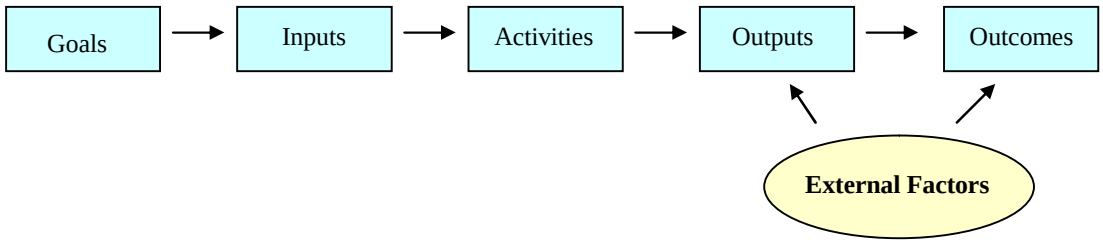
III. Using Performance Monitoring Systems to Examine Linkages in MJTF Logic Models

The new template provides measures that support examining the linkages between and among components of an MJTF logic model. A number of analytic techniques can be applied, but the basic approach involves examining the relationship between measures of different components of the model. For example, as the number of staff (a program input) increases, what happens to arrests and investigations? If the correlation between these two variables is positive (i.e., as MJTF staffing rises, arrests increase correspondingly), it provides evidence that the program is having its desired effect. However, one cannot assume that program results and outcomes are determined solely by the efforts of the program, for there may be other external forces at work that obstruct or contribute to a program's successes. For example, the creation or expansion of MJTF programs may result in a dramatic increase in numbers of drug traffickers arrested. But if the U.S. drug market has experienced a surge in the availability of a particular drug, illicit drug use may increase even as task forces reach their goals of increased numbers of arrests and convictions.

Determining with high levels of certainty whether this is a causal relationship (attributing the increase in arrests to the MJTF program, rather than to other causes as depicted in Figure 3.14) is an involved process that normally requires highly sophisticated research designs and statistical techniques that are beyond the scope of this Guide and usually require contracting with evaluation specialists. However, the application of simpler monitoring techniques can provide valuable feedback about whether the program appears to be on the right track.

Figure 3.14

Role of External Factors



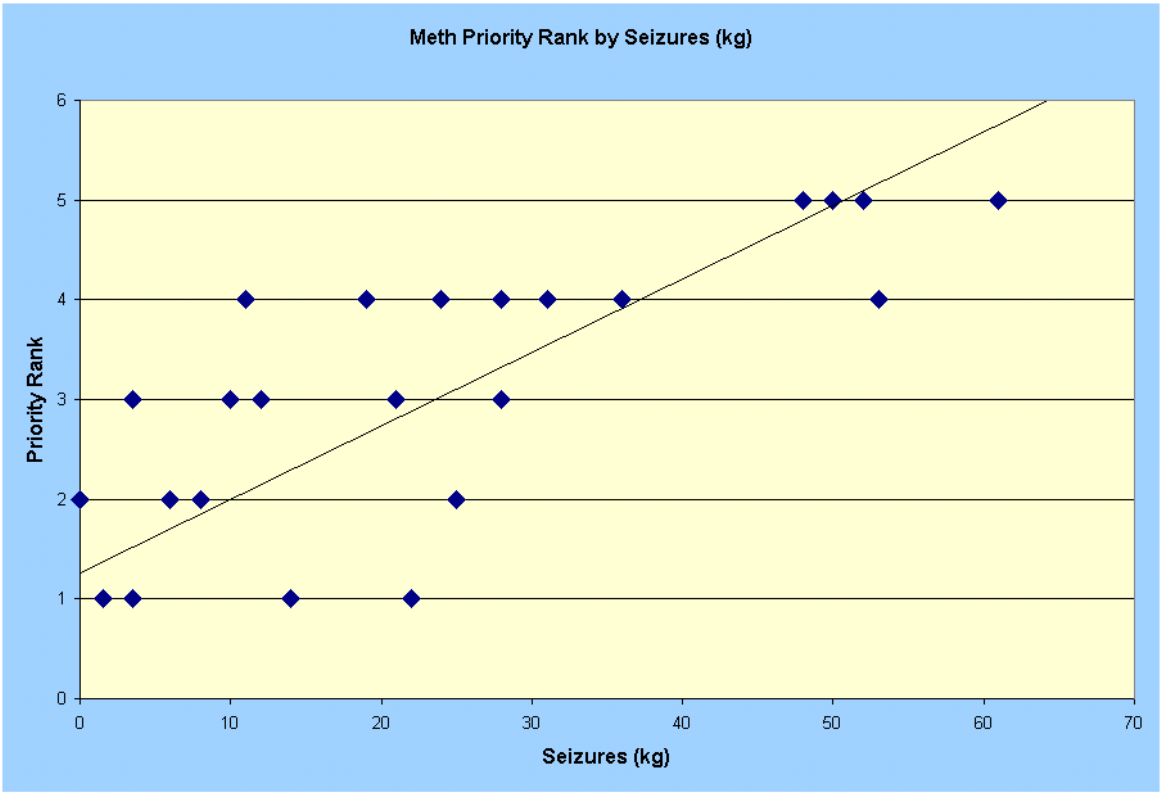
Examining Links Between MJTF Objectives and Outputs

The template identifies measures of drug-specific task force objectives (measurable derivatives of goals) and activities (drug seizures). Figure 3.15 provides an illustration. One would expect that greater amounts of methamphetamine would be seized by task forces that are focused on the objective of combating problems associated with that drug. To examine whether this is so, one can produce a scatter-plot in which the priority rank that each MJTF has given to meth is on one axis, and the amount of meth seized that year for each task force is on the other axis. Each point on the plot represents the intersections of meth priority rank scores and meth seized for each of 25 hypothetical MJTFs. The diagonal "regression line" represents the relationship between the two variables. The

line rising from left to right in Figure 3.15 means that as an MJTF's meth priority rank increases, the amount of meth seized by that MJTF tends to increase.

Figure 3.15

Correlating MJTF Objectives with Drug Seizures



Examining Links between MJTF Resources and Activities and Outputs

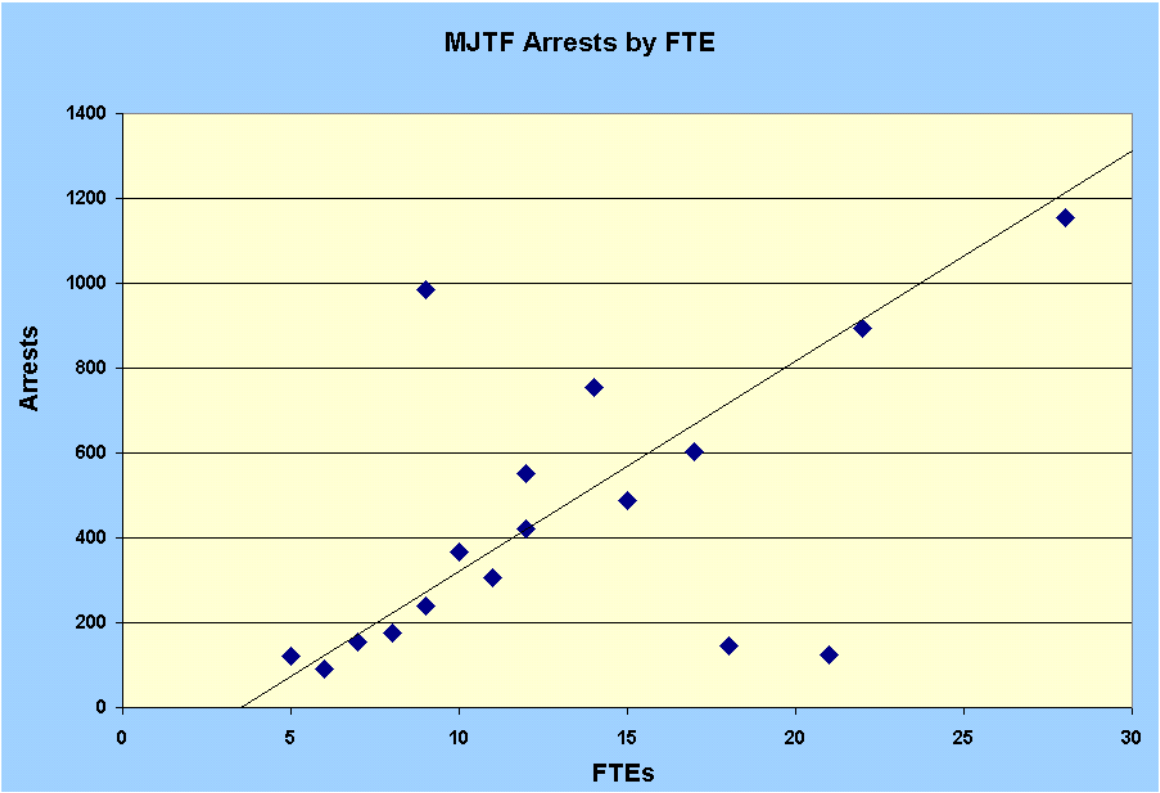
The MJTF template provides data supporting examination of linkages between several kinds of program inputs and activities or outputs. For example, the key resource devoted to MJTFs is staff, which the performance measurement template operationally defines as the number of full time equivalent positions (FTEs). To examine the relationship between this resource and the MJTF activities, one can produce scatter plots of FTEs and measures of task force activity, such as the number of submissions to intelligence databases, number of trainings conducted, and investigations opened; one might make a similar linkage with measures of task force outcomes, such as individuals arrested or drugs seized. As staff levels increase, most indicators of MJTF activity should increase correspondingly, else one would question whether the additional investment was worthwhile.

Time-series analysis – examining how output and outcomes vary over time – is a variation on simple cross tabulation. The difference is that the horizontal axis would show variation in FTEs for an individual task force. The presumption is that the outputs and outcomes would improve as FTEs increase, and outputs and outcomes should not deteriorate if FTEs remain constant. Evidence that contradicts these expectations calls for closer examination.

Figure 3.16 illustrates the relationship between hypothetical staffing levels and arrests. Each point on the plot represents an MJTF, and its position in the plot is determined by the number of arrests made on one axis, and its FTEs on the other. The diagonal regression line represents the relationship between FTEs and arrests. How tightly the points are grouped around the regression line can usually be interpreted as indicating the strength of the relationship between the two variables. If there is a weak relationship, the points on the scatter plot will be widely dispersed around the diagonal line. When the relationship is stronger, the points will be more clustered around the line. The results

Figure 3.16

Correlating MJTF Staffing Levels with Arrests



displayed in Figure 3.16 are consistent with expectations, where MJTFs with more FTEs tend to make more arrests. The line rising from left to right in Figure 3.16 means that as an MJTF's FTEs increase, the number of arrests by that MJTF tends to increase.

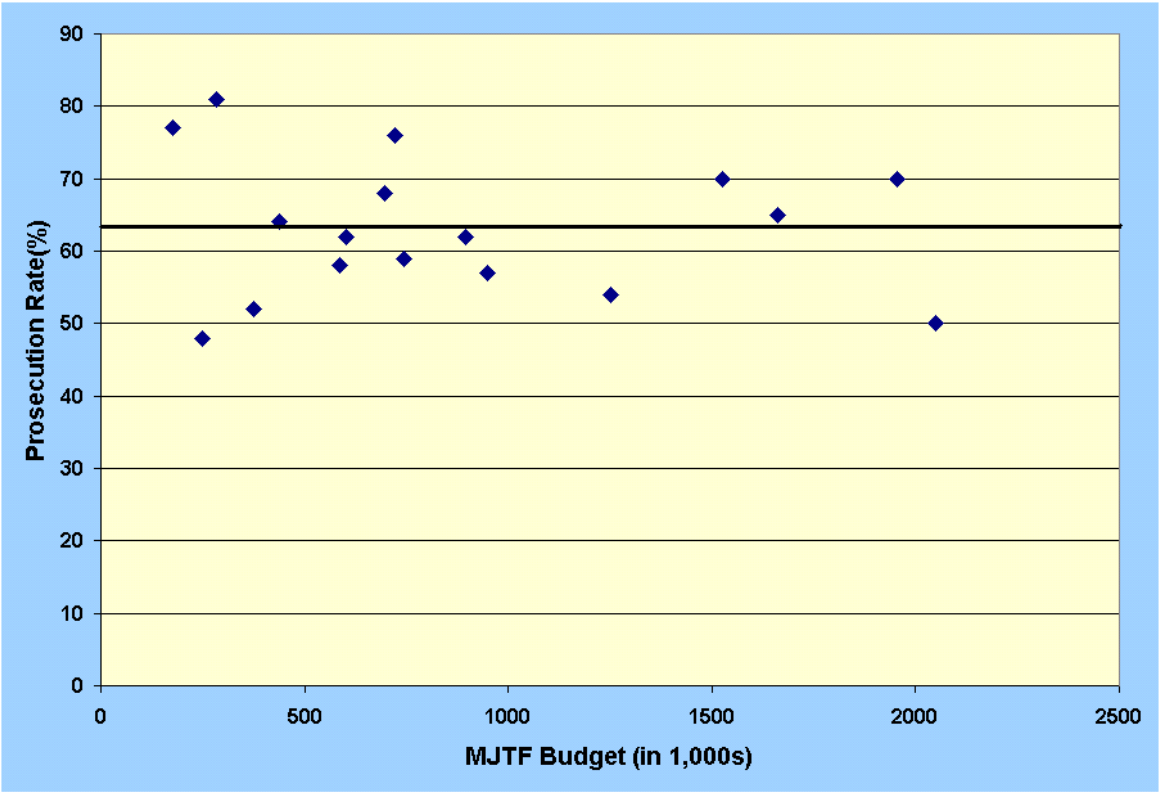
Figure 3.17 illustrates the relationship between task force budgets and prosecution acceptance rates. One might expect that as more is invested in MJTFs, the quality of cases might rise and result in higher rates of acceptance for prosecution. To examine this, one can plot total budgets and the rates of case acceptance for prosecution for each MJTF. Budgets are presented in thousands, and acceptance rates as the percentage of referred task force cases that are accepted for prosecution.

Figure 3.17

Correlating MJTF Budget and Prosecution Acceptance Rates

Figure 3.17

Correlating MJTF Budget and Prosecution Acceptance Rates



In this example, there is variation in both task force budgets and in prosecution acceptance rates, but there is no evident correlation between task force budgets and prosecution acceptance rates. Note that the regression line is horizontal, meaning that as the value of one variable (budgets) increases, there is no corresponding increase in the values of the other variable (acceptance rate). A plausible explanation for this result is that the number of prosecutions may increase with FTEs while the percentage of cases remains constant. That is, large MJTFs have the same acceptance rate as do small MJTFs. The former simply have more cases.

Other correlations that SAA can consider tracking using performance measures collected by the new template include:

- Budget by number of defendants referred for prosecution
- Budget by conviction rate
- FTEs and sentence length
- FTEs and number of offenders convicted
- Per-investigator rate of submissions to intelligence databases (# of submissions divided by investigator FTE) by number of defendants referred for prosecution

Some kinds of data are most effectively presented in the form of cross-tabulations, which convey the relationship between variables by showing how they are distributed when considered simultaneously or jointly. For example, to learn more about influences on prosecution acceptance rates, one can examine the relationship between prosecutor participation in MJTFs and MJTFs having written policies and procedures for targeting and approving cases. One might expect that prosecution rates would be higher when district attorneys or other prosecutors are formally involved in MJTFs, and when MJTFs have clearly articulated standards and procedures for approving cases for investigation. To examine this, one can produce a table like that presented in Table 3.4. This table presents the prosecution acceptance rates for MJTFs with and without prosecutors as formal task force partners, and those with and without formal, written policies for targeting and approving cases. As would be expected, prosecution rates are higher for MJTFs with both formal policies and formal involvement of prosecution.

Table 3.4

Prosecution Acceptance Rates by Having Prosecutors as Formal MJTF Partners and Having Formal Policies for Case Targeting and Approval

		MJTF Has Prosecutor as Formal Partner	
MJTF Has Case Targeting & Approving Policies		Yes	No
	Yes	78%	62%
	No	66%	54%

Monitoring Using Tabulations versus Monitoring Using Statistical Analysis

This chapter has demonstrated how a performance monitoring system can make simple but powerful use of data from a performance measurement system. Advanced statistical analysis can move a performance monitoring system from description and simple associations between variables toward evaluation. The next Chapter discusses some elements of statistical analysis that will prove useful as the performance measurement system matures. It illustrates some aspects of that analysis using retrospective data from four task forces.

Chapter 4: Performance Monitoring: Practical Steps and Advanced Analysis

Monitoring the performance of MJTFs involves not only looking at the relationships between various aspects of their operations and the outcomes of those operations, but also examining the statistical significance of observed changes over time. Sometimes the task is simply to consider the association between two parts of a program model, such as the relationship between the number of cases successfully prosecuted and the number of FTEs assigned to case development. Other times the dynamic nature of the system is of interest and the task is to examine the components of the model in a time series. Examples of both of these analyses are examined in the Workbook provided in this chapter.

The three previous chapters provided the core for a performance monitoring system. As described in Chapters 1 and 2, implementing the performance measurement component should be straightforward, although not necessarily easy. As outlined in Chapter 3, using basic performance monitoring data in tabulations, cross-tabulations, and graphics should be within the reach of most SAAs. However, applying advanced statistical analysis to monitor and evaluate MJTFs is more challenging. As described in this chapter, some of that analysis should be achievable by BJA and SAA staffs, and the rest may require the use of consultants with advanced analysis skills.

This chapter first addresses issues that the analyst needs to consider when moving from a display of data to statistical interpretation of those data and provides a spreadsheet for preparing data and making common calculations that are likely of interest to SAAs. The Workbook first focuses on the ever-present difficulty of missing data and provides guidance on how to address imputation. It then walks though a spreadsheet designed to assist in analyzing trends and calculate correlations between inputs and outputs. Finally, the Workbook provides a guide to looking at performance data as a time series.

Accumulating data using the template presented in Chapter 2 will provide the basis for these analyses. This spreadsheet-based workbook should be usable for SAAs in applying statistical analysis to performance monitoring data now and as they gather data in the future. The example analyses presented here use seven years of retrospective data from three task forces and rely on spreadsheet tools. Without burdensome manipulation, however, spreadsheets cannot support more advanced statistical analysis. Consequently, this chapter also makes recommendations for performing more advanced analysis, but actual implementation is likely to require a social science researcher and statistical computing software.

The Workbook

The **Workbook** file is an Excel workbook associated with this guide. It comprises multiple spreadsheets. The first spreadsheet (**DATA**) is a selection of seven years (2000 to 2006) of retrospective data from various Multi- Jurisdictional Task Forces in Tennessee, Colorado, and Illinois. (Georgia was unable to provide the requisite data for the full period.) The original data provided by the Task Forces appear as normal (uncolored) numbers under the “Original Data”

header.[5] Although the data are "real", the specific data were selected for purposes of illustration. Readers should not attach much meaning to the substantive findings from these "real" data. They are provided as illustrations only.

The variables are:

INPUT VARIABLES

YEAR	This is the year.
MJTF	This specifies the MJTF. In order to simplify the structure of this spreadsheet and to ensure anonymity, MJTFs were randomly assigned a number and not otherwise identified.
FTE	This is the full time equivalent task force personnel.
BUDGET	This is the MJTF budget in dollars.

OUTPUT VARIABLES

OPENED	This is the number of cases opened by the task force.
PROSECUTED	This is the number of cases prosecuted by the task force.
DROPPED	This is the number of cases dropped by the task force.
ARRESTS	This is the number of arrests made because of task force investigations.

These variables do not exhaust all types of retrospective data, but they will be used to demonstrate how monitoring data might be applied to more sophisticated analyses.

Imputing Missing Data

Even in the best monitoring systems, it is not uncommon to find that data are missing for some variables for some jurisdictions. For some purposes, it is convenient to impute responses for missing data. For example, suppose that an analyst wanted to plot trends in arrests over the last seven years: 2000 through 2006. When some of the MJTFs fail to report arrest data for some of the years, the year-by-year reports would fluctuate with missing responses and trends could be misleading. There are a number of ways to deal with this problem. One approach is to drop every district that has missing observations for *any* years, but that approach gives an incomplete picture and can be misleading if many MJTFs are excluded from the analysis. A second approach is to impute responses when the responses are missing. Imputation requires some assumptions, and one reasonable assumption is that the ratio of the variable of interest (call it X) is roughly proportional to the budget. An alternative assumption might be that X is proportional to FTEs. This proportionality can be computed when data are provided, and then the missing data can be replaced with imputed values. Columns J through P in spreadsheet **DATA** include these imputations in red and are under the "Imputed Data" header.

The formulas embedded in the spreadsheet show how the imputations were actually done. The numbers within the grey highlighted areas in the spreadsheet are the proportions calculated for the input variable (BUDGET) and the respective output variables' (X) missing data. In order to calculate this proportion, the years that contain known data for both variables are summed for each of the two

[5] We have also attached spreadsheets for each of the three states participating in the BJA study designated as MJTF Template for IL, MJTF Template for TN and MJTF Template for CO.

variables separately. The output variable's sum is then divided by the sum of the input variable. Below we present an example from the spreadsheet:

> Example: Task Force 2 in the spreadsheet is missing data for the variable DROPPED for 2000 (cell O11). To impute the missing value for this cell the following calculations are made:
>
> 1. The known DROPPED data for 2001 to 2006 are summed
>   DROPPED Sum = O12 + O13 + O14 + O15 + O16 + O17
> 2. The corresponding data for BUDGET is also summed
>   BUDGET Sum = L12 + L13 + L14 + L15 + L16 + L17
> 3. The sum of the known DROPPED data is then divided by the sum of the known BUDGET data
>   O18 = DROPPED Sum/ BUDGET Sum
>   This number is represented in the spreadsheet by the number highlighted in grey at the bottom of the DROPPED variable column in cell O18.
> 4. This proportion (O18) is then applied in order to determine the missing data for DROPPED in 2000 by multiplying it by BUDGET for 2000
>   O11=O18 x L11

While the above example provides a standard methodology, various circumstances may complicate calculations. The following discussion attempts to address issues that may arise that require an alternative methodology than that presented above.

Sometimes the BUDGET may be missing data for one or more years. In this case, FTE (rather than BUDGET) may be used as the input variable in order to impute an output variable. For example, in Task Force 7, BUDGET and PROSECUTED are both missing data for 2003. Since BUDGET was known for 2000 to 2002 and 2004 to 2006, PROSECUTED could be imputed for 2000 to 2002 and 2004 to 2006 using BUDGET. However, since BUDGET is missing for 2003 FTE is used to compute the imputed value.

Known FTE values may also be used to impute unknown BUDGET values. In Task Force 5, for example, BUDGET is missing for 2005. The sum of corresponding available data for each of the input variables is calculated and then the BUDGET sum is divided by the FTE Sum to establish the proportion to be used in the imputation (L42). Finally, this proportion is multiplied by the FTE data available for 2005 (K40) in order to establish the imputed BUDGET for 2005 (L40).

BUDGET Sum = L35 + L38 + L39

FTE Sum = K35 + K38 + K39

L42 = BUDGET Sum/ FTE Sum

L40 = L42 x K40

The same method can be followed when imputing an unknown FTE value from a known BUDGET value.

Note: *This method of imputation can only be used when known corresponding input variable data exists (e.g. both BUDGET and FTE are known for the same year, at least once).*

Note: *If a BUDGET (or any other variable) value has been imputed, that specific value should never be used to impute another value.*

Another issue that may arise is that both the BUDGET and FTE variables for a given task force are missing data for the same year(s). These variables may be imputed using a known output variable. Task Force 5 provides an example of this situation for 2006. The second proportion highlighted in grey for both BUDGET and FTE (L43 and K43, respectively) is established using known ARREST data. The sum of the known BUDGET data is divided by the sum of corresponding known ARREST data. This proportion is then applied to the missing BUDGET data for 2006 by multiplying it by the ARREST data for 2006. The same process is followed when computing the unknown FTE variable. It should not be computed with the imputed calculations for BUDGET.

BUDGET Sum = L35 + L36 + L37 + L38 + L39

FTE Sum = K35 + K 36 + K 37 + K 38 + K 39

ARREST Sum = P35 + P36 + P37 + P38 + P39

L43 = BUDGET Sum/ ARREST Sum

K43 = FTE Sum/ ARREST Sum

If data for a specific variable are missing for every years (e.g. MJTF 1's PROSECUTED and DROPPED columns) then no data can be imputed. Similarly, if there is no known data across all variables for one or more years (see Task Force 6 for example) no values can be imputed for these years. The spreadsheet converts these missing responses to zero; it identifies them with yellow shading. The zeros will not add anything to a summation, so those districts will be excluded from the tabulations for specific variables.

This approach to imputations is reasonable for most tabulations provided that missing data are not a pervasive problem. The more missing data, the less compelling any argument is that relies on data imputations. However, simple imputations will not work for statistical Inference. Consequently this chapter returns to whether or not missing responses should be imputed when applying statistical analysis.

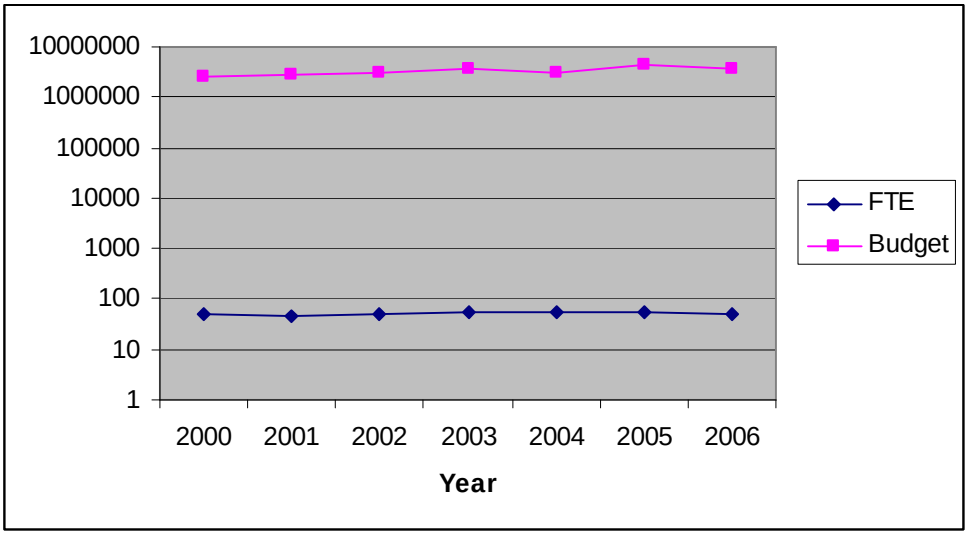
Trends in Inputs

For some purposes, the analyst may want to sum variables across the MJTFs for a particular year. As already noted, these summations use the imputations, and the summations altogether exclude MJTFs that failed to report a single value for a variable over all seven years. The resulting tabulations appear in the second spreadsheet called **TRENDS**.

Using the observed data and imputations reported in the previous section, **Figure 4.1** shows trends in annual budget and full-time equivalent employees across the seven years. Because of the differences in scale for budgets and employees, the figure shows the trends on a logarithmic scale.

Figure 4.1

Trends in Full-Time Equivalent Employees and Budgets, 2000 – 2006



Although logarithms are convenient for drawing the figures, they are not especially convenient for reporting numbers. **Table 4.1** is a substitute.

Table 4.1

Annual Budget and FTEs with Imputations for Missing Data

YEAR	FTE	BUDGET
2000	50	2,656,504
2001	46	2,686,101
2002	51	3,161,696
2003	53	3,635,644
2004	53	3,200,027
2005	55	4,287,799
2006	50	3,522,176

Based on Figure 4.1 and Table 4.1, the conclusions seem clear. Although FTEs fell between 2000 and 2001 and between 2005 and 2006, on average the overall FTE only increased by about one percent per year. Budgets also fell during two years (2004 and 2006), but on average grew by about one percent per year. Data were selected for illustration, so a reader should not draw a substantive conclusion from this illustrative analysis.

Correlations between Inputs and Outputs

Continuing to rely on imputations, **Figure 4.2** shows correlations between BUDGET and selected outputs. **Figure 4.3** shows the comparable correlation between FTE and the same selected outputs. An analyst should exercise care when drawing trends, because when an imputation of 0 appears in the first spreadsheet (these are denoted with yellow and occur because the jurisdiction failed to report this variable over the entire seven year period), the analyst cannot use the 0 value and should not use the associated BUDGET/FTE in the tabulations. The necessary adjustments appear in spreadsheet 3, called **CORRELATIONS**. The spreadsheet also shows how to compute the figures shown below.

The interpretation seems straightforward. Outputs (i.e., numbers of arrests, cases opened, prosecutions) are clearly a function of operating budgets and FTEs (inputs). Again, these data are illustrative, so a reader should not draw conclusions about MJTF performance from this analysis.

Statistical Significance of Correlations between Inputs and Outputs

The historical data comprise a *time-series* of *cross-sections.* By cross-section, we mean that the data comprise multiple MJTFs; each MJTF is an element of the cross-section. By *time-series*, we mean that the data comprise multiple years. Data with cross-sectional and time-series elements are often called panel data, and a number of techniques are available for dealing with panel data.

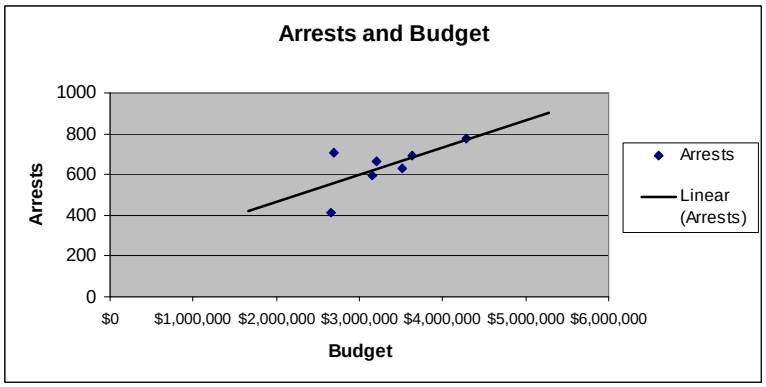
Two techniques are especially useful. The first is known as a fixed-effect regression model. This model seeks to explain how an outcome variable (such as arrests) varies with an input variable (such as the budget) taking into account the identity of the specific MJTF and the year. Evaluators recognize this as a least squares regression that has the outcome as the dependent variable (what is predicted), the input as the principal independent variable (what is used to predict), and dummy variables for each MJTF and each year. This model leads to an estimate of how the input variable affects the outcome variable.

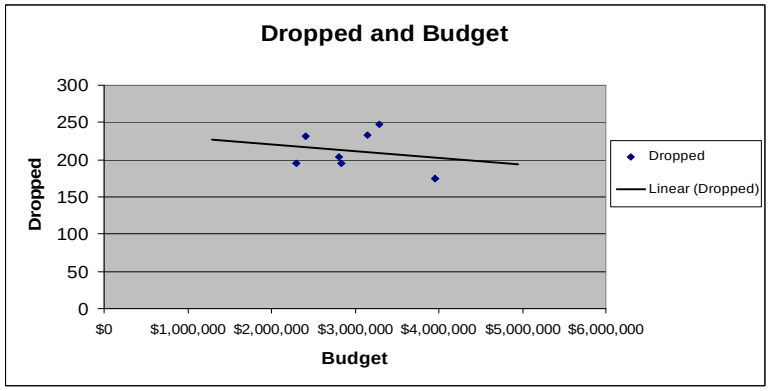
A second model—a random effects model—is similar, but instead of fixed effects for each MJTF, the random effects model assumes that the differences across districts can be captured by a random variable in lieu of a the dummy variable for the jurisdictions. The random effects model makes more efficient use of data, so it is preferable to the fixed-effect model *when the random effects model is valid* to use. It is not always valid, and tests are available to assess whether results from the random effects model are trustworthy.[6]

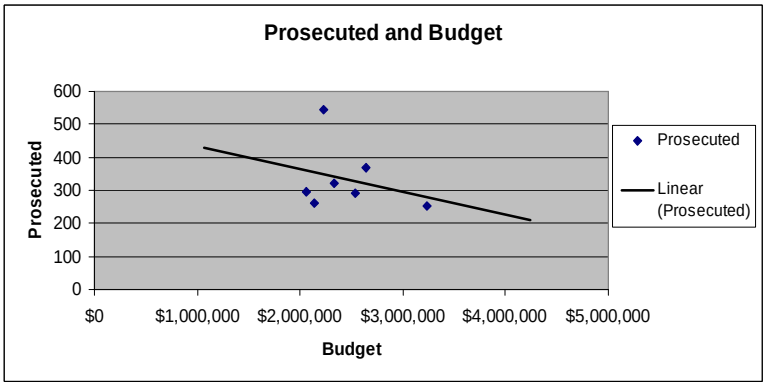
[6] Based on the fixed effect estimator, the inference comes entirely from variation over time within each jurisdiction. The jurisdiction trends are then averaged over the jurisdictions. Based on the random effect estimator, the inference comes from combining the inference based on the above time-series variation and an inference based on cross-sectional variation. Thus, the random effects model uses more information than does the fixed-effect model. The random effect model would always be preferable, except the inference based on the cross-sectional variation may be biased. The Hausman test is used to detect that bias—actually known as consistency in the econometrics literature.

Figure 4.2

Correlations between Selected Outputs and Budget with Imputations







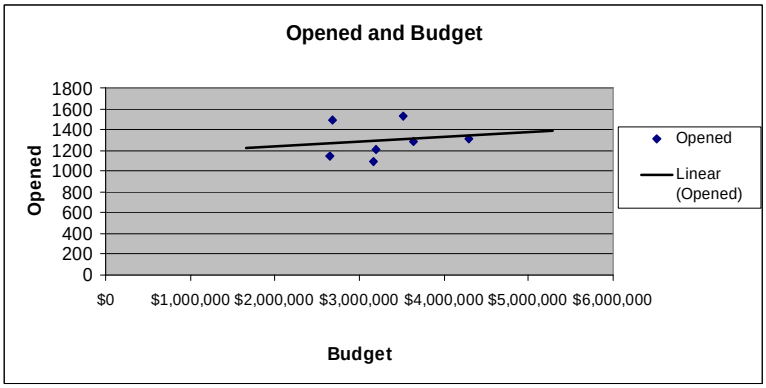
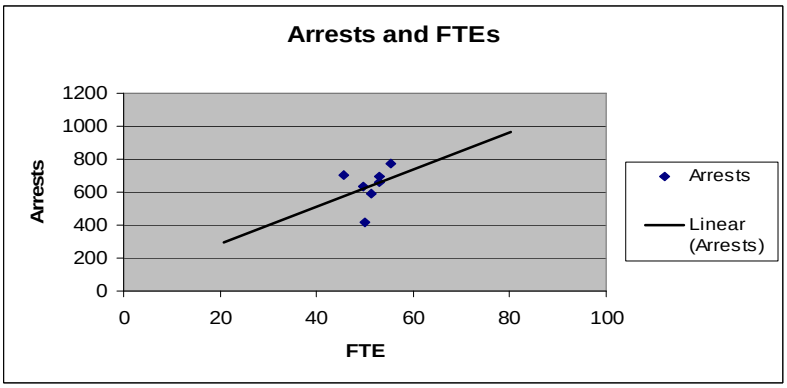
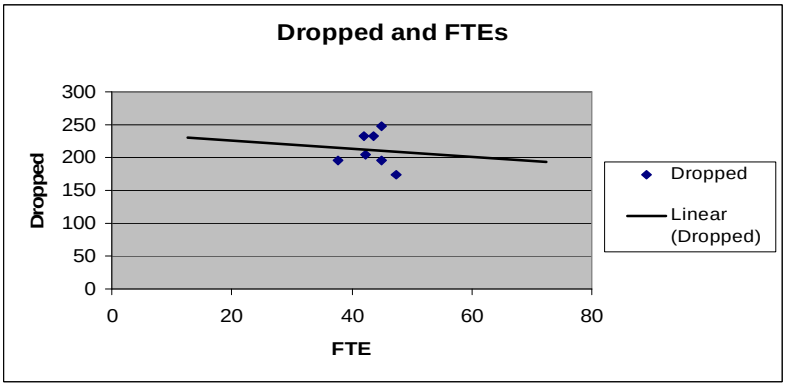
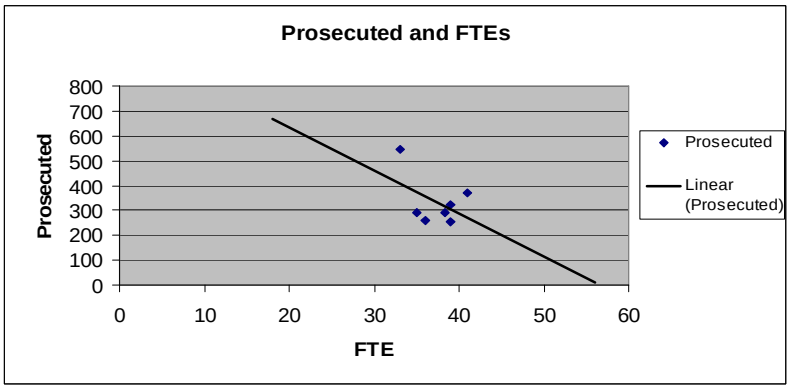


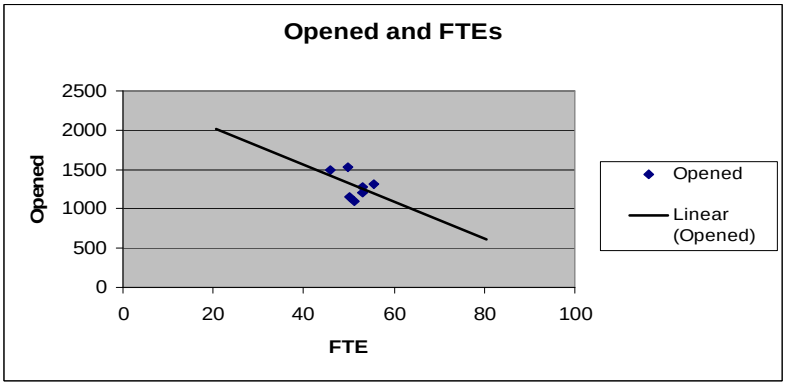
Figure 4.3

Correlations between Selected Outputs and FTEs with Imputations









Though this may appear to be abstract, a way to proceed is to follow these steps:

- Estimate the fixed-effects model and test for the statistical significance of the relationship between the outcome variable and the input variable; and
- Estimate the random effects model and test whether the model is valid. This test is known as a Hausman test. If the random effects model is invalid, then use the results from the fixed effects model; or
- If the random effects model is not invalidated, use the results from the random effects model.

Statistical software for fixed effect and random effect estimation is widely available, but application and interpretation may require a social scientist trained in statistical analysis.

While the random effects model requires the use of specialized software, the fixed effect model can be estimated with a spreadsheet. If Y is the dependent variable (for example, the number of arrests) and if X is the independent variables (for example, the budget), then the fixed-effect model can be estimated by following three steps:

- Compute the mean value for Y and X across the seven years for a specific district. Call these $\overline{Y}$ and $\overline{X}$. Imputed responses do not enter into these calculations.
- Create new variables y and x by subtracting the district mean from the original variable. These new variables equal $y = Y - \overline{Y}$ and $x = X - \overline{X}$.
- Estimate a regression by regressing y onto x. This regression has no constant.
- The regression will not report the correct tests of significance. The tests can be corrected by adjustments done within the spreadsheet. The adjustment is to reduce the degrees of freedom by the number of districts minus 1.

Spreadsheet 4 (Fixed Effects) illustrates this approach. First, the spreadsheet carries all the data forward from spreadsheet 1. This transfer preserves all the missing variables, because imputations should not be used in the analysis. Second, the spreadsheet copies four variables (year, district, budget and arrests) into a lower part of the spreadsheet, and it creates dummy variables for the years. The dummy variables are coded 1 if the observation pertains to the specified year and 0 otherwise. The year 2000 does not get a dummy variable.[7] Missing data, which are identified with yellow highlights, appear in these data.

The third step has several subcomponents. The first subcomponent carries forward only those observations that have no missing data. All others are deleted. See cells A120:J157; these are shaded with light grey. Alternatively, an analyst might delete only those cases that contain missing data for the variables used in the analysis. The second subcomponent computes the means by jurisdictions across years. These means are reported in cells K120:R157, also shaded with grey. The third

[7] Although the regression is estimated without a constant, each jurisdiction is represented by a constant, so including dummy variables for every year would impose perfectly collinearity on the estimation. This is the reason for dropping one of the year dummy variables. Any year could have been selected.

subcomponent subtracts the computed means from the original data. The results are reported in S120:Z157, also shaded with grey. Cells S120:Z157 provide the data used in the regression.

An analyst can estimate a linear regression using Excel's function LINEST. This function has four arguments. The first argument is "known_Y's". For this problem, the known Y's comprise the column of arrests minus the mean arrests for a district: Cells S288:S413. The second argument is "known X's". For this problem, the known X's comprise the cells T288:Z413. The third argument is Const; the argument should be "False". The fourth argument is Stats; the argument should be "True".

Excel documentation explains that this function must be treated as an array formula—a definition that is specific to Excel. Entering an array formula requires some steps. First, enter the formula. Second, identify a range by holding the shift key and highlighting the range coded in pink. Hit the F2 function key, and then while holding the shift and control keys, hit the enter key.

Within the block of cells highlighted with pink, the three cells of greatest interest are highlighted with red. The first is the effect that the budget has on arrests: -0.0000548. The second is the standard error for that estimated effect: 0.0000345. As an estimate of the standard error, this is too small. The problem is that the third cell, which reported the degrees of freedom as 30, is too large because it does not account for estimating the means for each district.

Three other spreadsheets are appended to this workbook. Those additional spreadsheets pertain to Tennessee, Colorado and Illinois. Georgia was unable to provide similar data. The state-specific spreadsheets demonstrate the use of the procedures described in the workbook spreadsheet using real data from Tennessee, Colorado and Illinois. We invite the reader to review findings from those sites.

Extended Analysis

Using a spreadsheet as a statistical computing package is convenient because spreadsheets are widely used but limiting because spreadsheets lack the statistical computing routines provided by specialized software package. We demonstrate the use of one statistical computing package – Stata – for extending the analysis presented above. We show and discuss Stata code for making these computations. Of course users might prefer different statistical computing software.

We elect to begin our Stata sessions with the following lines of code:

```
#delimit ;
clear;
set memory 100m;
set more off;
```

Stata needs a character denoting the end for a line of instructions. By default this is a carriage return. We replace the default with ";" by using the **delimit** command. The next line **clear**s the memo. This line drops all variables that remain from earlier runs. The default memory allocation is sometimes insufficient in State, so we increase it to 100 megabytes using **set memory 100m**. The next command, **set more off**, simply changes the way that Stata reports output to the screen and reduces the need for user intervention. The next line reads data into memory:

```
use file,clear;
```

This assumes that the data have been converted into a Stata file structure. The term *file* denotes the path and file name. Many users will read the data multiple times. If the data have already been read once, however, Stata will refuse to read them again unless the user tells Stata to overwrite the file in memory. The option **clear** provides Stata with that instruction. From this point forward, we will use the Tennessee data, which are provided as part of this report. Variables are:

year	years from 2000 through 2006
mjtf	the identity of the task force
fte	full time equivalent employees
budget	task force budget in thousands
opened	number of cases opened during the year
prosecuted	number of cases prosecuted during the year
dropped	number of cases dropped during the year
arrests	number of arrests during the year

A good starting point is to describe these data. This can be done with the Stata command:

summarize, separator(10);

The command **summarize** will summarize the eight variables in the analysis file. The option **separator(10)** affects the format of the reported table. Using the Tennessee data, we see:

```
. summarize, separator(10);
```

Variable	Obs	Mean	Std. Dev.	Min	Max
year	140	2003	2.007181	2000	2006
mjtf	140	13.8	8.448064	1	30
fte	121	6.38843	7.660637	0	49
budget	140	285.2941	292.1967	78.864	2452.556
opened	119	192.6639	143.5531	20	638
prosecuted	100	127.26	88.99491	3	355
dropped	101	22.92079	35.75044	0	142
arrests	126	124.9048	97.29785	3	500

Here and elsewhere we will report the screenshots from Stata output rather than formatting output in Word tables. There are 140 observations in the Tennessee data file. There are 20 task forces. Each task force reports for seven years: 20x7=140. However, data are sometimes missing. While we always know the budget, we know full time equivalent employees for only 131 task force/year combinations – and for one of those reports there were no full time equivalent employees. Likewise we do not always know the number of opened cases, the number of arrests, and the number of prosecutions or dropped cases.

We might attempt to impute responses for missing variables, but we will not do that for this analysis. First, imputation methodology is complicated, and second, we can draw inferences without performing imputations.

A second useful step is to summarize the data by MJTF. To request that summary, type:

by mjtf:summarize fte budget;

This statement again uses the **summarize** command, but now we limit the summary to two variables: fte and budget. More importantly we have added the prefix: by **mjtf:**. This prefix will provide a summary within each of the thirty task forces. The resulting output is lengthy, so we do not report it in its entirely here. We do report some parts of that output. For the first MJTF, Stata reports:

```
-> mjtf = 1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
fte	7	7.857143	.8997354	6	9
budget	7	343.2717	47.74866	243.948	394.381

Both the full time equivalent employees and the budget are reported for all seven years. That is, there are no missing data. The number of full time equivalents varies between 6 and 9. The budget varies between $244,000 and $394,000. Across the MJTFs in Tennessee, the full time equivalent employees is seldom above 10, but in one MJTF, it is never fewer than 13 and was as many as 49. That exceptional district averaged 32 full time equivalent employees and averaged a budget of nearly $1.2 million over seven years. No other MJTF averaged more than 9 full time equivalents, and none had a budget that averaged more than $510,000. That exceptional district was so different from the other districts that we delete it from the data file. It was district 30 (the district identifications are not numbered from 1 to 20), so we drop it with the statement:

```
drop if MJTF == 30;
```

The descriptive statistics tell another useful story. Consider the results from MJTF 19:

```
-> mjtf = 19
```

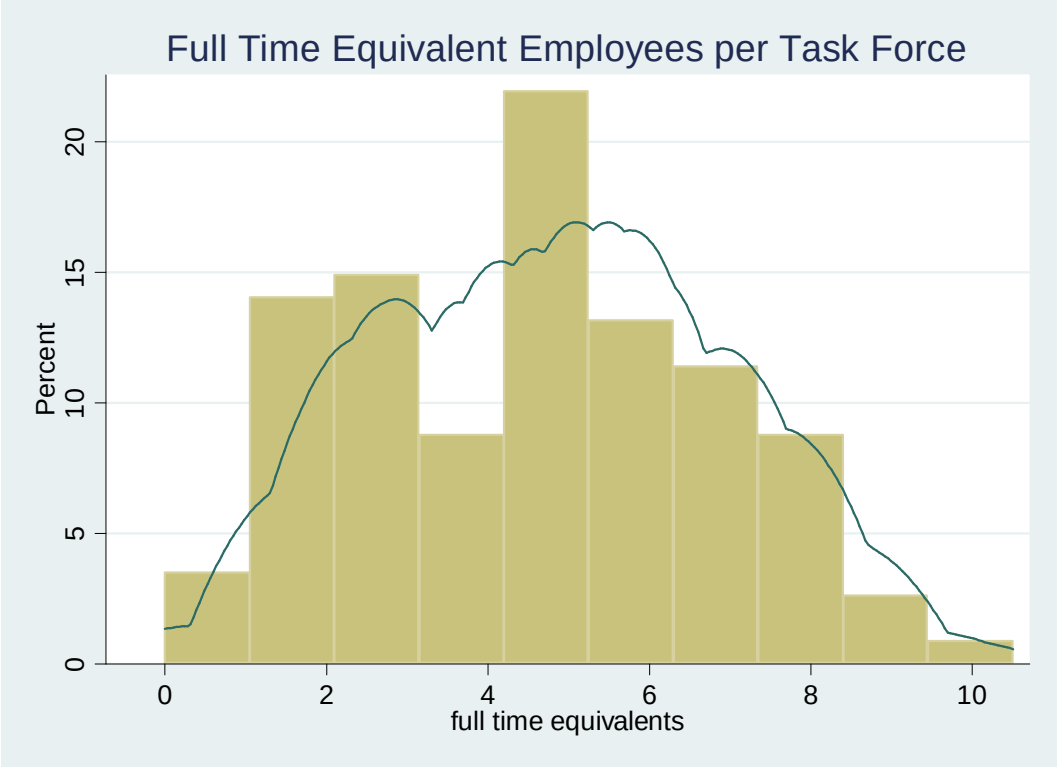
Variable	Obs	Mean	Std. Dev.	Min	Max
fte	7	5	0	5	5
budget	7	146.1006	4.540284	141.598	154.352

Data are never missing, but there is no variation over time in the number of full time equivalents, and there is very little variation in the budget. Nothing can be learned by comparing the variation of full time equivalents and arrests in this district, simple because there is no variation in full time equivalents. Likewise, little can be learned from comparing the variation between arrests and budget, because the variation in budget is immaterial.

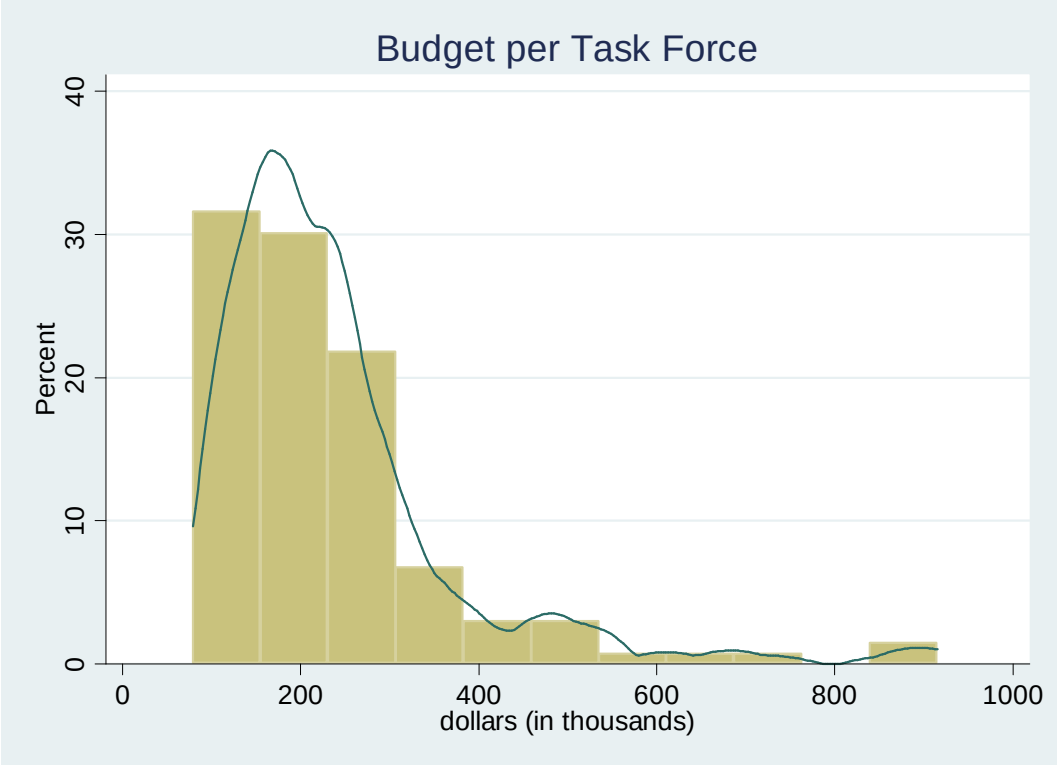
After excluding the one MJTF, we can insect the distribution of full time equivalents across MJTF using a histogram. To request a histogram:

```
histogram fte,kdensity percent xtitle("full time equivalents") title("Full Time Equivalent Employees per Task Force");
```

The basic command is **histogram fte** to request a histogram of full time equivalents. We added a relatively smooth line with the statement **kdensity** . The option **percent** tells Stata to report the histogram as percentages. The **xtitle** and **title** statements add titles. The figure appears below.



After deleting the one MJTF, the remaining MJTFs look similar at least with respect to the number of full time equivalent statements. We also examine the distribution for budgets measured in thousands. The statement is the same, except that we substitute budget for fte. The resulting histogram is:



We might be concerned that a few MJTF/ year combinations have dollar amounts that are considerably larger than other MJTF/year combinations, but we will ignore that problem in what follows.

At this point, one might turn to analysis using statistical models that are appropriate for analyzing panel data. State requires a user to identify cross-sections (the MJTFs in these data) and individual observations (the time-series in these data). The statement is simple but must appear in the file:

xtset mjtf year;

Now the user can estimate models that are appropriate for analyzing panel data. First estimate a fixed effect model using the statement:

xtreg arrests fte year,fe;
estimates store FE;

The statement **xtreg** tells Stata to estimate a regression. The variable to be explained (**arrests**) is the first variable is a list; the explanatory variables (**fte** and **year**) are the rest of the variables in the list. The year variable appears so that we can test for a trend that has nothing to do with the number of full time equivalent employees. Note that this statement uses the option **fe**, which follows a comma. This tells Stata to estimate a fixed effect model. We will discuss the **estimates** statement later. The output is:

```
Fixed-effects (within) regression               Number of obs      =       105
Group variable: mjtf                            Number of groups   =        19

R-sq:  within  = 0.0396                         Obs per group: min =         1
       between = 0.4208                                        avg =       5.5
       overall = 0.2599                                        max =         7

                                                F(2,84)            =      1.73
corr(u_i, Xb)  = 0.3945                         Prob > F           =    0.1834

------------------------------------------------------------------------------
     arrests |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fte |    8.238908   6.397876     1.29   0.201    -4.483969    20.96179
        year |    2.065148   2.522229     0.82   0.415    -2.950582    7.080878
       _cons |   -4066.265   5041.383    -0.81   0.422    -14091.61    5959.078
-------------+----------------------------------------------------------------
     sigma_u |   106.14104
     sigma_e |   47.375221
         rho |   .83387444   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(18, 84) =    10.62                 Prob > F = 0.0000
```

Because we have eliminated one MJTF from the analysis, there are only 19 remaining MJTFs in the analysis. Stata calls these *groups*. If every MJTF had data for each of the seven years, there would be 19x7=133 observations, but there are only 105 observations. This is because missing data reduce the available data to 5.5 observations per MJTF. The relationship between full time equivalent employees and arrests is not statistically significant. (It has a P-value of 0.201. Typically an analyst would require a P-value to be smaller than 0.10 or perhaps 0.05.) On the other hand, there is an apparent time-trend toward more arrests. Plausibly this is an accumulating effect from task force presence although that explanation is speculative.

So, is there no relationship between the number of full time equivalent employees and arrests? Possibly, but there is a problem. A fixed effect model only exploits the information within each task force; it discards all information from across task forces. The problem is that we know from the earlier analysis using the by *mjtf:summarize fte budget; c*ommand that fte's are constant or nearly constant across several task forces. Such task forces can contribute nothing to the estimates based on the fixed-effect model. The statement xtsum provides some insight into the issue of within MJTF variance and across MJTF variance. Request:

xtsum arrests fte;

This produces a table:

Variable		Mean	Std. Dev.	Min	Max	Observations
arrests	overall	117.479	88.78453	3	500	N = 119
	between		71.85815	11.71429	255.2	n = 19
	within		54.97922	-21.72101	362.279	T-bar = 6.26316
fte	overall	4.807018	2.168263	0	10.5	N = 114
	between		2.145397	2	9	n = 19
	within		.8202299	2.66416	7.66416	T-bar = 6

This table shows that the total variation is arrests can be described by a standard deviation of 88.8. Within MJTFs, the standard deviation is 55.0, but across MJTFs it is 71.9. There is considerable variation across MJTFs, and the fixed effect estimator does not exploit that. A random effect estimator does exploit that additional variation. Use the command:

xtreg arrests fte year,re;
estimates store RE;

This command is the same as the earlier command except that the fe has been replaced with a re. The output is:

```
Random-effects GLS regression                   Number of obs      =       105
Group variable: mjtf                            Number of groups   =        19

R-sq:  within  = 0.0372                         Obs per group: min =         1
       between = 0.3910                                        avg =       5.5
       overall = 0.2647                                        max =         7

Random effects u_i ~ Gaussian                   Wald chi2(2)       =      9.61
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0082
```

arrests	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
fte	15.10009	5.446569	2.77	0.006	4.425016	25.77517
year	1.439133	2.499151	0.58	0.565	-3.459114	6.337379
_cons	-2826.099	4998.798	-0.57	0.572	-12623.56	6971.365
sigma_u	83.531516					
sigma_e	47.375221					
rho	.75662225	(fraction of variance due to u_i)				

These results are more reasonable given an expectation that MJTF resources should account for something of value. In fact the number of arrests appears to increase by about 15.1 per year per FTE added to a task force. The effect is statistically significant but it is measured with considerable imprecision. A 95% confidence intervals is 4.4 to 25.8 additional arrests per year per FTE.

Why would an analyst ever use the fixed effect model if it uses less information than does the random effect model? That question is the sort of problem that occupies statisticians, but to encapsulate: The random effect model rests on some strong assumption that, if not true, will lead to biased estimates. The fixed-effect model is often used to guard against that bias.

How can we tell whether or not the random effect model is valid? Regrettably the answer is not straightforward. Analyst sometimes use a Hausman test, which is available in Stata. Anticipating the use of a Hausman test, we introduced but did not explain the lines *estimates store FE* and estimates store *RE*. To apply the Hausman test, run the following lines:

```
xtreg arrests fte year,fe;
estimates store FE;
xtreg arrests fte year,re;
estimates store RE;
hausman FE RE,sigmamore;
```

The command **estimates store FE** and **estimates store RE** save results for use by the Hausman test. The last line is new. It invokes a Hausman test. If the random effect model is valid, the parameter estimates from the two models will not be significantly different. The results of the Hausman test are:

```
                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       FE           RE         Difference          S.E.
-------------+------------------------------------------------------------
         fte |  8.238908     15.10009       -6.861186       3.609686
        year |  2.065148     1.439133        .6260153       .6242975
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =         6.90
                Prob>chi2 =       0.0318
```

Unfortunately the parameter associated with the full time equivalent variable is significantly different across the two estimators. We know this by dividing -6.861 by 3.610 to get a z-score of 1.90, which will be statistically significant according to a table of standard normal variables. This leaves us in an uncertain position, but the evidence seems to indicate that arrests increase with full time equivalent employees. The best estimate is that the increase equals 8.2 arrests per year per full time equivalent in the fixed-effect model and 15.1 arrests per year per full time equivalent in the random effect model.

We can estimate the same regressions after substituting budget for FTEs. The code is:

```
xtreg arrests budget year,fe;
estimates store FE;
xtreg arrests budget year,re;
estimates store RE;
hausman FE RE,sigmamore;
```

The results follow:

```
Fixed-effects (within) regression               Number of obs      =       119
Group variable: mjtf                            Number of groups   =        19

R-sq:  within  = 0.1163                         Obs per group: min =         3
       between = 0.2520                                        avg =       6.3
       overall = 0.1873                                        max =         7

                                                F(2,98)            =      6.45
corr(u_i, Xb)  = 0.1512                         Prob > F           =    0.0023
```

arrests	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
budget	.1971487	.0698045	2.82	0.006	.0586239	.3356734
year	4.136974	2.655187	1.56	0.122	-1.132159	9.406107
_cons	-8215.254	5315.16	-1.55	0.125	-18763.02	2332.508
sigma_u	63.008884					
sigma_e	56.713665					
rho	.55243675	(fraction of variance due to u_i)				

```
F test that all u_i=0:     F(18, 98) =     7.56                 Prob > F = 0.0000
```

```
Random-effects GLS regression                   Number of obs      =       119
Group variable: mjtf                            Number of groups   =        19

R-sq:  within  = 0.1158                         Obs per group: min =         3
       between = 0.2536                                        avg =       6.3
       overall = 0.1891                                        max =         7

Random effects u_i ~ Gaussian                   Wald chi2(2)       =     18.27
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0001
```

arrests	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
budget	.2201054	.060246	3.65	0.000	.1020255	.3381853
year	3.861643	2.625125	1.47	0.141	-1.283509	9.006794
_cons	-7669.066	5255.767	-1.46	0.145	-17970.18	2632.049
sigma_u	60.495378					
sigma_e	56.713665					
rho	.53223114	(fraction of variance due to u_i)				

	(b) FE	(B) RE	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
budget	.1971487	.2201054	-.0229567	.0346062
year	4.136974	3.861643	.2753316	.3048018

```
                b = consistent under Ho and Ha; obtained from xtreg
 B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                        =        0.85
              Prob>chi2 =      0.6538
```

The fixed-effect model now indicates that $1000 of expenditures increases arrests by about 0.20 arrests. The random-effects model indicates that $1000 of expenditures increases arrests by about 0.22. The Hausman test provides no evidence that would cause us to dismiss the random effect model

as being inappropriate, so we might conclude that an additional $1000 of expenditures increases arrests by 0.10 to 0.34 arrests based on a 95 percent confidence interval.

Statisticians often worry about other assumptions that could lead to misleading inferences based on statistical analysis. Dealing with these assumptions requires advanced knowledge of statistical models applicable to panel data. We cannot cover those topics here, but we will consider one issue. Panel data almost always include data that are correlated across time, which is the reason why statisticians use fixed-effect and random-effect models. But sometimes that correlations take on a specific form including autocorrelation. Dealing with autocorrelation is straightforward in balanced panels. Balanced panels have the same number of data points per MJTF. Our data are not balanced principally because of missing data patterns.

To introduce first-order autocorrelation (a common assumption about autocorrelation) into the model, use the statements:

```
xtreg arrests fte year,pa corr(ar 1) vce(robust);
matrix list e(R)
xtreg arrests fte year,pa corr(ar 1) vce(robust);
matrix list e(R)
```

These statements introduce some new terms. First, **pa** substitute for **fe** and **re**. The option **pa** stands for *population averaged* and is in fact a very different estimator than those considered earlier. The term corr(ar 1) presumes that there exists first-order serial correlation and requires the model to estimate a parameter describing that first-order serial correlation and correct for its presence. The vce(robust) term corrects for other problems including clustering; those problems motivated the use of the fixed-effect and random-effect models in the first place. The command matrix list e(R) provides an estimate of the correlation across observations.

Output from these two models is reported below, first when the explanatory variable is full time equivalents and second when the explanatory variable is budget. The notes appearing with the output explain that some MJTFs had to be eliminated from the data because the data were *unbalanced.* In this context, unbalanced means that the data skipped years, so Stata could not compute the autocorrelation coefficient. Consequently, these results do not just rest on a different statistical model; they also rest on different data. Note that only 91 observations from 13 MJTFs entered this analysis.

```
note:  observations not equally spaced
       modal spacing is delta year = 1 unit
       3 groups omitted from estimation
note:  some groups have fewer than 2 observations
       not possible to estimate correlations for those groups
       3 groups omitted from estimation


Iteration 1: tolerance = .53818221
Iteration 2: tolerance = .0033746
Iteration 3: tolerance = .00005224
Iteration 4: tolerance = 8.174e-07

GEE population-averaged model                 Number of obs      =         91
Group and time vars:              mjtf year   Number of groups   =         13
Link:                               identity  Obs per group: min =          7
Family:                             Gaussian                avg =        7.0
Correlation:                           AR(1)                max =          7
                                              Wald chi2(2)       =      17.70
Scale parameter:                    4303.807  Prob > chi2        =     0.0001
```

arrests	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
fte	19.72863	4.880275	4.04	0.000	10.16347	29.29379
year	2.983045	4.1546	0.72	0.473	-5.159822	11.12591
_cons	-5972.72	8318.953	-0.72	0.473	-22277.57	10332.13

```
.                matrix list e(R);

symmetric e(R)[7,7]
           c1         c2         c3         c4         c5         c6         c7
r1          1
r2  .60652765          1
r3  .36787579  .60652765          1
r4  .22312684  .36787579  .60652765          1
r5   .1353326  .22312684  .36787579  .60652765          1
r6  .08208296   .1353326  .22312684  .36787579  .60652765          1
r7  .04978559  .08208296   .1353326  .22312684  .36787579  .60652765          1

note:  observations not equally spaced
       modal spacing is delta year = 1 unit
       6 groups omitted from estimation

Iteration 1: tolerance = .6283854
Iteration 2: tolerance = .02921007
Iteration 3: tolerance = .00179446
Iteration 4: tolerance = .0001103
Iteration 5: tolerance = 6.777e-06
Iteration 6: tolerance = 4.171e-07

GEE population-averaged model                 Number of obs      =         91
Group and time vars:              mjtf year   Number of groups   =         13
Link:                               identity  Obs per group: min =          7
Family:                             Gaussian                avg =        7.0
Correlation:                           AR(1)                max =          7
                                              Wald chi2(2)       =       6.74
Scale parameter:                    5657.55   Prob > chi2        =     0.0344

                                  (Std. Err. adjusted for clustering on mjtf)
```

arrests	Coef.	Semi-robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
budget	.1152991	.0707715	1.63	0.103	-.0234105	.2540088
year	4.56819	2.857893	1.60	0.110	-1.033177	10.16956
_cons	-9069.581	5712.346	-1.59	0.112	-20265.57	2126.411

```
.                matrix list e(R);

symmetric e(R)[7,7]
           c1         c2         c3         c4         c5         c6         c7
r1          1
r2  .67751601          1
r3  .45902794  .67751601          1
r4  .31099878  .45902794  .67751601          1
r5  .21070665  .31099878  .45902794  .67751601          1
r6  .14275713  .21070665  .31099878  .45902794  .67751601          1
r7  .09672024  .14275713  .21070665  .31099878  .45902794  .67751601          1
```

From this analysis, it seems that the addition of a full time equivalent employee increases the number of arrest from between 10.2 and 29.3 – the best estimate is nearly 20 arrests per full time equivalent. The autocorrelation structure appears to be material. The best estimate is that an additional $1000 expenditure increases the number of arrests by about 0.12. The effect is not quite significant at 0.10 in a two tailed test of significance. One might argue for using a one-tailed test of significance given that the direction of the effect seems predictable. Then the effect would be significant at 0.10 but not quite at 0.05.

One might employ still more sophisticated (and perhaps more compelling) statistical models, but there are two caveats. Many more compelling models require longer time-series than are provided by these data, so they may not be useful for these data, although they may be useful for future data. Second, the quality of these data – which undoubtedly include data that suffer from measurement error – may preclude reaching strong conclusions.

This last caveat aside, we close this section with a summary analysis of the data. Table X presents results of how full time equivalents and budget separately affect four outcomes:

- arrests
- opened cases
- prosecuted cases
- proportion prosecuted

Tables 4.2 to 4.4 report results for three models: fixed effect, random effect and population averaged with first-order auto correlation. These are the three models described above.

The tables are organized according to state: Tennessee, Illinois and Colorado. Georgia could not provide comparable data. For each state, the tables are divided into the two input measures: full time equivalent employees and MJTF budget. For each state/input the tables are divided into four outcome measures: arrests, cases opened, prosecutions and proportion of cases that were dropped without prosecution. The columns report the parameter estimate for the regression of the outcome on FTE/budget and year. A P-value for a two-tailed test of significance appears below the parameter estimate. Readers who prefer a one-tailed test based on a prior belief about the direction of causation from inputs and outputs can divide the P-values by two.

Evaluators will often impose a penalty for when making *multiple comparisons*. When a table reports parameter estimates for multiple analyses (Ninety-six tests in the tables), some would be statistically significant by chance, so researchers adjust the P-values to accommodate that bias. We have not done that.

Readers can draw their own conclusions, but to summarize our conclusions: In Tennessee, the evidence is compelling that the number of cases, arrests and prosecutions increases with the number of MJTF full time equivalent staff. The evidence is less compelling – but still appreciable – that cases, arrests and prosecutions increase with budgets. There is no evidence that the proportion of dropped cases increases or falls with either budget or full time equivalent employees.

The conclusions are similar in Illinois. However, in Illinois, the fixed-effect model seems to provide no evidence that either full time equivalent employees or task force budgets increase outputs. This may be because the inputs do not vary much over time within each task force, but the evidence is to the contrary as the number of full time equivalent employees was constant in only one task force. The random effect model does point toward improved outcomes based on expanded resources. Of course, this may be because the random effect model is misspecified. This is possible because we see that the Hausman test calls into question the bias/consistency of the random effect model.

Results from Colorado are very similar to the results from Tennessee provided we examine how full time equivalents affect outcomes. The Hausman test suggests that the findings from the random effect model are inconsistent, but nevertheless, there is little substantive difference between the findings from the fixed-effect and random-effect models. In contrast there is no evidence that budgets have a large impact on outcomes.

Table 4.2

Results from Estimating the Fixed-Effect and Random-Effect Models in Tennessee

	Parameter Estimates and P-Values for Two-Tailed Tests			
	Fixed-Effect Model	**Random-Effect Model**	**P-Value for the Hausman test**	**Population Averaged Model**
Tennessee				
FTE				
Arrests	8.24	15.10	0.029	19.73
	0.201	0.006		0.000
Opened	15.81	22.72	0.256	13.79
	0.029	0.000		0.052
Prosecuted	14.05	18.91	0.077	14.56
	0.020	0.000		0.011
Dropped	-0.01	-0.01	0.147	0.00
	0.346	0.273		0.759
Budget				
Arrests	0.20	0.22	0.654	0.12
	0.000	0.000		0.103
Opened	0.30	0.31	0.333	0.15
	0.000	0.000		0.162
Prosecuted	0.17	0.21	0.448	0.18
	0.059	0.017		0.193
Dropped	0.00	0.00	0.408	0.00
	0.614	0.446		0.002

Table 4.3

Results from Estimating the Fixed-Effect and Random-Effect Models in Illinois

	Parameter Estimates and P-Values for Two-Tailed Tests			
	Fixed-Effect Model	Random-Effect Model	P-Value for the Hausman test	Population Averaged Model
Illinois				
FTE				
Arrests	2.13	9.71	0.006	8.03
	0.631	0.004		0.002
Opened	-0.71	9.08	0.015	13.07
	0.904	0.020		0.003
Prosecuted	2.31	11.82	0.000	12.40
	0.858	0.000		0.000
Dropped	0.00	0.00	0.270	0.01
	0.997	0.733		0.454
Budget				
Arrests	0.03	0.06	0.475	0.04
	0.581	0.147		0.451
Opened	-0.04	0.04	0.061	0.08
	0.603	0.393		0.143
Prosecuted	-0.12	-0.01	0.453	0.00
	0.147	0.858		0.966
Dropped	0.00	0.00	0.467	0.00
	0.941	0.320		0.971

Table 4.4

Results from Estimating the Fixed-Effect and Random-Effect Models in Colorado

	Parameter Estimates and P-Values for Two-Tailed Tests			
	Fixed-Effect Model	Random-Effect Model	P-Value for the Hausman test	Population Averaged Model
Colorado				
FTE				
Arrests	12.59	14.73	0.046	9.68
	0.000	0.000		0.102
Opened	10.29	12.49	0.000	6.49
	0.003	0.000		0.174
Prosecuted	8.74	10.39	0.004	11.81
	0.002	0.000		0.053
Dropped	0.02	-0.01	0.000	0.01
	0.663	0.685		0.001
Budget				
Arrests	0.03	0.04	0.494	0.04
	0.598	0.500		0.556
Opened	0.04	0.05	0.376	-0.01
	0.473	0.382		0.885
Prosecuted	0.219	0.027	0.495	0.00
	0.68	0.58		0.950
Dropped	0.000	0.000	0.235	0.00
	0.596	0.771		0.678

One interpretation is that MJTFs had their expected effect on outputs. They seem to have caused an increase in cases, arrests and prosecutions. There is no evidence that MJTFs lead to a higher proportion of prosecutions per case. We might have expected this outcome without statistical analysis, although without some formal analysis, we could not have quantified that effect. A reader must be cautious. Data assembly was complicated and costly, and likely these data include an unquantifiable amount of measurement error.

Conclusion

The goal of this guide is to provide a template that BJA and state agencies might follow to develop performance monitoring systems. A well-developed performance monitoring system will allow state and federal agencies to hold MJTFs accountable for how grant funds are used and to provide solid evidence of their activities and effectiveness. Programs unable to provide this kind of evidence are at risk of having their resources cut or diverted to other programs that can better demonstrate success.

For states without existing performance monitoring systems, the guide provides a thoughtful approach to building a system that supports public decision making. By starting with a logic model that represents the goals, resources, activities, outputs, outcomes, and impacts of the MJTFs in the state and building a performance measurement component that is developed from the model data reporting template, the SAA can be confident that the resulting performance monitoring system will provide sufficient data to monitor and measure performance.

SAAs with existing performance monitoring systems can also benefit from this guide. Existing systems can always benefit from review, and this guide supports that process by promoting the use of logic models to assess MJTF goals, resources, activities, outputs, outcomes and impacts to ensure performance measures are appropriate for tracking and measuring MJTF performance. The model data reporting template may also provide additional measures not previously collected by the SAA or a different way to collect information previously of interest to the SAA. The model analysis plan can help SAAs identify new ways of analyzing data that would be more useful to public decision makers.

Whether the state is establishing a new system or improving upon an existing performance monitoring system, the guidance on advanced statistical analysis may assist the SAA in conducting their own analysis or to better instruct consultants trained in social science statistics.

To summarize, the guide has broad application to multiple audiences that may or may not have experience with performance monitoring. Promoting accountability among MJTFs is critical to ensuring continued support for this important first line of defense in reducing the availability of illegal drugs on American streets.

Appendix 1: References

Baltagi, B. (2005) Econometric Analysis of Panel Data, John Wiley & Sons, 3rd Edition.

Greene, W. (2008) Econometric Analysis, Prentice Hall, 6th Edition.

Heckman, J. and Leamer, E. (2007) Handbook of Econometrics Volume 6B, North-Holland Press, Amsterdam.

King, G., Honaker, J., Joseph, A. and Scheve, K. (2001) Analyzing Incomplete Political Science Data: An Algorithm for Multiple Imputations, American Political Science Review, Vol. 95(1).

Morgan, S. and Winship, C. (2007) Counterfactuals and Causal Inference: Methods and Principals for Social Research. Cambridge University Press.

Rosenbaum, P. (2002) Observational Studies, 2nd ed. Springer-Verlag.

Schafer, J. (1997) Analysis of Incomplete Multivariate Data, Chapman and Hall.

Wooldridge, J. (2002) Econometric Analysis of Cross Section and Panel Data. The MIT Press. Cambridge, MA.

Appendix 2: Performance Measurement Template

MJTF Performance Monitoring Tool

Definitions

Partner Agencies: Those agencies who participate as members of the Task Force in Memoranda of Understanding or Interagency Agreements and/or have assigned staff to the Task Force.

Average FTEs: Comprises all sworn MJTF staff (e.g., police officers, sheriff's deputies, investigators) and non-sworn MJTF staff who provide substantive support to field operations and investigations (e.g., crime analysts, intelligence analysts). Administrative assistants and clerical staff are excluded. It includes Task Force staff who are funded by BJA, other grant, through agency budgets, or serve on the Task Force as matching or in-kind contributions. This should be reported in terms of full-time equivalents (i.e., 1 full-time person = 1 FTE; 1 full-time person and 1 half-time person = 1.5 FTEs). Since staffing can fluctuate over time, please update any changes to the average FTEs each reporting period.

Task Force Annual Budget: This should include a breakdown of all funds and contributions made to the Task Force for that fiscal year. Therefore, federal funds include BJA/JAG funding as well as any other federal funds used to support Task Force operations. State, local, and/or tribal funds should include any contributions made by any of these sources. Task Force revenue should include any planned revenue from forfeiture proceeds, fines, or any other sources of income used to support the Task Force. In-kind contributions (personnel, cruiser, Task Force office space, etc.) are also captured, which should include personnel loaned to the Task Force from a law enforcement agency partner.

Total Arrests: Total number of arrests conducted by the Task Force (one body equals one arrest). Information should be reported by the most serious charge that applies at the time, using the following broad categories: felony/misdemeanor, and possession/delivery of controlled substance (i.e., all substances that are federally regulated). Therefore, each arrest will fall in either felony possession of controlled substance, felony delivery of controlled substance and so on. Although statutory definitions and terminology will differ from state to state, the delivery category should include any charge other than simple possession, including: distribution, manufacture, conspiracy and possession with intent to deliver. Similarly, the controlled substance category should include all controlled substances, including cannabis.

Gang: An ongoing organization, association, or group of three or more persons that have a common interest and/or activity characterized by the commission of or involvement in a pattern of criminal or delinquent conduct. (NCIC).

Drug Seizure Amounts and Units: The total amount of each type of drug (marijuana, cocaine, crack, heroin/opiates, methamphetamine, and other drugs) seized by the Task Force. This is a sum of the amount of each drug seized reported in each Task Force's quarterly report. For each drug type listed, please record based on the units provided. If you have drugs reported in other metrics, use the following to convert them: 1kg=1000g; 1kg=2.2lbs; 1kg=35.2oz; 1lb=16oz; and 1lb=453.6g. For Task Forces seizing other types of drugs (e.g., ecstasy, LSD, prescription drugs), space is included to identify four additional drugs for which the greatest amount was seized, and to provide the amounts and units of measurement for each.

National Integrated Ballistic Information Network (NIBIN): An intelligence tool provided by the ATF that deploys Integrated Ballistic Identification System (IBIS) equipment into federal, state and local law

enforcement agencies for their use in imaging and comparing crime gun evidence. IBIS equipment is used to compare firearm related evidence to identify possible links to a crime investigation. (ATF)

Disrupted: Impeding the normal and effective operation of the targeted organization, as indicated by changes in organizational leadership and/or changes in methods of operation, including, for example, financing, trafficking patterns, communications, or drug production. (HIDTA)

Dismantled: Destruction of the organization's leadership, financial base, and supply network such that the organization is incapable of operating and/or reconstituting itself. (HIDTA)

Drug Trafficking Organization (DTO): An organization consisting of five or more persons that (1) has a clearly defined hierarchy or chain of command and (2) whose principal activity is to generate income or acquire assets through continuing illegal drug production, manufacturing, importation, or distribution activities. (Please note: a hierarchy or chain of command can be as minimal as one recognized leader for all the participants. At least one recognized leader must be present to constitute an organization). (ONDCP)

Total Length of Prison Sentence: The mid-point between the minimum and maximum sentence for each defendant. For example, a sentence of five to fifteen years would be ten years. For defendants who receive a life sentence, assume life equals 25 years.

Reporting Cycle: *[Note: This will be pre-filled for respondents.]*	
Start Date:	End Date:

A. Task Force Characteristics
(Please complete the following items or update as necessary.)

[Note: This will be reported initially and updated as needed, but at a minimum annually. This applies to sections A-C.]

A.1 Date completed: _____

A.2 Task Force name: _____

A.3 Task Force jurisdiction (*list state and/or all counties/judicial districts covered by the Task Force*):

_____

_____

_____

_____

_____

_____

_____

A.4 Partnering agencies (*please list the names of each partner agency*):

a. Federal law enforcement agencies:

_____

_____

b. State law enforcement agencies:

_____

_____

c. Local law enforcement agencies

_____

_____

_____

_____

_____

_____

d. Tribal law enforcement agencies:

_____

_____

_____

e. Other non-law enforcement agencies (e.g., District Attorney's Office):

_____

_____

_____

_____

A5. a. Please rank the priority that your Task Force has given to interdicting the following drug types during the reporting period. (*1=highest priority, the others in descending order*):

Rank:

_____ Marijuana

_____ Crack Cocaine

_____ Powder Cocaine

_____ Heroin/Opiates

_____ Methamphetamine

b. Please rank the priority that your Task Force has given to the following issues during the reporting period (*1=highest priority, the others in descending order*):

Rank:

_____ Disrupting organized crime drug trafficking networks

_____ Disrupting street-level drug dealers

_____ Highway interdiction

_____ Public education/outreach

_____ Violent crime

_____ Other *(Please list:)* _____

_____

A.6 Full-time equivalent (FTE) Task Force staff assigned to the Task Force:

a. Average number of federal law enforcement ..... _____

b. Average number of state law enforcement ..... _____

c. Average number of local law enforcement..... _____

d. Average number of tribal law enforcement ..... _____

e. Average number of other non-law enforcement (e.g., prosecutor, crime analyst) ..... _____

B. Task Force Annual Budget Information
(Please complete the following items or update as necessary.)

B1. Please report a breakdown of your Task Force's annual budget for the current fiscal year, separating personnel from other costs.

		Federal funds	State, local and/or tribal funds (including local match)	Task Force revenue (e.g., fines, forfeitures)
a.	**Personnel**	$_____	$_____	$_____
b.	**Non-personnel** *(e.g. equipment, supplies, office space, technology, and other non-personnel costs)*	$_____	$_____	$_____

B.2 Specify any in-kind contributions expected or made during the current fiscal year *(check all that apply)*.

- ❑ Personnel
- ❑ Technology
- ❑ Equipment
- ❑ Supplies
- ❑ Other *(please specify:)* _____

C. Task Force Structure and Operations

(Please complete the following items for the current fiscal year or update information as necessary.)

C.1 The Task Force has *(check all that apply)*:

- ❑ A fully executed (i.e., current and signed) written Interagency Agreement or Memorandum of Understanding.

- ❑ A Task Force Policy Board (i.e., official board of law enforcement executives that oversee formulation and adherence to policies and operating procedures).

  If you have a ***Task Force Policy Board***, how often does the policy board meet? *(Check all that apply.)*

  - ❑ Annually
  - ❑ Quarterly
  - ❑ Monthly
  - ❑ As needed

C.2 The Task Force has written protocols that include any of the following *(check all that apply)*:

- ❑ Task Force chain of command
- ❑ Task Force mission statement
- ❑ Task Force goals and objectives
- ❑ Policies and procedures

  If you have written ***policies and procedures***, do they include any of the following? *(Check all that apply.)*

  - ❑ Standard Operating Procedures for agents
  - ❑ A policy for targeting and approving cases
  - ❑ A policy for information referrals
  - ❑ Other

C3. How has this documentation been communicated to Task Force members? *(check all that apply)*

- ❑ Share with Task Force personnel
- ❑ Task Force training
- ❑ Require Task Force personnel to sign copy
- ❑ Share with partner agencies
- ❑ Require partner agencies to sign
- ❑ None of the above

C.4 The Task Force has a physical space designated to the Task Force. *(Check if yes.)* ..... ❑

C.5 The Task Force has a prosecutor assigned to work with the Task Force. *(Check if yes.)* ..... ❑

C.6 The Task Force has access to the following criminal intelligence support systems *(check all that apply)*:

- ❑ State level (e.g., Pennsylvania Justice Network)
- ❑ Regional level (e.g., RISS)
- ❑ National (e.g., Homeland Security Information Network)
- ❑ Other

C.7 The Task Force utilizes the results from its local annual threat assessment to set priorities. *(Check if yes.)*..... ❑

C.8 The Task Force has a deconfliction system (i.e., a system to identify whether a Task Force investigation conflicts with another agency's investigation). *(Check if yes.)* ..... ❑

If the Task Force has a ***deconfliction system***, indicate the number of times during this reporting period the deconfliction system identified that a Task Force investigation conflicted with another agency's investigation:..... _____

D. Task Force Activities

(Please complete the following items for the current reporting period.)

[Note: Sections D and E would be reported periodically.]

Information and Intelligence Sharing

D.1 Number of submissions to the following databases:

a. State intelligence database (e.g., Pennsylvania Justice Network ..... _____

b. Regional intelligence database (e.g., Citizen and Law Enforcement Analysis and Reporting–CLEAR) ..... _____

c. National intelligence database (e.g., Homeland Security Information Network) ..... _____

D.2 Number of counterterrorism referrals to other agencies (e.g., Joint Terrorism Task Force) ..... _____

D.3 Number of drug-endangered child referrals to appropriate child protective services agencies ..... _____

D.4 Number of tips/leads referred to other agencies (other than counterterrorism or child protective) ..... _____

Education and Awareness

D.5 Number of specialized training programs received by Task Force personnel (e.g., investigation techniques, undercover operations, clandestine lab identification, drug interdiction)..... _____

a. Number of participants trained..... _____

D.6 Number of community prevention/awareness trainings conducted (e.g., clandestine lab identification)..... _____

a. Number of participants trained..... _____

D.7 Number of training programs provided by Task Force personnel to other law enforcement personnel..... _____

a. Number of participants trained..... _____

Investigation Activity Initiated by Task Force Personnel

D.8 Investigation activity:

a. Number of new investigations opened..... _____

b. Number of investigations closed..... _____

c. Number of investigations dropped (i.e., not resulting in arrest)..... _____

d. Number of investigations pending ..... _____

D.9 Number of search warrants served:

a. Federal..... _____

b. State..... _____

D.10a Number of individuals arrested for:

a. Felony possession of controlled substance..... _____

b. Felony delivery of controlled substance..... _____

c. Misdemeanor possession of controlled substance..... _____

d. Misdemeanor delivery of controlled substance..... _____

D.10b Of the individuals arrested (from D10 above), how many are documented criminal gang members?..... _____

Of those who are documented criminal gang members, how many were arrested for:

a. Felony possession of controlled substance..... _____

b. Felony delivery of controlled substance..... _____

c. Misdemeanor possession of controlled substance..... _____

d. Misdemeanor delivery of controlled substance..... _____

Seizures and Forfeitures Initiated by Task Force Personnel

Drug Seizures

D.11 Drug seizure amount and units by drug: *(For each type of drug, please indicate the total amount of drugs seized by your Task Force during the previous fiscal year. If applicable, please list four "other" drugs for which the greatest amount was seized each year, and provide the amounts and units of measurement for each.)*

Drug	Amount	Unit
Marijuana	Amount = _____	pounds
Powder Cocaine	Amount = _____	ounces
Crack Cocaine	Amount = _____	grams
Heroin/Opiates	Amount = _____	grams
Methamphetamine	Amount = _____	grams
Other: _____	Amount = _____	Unit = _____
_____	Amount = _____	Unit = _____
_____	Amount = _____	Unit = _____
_____	Amount = _____	Unit = _____
_____	Amount = _____	Unit = _____

D.12 Number of methamphetamine labs seized..... _____

D.13 Number of cultivated marijuana plants seized..... _____

D.14 Number of cannabis-growing operations dismantled..... _____

Drug Seizures

D.15 Number of firearms seized ..... _____

a. Number of firearms seized that were reported to NIBIN ..... _____

b. Number of hits in NIBIN ..... _____

Property Seizures and Forfeitures

D.16 Value of assets seized:

a. Real property (e.g., house, business) ..... $_____

b. Cash ..... $_____

c. Other (e.g., vehicles, jewelry) ..... $_____

D.17 Number of forfeiture cases filed:

a. Federal ..... _____

b. State ..... _____

D.18 Value of assets forfeited:

a. Real property (e.g., house, business) ..... $_____

b. Cash ..... $_____

c. Other (e.g., vehicles, jewelry) ..... $_____

Other

D.19 Number of hazardous sites (i.e., methamphetamine labs) identified ..... _____

D.20 Number of hazardous sites (i.e., methamphetamine labs) cleaned ..... _____

D.21 Number of disrupted drug trafficking organizations ..... _____

D.22 Number of dismantled drug trafficking organizations ..... _____

E. Task Force Outcomes

(Please complete the following items for the current reporting period.)

Prosecutions of Task Force Initiated Cases

E.1 Number of defendants referred for prosecution ..... _____

a. Number of defendants dropped (i.e., not resulting in prosecution)..... _____

	Federal	State
b. Number of defendants accepted for prosecution.....	_____	_____
Of the number of defendants ***accepted for prosecution***, as reported in b above, how many defendants are documented criminal gang members?	_____	_____
c. Number of defendants deferred from prosecution (i.e., referred to drug court, etc.)	_____	_____

Convictions & Sentencing of Defendants in Task Force Initiated Cases

	Federal	State
E.2 Number of defendants convicted.....	_____	_____
a. Number of defendants convicted, as reported in E2 above, who are ***documented criminal gang members***.....	_____	_____
E.3 Number of convicted defendants sentenced to prison	_____	_____
E.4 Total length of prison sentences imposed on convicted defendants.....	_____	_____
a. As reported in E4 above, the total length of prison sentences imposed on convicted defendants who are ***documented criminal gang members*** .	_____	_____

E.5 Use the space below to provide information regarding specific cases/convictions that clearly demonstrate a documented impact and benefit to public safety in the targeted jurisdiction. (i.e., disruption of a specific criminal gang/enterprise that results in the documented and sustained decrease in violent crime in the affected target area.)

_____

_____

_____

_____

_____

_____

Appendix 3: Analysis Worksheet

Attached.