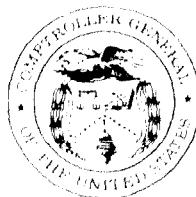United States General Accounting Office

GAO

Report to the Ranking Minority Member, Subcommittee on Courts and Administrative Practice, Committee on the Judiciary, U.S. Senate
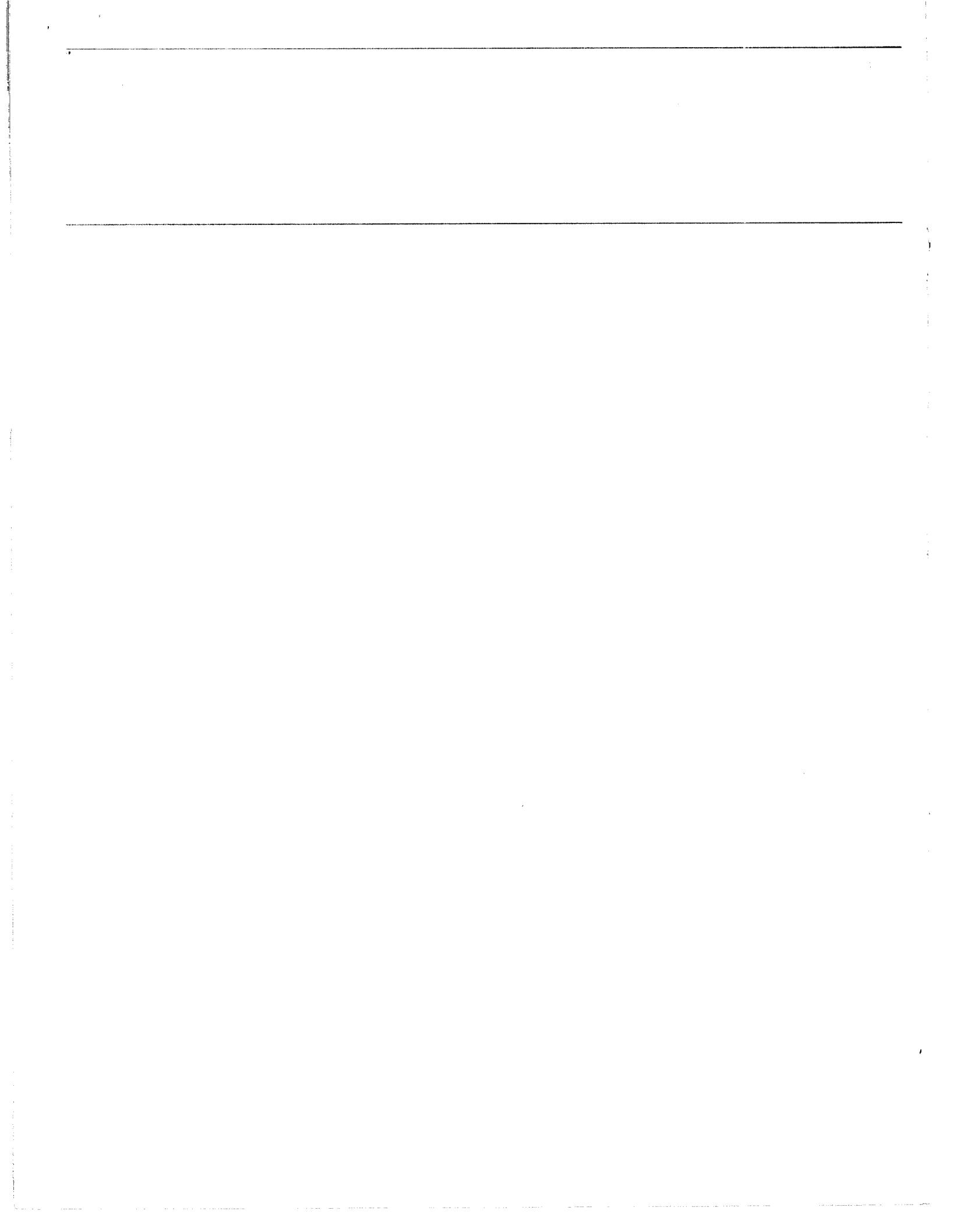
March 1991

U.S. ATTORNEYS

Better Models Can Reduce Resource Disparities Among Offices

143399

GAO

United States
General Accounting Office
Washington, D.C. 20548

General Government Division

B-241124

March 6, 1991

The Honorable Charles E. Grassley
Ranking Minority Member
Subcommittee on Courts and
  Administrative Practice
Committee on the Judiciary
United States Senate

Dear Senator Grassley:

This report responds to your request that we review the Department of
Justice's process for allocating attorneys among the 94 United States
Attorney offices. Through discussions with your office, we developed
two objectives: (1) to determine whether attorney staffing disparities
exist among the offices, and if so, (2) to identify attorney staffing alloca-
tion methods that would substantially reduce disparities among offices
and make the staff allocations more economically efficient.

To accomplish our objectives, we developed a model (the "workload
model") that seeks to account for differences in the workloads of the
U.S. Attorney offices. We used this model to assess attorney staffing dis-
parities among U.S. Attorney offices.[1] We then developed a second
model (the "allocation model") to allocate new attorney positions in
such a way as to reduce staffing disparities identified by the workload
model.

Results in Brief

The Justice Department's allocation process does not adequately
account for differences in complexity of legal workload among U.S.
Attorney offices. There are many factors, including type of case (e.g.,
drugs, organized crime), number of defendants, and whether a trial or
indictment occurred, that cause some cases to be more complex—taking
more time and effort to litigate—than others. Yet Justice's allocation
process seeks to measure only a few of these factors.

The results of our workload model suggest that resource disparities
exist among the U.S. Attorney offices. On the basis of fiscal year 1989
data, the model results showed 44 offices that expended significantly
less criminal attorney time than expected and 35 offices that expended

---

[1]The workload model addresses only the question of the relative resource requirements of the U.S.
Attorney offices. We did not determine the absolute resource requirements of U.S. Attorney offices,
i.e., how many total attorney staff each office requires to maximize its efficiency and effectiveness.

significantly less civil attorney time than expected. Conversely, 22 offices used significantly more criminal time than expected, and 37 offices used significantly more civil time than expected.

Our allocation model distributed additional positions to the attorney offices in such a way as to reduce staffing disparities identified by the workload model. Unlike the model currently used by the Justice Department, our allocation model can be applied either to total criminal or civil workloads or to more specific areas such as drugs. For the drug litigation area, we compared the model's allocation of 423 additional attorneys with the Justice Department's actual 1989 allocation of 423 drug crime attorneys. Our model's allocation showed a high level of agreement with the allocation Justice actually made for most offices, but the two allocations differed substantially for some offices.

Interpreted with due care, the workload and allocation models of this report could aid Department of Justice managers in assessing the staffing needs of U.S. Attorney offices by better measuring the complexity of workloads and, to reduce staffing disparities among the offices, by allocating additional positions to reduce such disparities.

Background

The 93 U.S. Attorneys[2] are the principal litigators for the U.S. government. They are presidential appointees who set enforcement priorities consistent with the Attorney General's law enforcement goals, operate largely autonomously, and control the use of staff resources allocated to their offices. Generally, each U.S. Attorney office has a criminal and civil unit staffed by Assistant United States Attorneys, and each unit is responsible for prosecuting and litigating cases within its respective area.

The Executive Office for United States Attorneys in Justice headquarters provides general assistance and nonlitigative oversight to the 94 offices. This assistance involves (1) allocating personnel and financial resources; (2) publishing the U.S. Attorneys' annual caseload statistics; (3) coordinating the relationships of the U.S. Attorney offices with other units of Justice; and (4) providing legal opinions, interpretations, and advice to U.S. Attorneys on the budget, legislation, ethics, and Justice guidelines.

---

[2]Because a single U.S. Attorney administers offices in both Guam and the Northern Mariana islands, there are 94 U.S. Attorney offices but only 93 U.S. Attorneys.

During the last decade, the number of cases handled by the U.S. Attorneys increased substantially, especially for white-collar crimes and drug crimes. Many U.S. Attorneys have advised the Department of Justice that their resources are inadequate to handle their growing caseloads. They also have asserted that serious disparities exist in the distribution of resources among offices.

Congress has responded to the resource concerns of the U.S. Attorneys by authorizing the appointment of 428 new federal prosecutors (Assistant U.S. Attorneys) in fiscal year 1989 and 687 in fiscal year 1990. This represents a 41-percent increase from the 2,720 attorney positions authorized in fiscal year 1988.

Approach and Methodology

Before developing our models, we met with officials at the Executive Office for United States Attorneys and 28 U.S. Attorney offices and reviewed the literature on case weighting models. At the Executive Office, officials from the Administrative Services and Legal Information Systems divisions explained the Department of Justice's methods for allocating resources. On the basis of our discussions with these officials and our literature review, we developed a set of hypotheses about the relative attorney time requirements of different types of cases. The hypotheses helped us to identify variables, measured by Department of Justice or Judicial Branch data systems, which could be used to classify cases according to their relative time requirements. However, several potentially relevant variables were not measured in the available data systems, such as grand jury and trial hours, and hence could not be used in modelling the time requirements of workloads (see app. I).

These officials also provided most of the data for our review, including (1) supporting documents on the Department of Justice's current resource allocation method; (2) statistics on resource allocations in prior years; (3) U.S. Attorneys' Resource Utilization Reports; and (4) automated criminal and civil master files for fiscal years 1987, 1988, and 1989. At 28 U.S. Attorney offices, we obtained information on the resource allocation process, the adequacy of resources, and the impacts of resource shortages on the legal system.[3]

---

[3]The 28 offices were a judgmental, or nonrandom, sample of the 91 offices in the continental United States and Puerto Rico. The sample was selected to include at least one district from each of the 12 judicial circuits. The offices visited accounted for 50 percent of the total criminal and civil cases filed in court by U.S. Attorneys in fiscal year 1989.

As part of the modelling work, we evaluated the completeness and accuracy of the Executive Office's central criminal database by comparing it with a criminal database maintained independently by the Administrative Office of the United States Courts. The results of the comparison are presented in appendix II of this report.

Appendixes I through VI present the workload and allocation models, instructions for using these models, and detailed documentation. Appendix I presents workload models for civil and criminal litigation. Appendix II discusses the accuracy and completeness of data we used to build the workload model. Appendix III shows how the workload model can be used to assess staffing disparities among U.S. Attorney offices. Appendix IV applies the allocation model to several funding scenarios based on appropriations recently approved by Congress, such as the 614 new attorney positions targeted in 1990 to prosecute drug and violent crimes. Appendixes V and VI provide more detailed technical documentation of the workload model and of the data sets and variables used in the analyses.
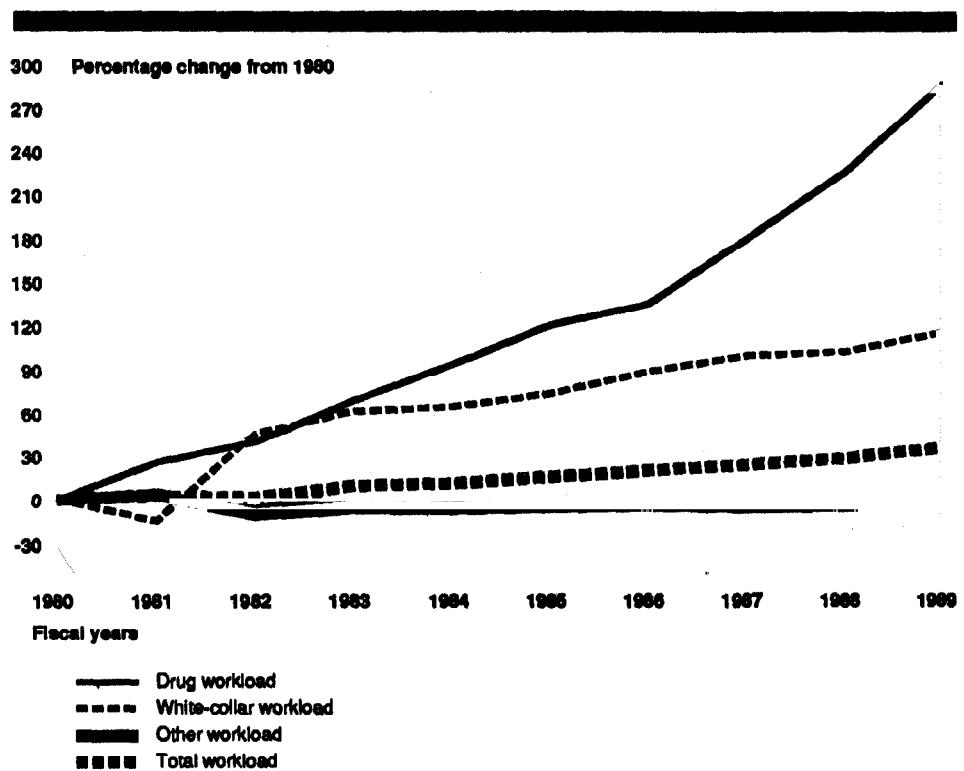
We did our work between January 1989 and April 1990 in accordance with generally accepted government auditing standards. We obtained official comments on this report from the Department of Justice and have included and evaluated them in appendix VII.

Overview of the U.S. Attorney Offices' Workload and Staffing During the 1980s

During the 1980s, the number of criminal matters[4] and cases reported by the U.S. Attorneys increased by 36 percent: 159,742 matters and cases were reported in fiscal year 1989 compared with 117,853 matters and cases in fiscal year 1980. Figure 1 shows trends in two categories of criminal litigation, drug cases and white-collar crime cases, that increased substantially between 1980 and 1989.

---

[4]A matter is a legal item being considered by the U.S. Attorney for prosecution or litigation in court.

Figure 1: U.S. Attorneys' Criminal Workload Change, Fiscal Years 1981-1989

300 Percentage change from 1980

270

240

210

180

150

120

90

60

30

0

-30

1980   1981   1982   1983   1984   1985   1986   1987   1988   1989

Fiscal years

——— Drug workload
■ ■ ■ ■ White-collar workload
▄▄▄▄ Other workload
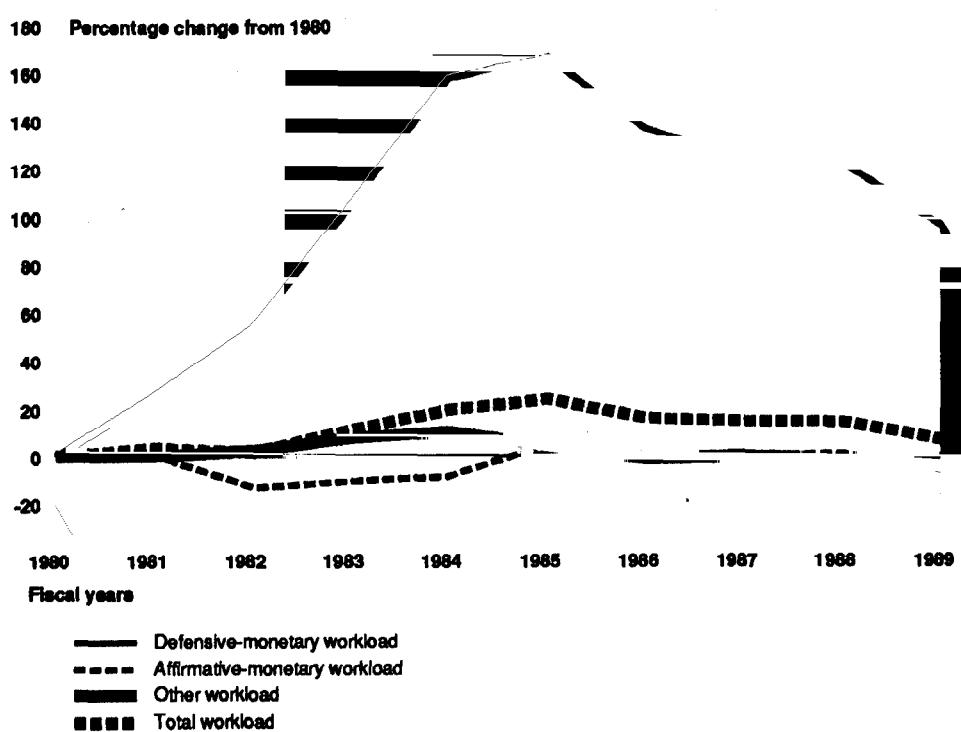■ ■ ■ ■ Total workload

Note: Percentage change calculated using 1980 totals as base.
Source: United States Attorneys' Statistical Report, fiscal years 1980 through 1989.

The number of civil matters and cases increased much less rapidly, by about 7 percent during the 10-year period: 261,947 matters and cases were handled in fiscal year 1989 compared with 245,755 matters and cases handled in fiscal year 1980. Unlike criminal matters and cases, civil matters and cases did not continually increase during the 1980s. As shown in figure 2, there was an increase of more than 100 percent between 1980 and 1989 in the reported number of monetary suits filed against the federal government (defensive litigation), but after 1985, the number of defensive monetary suits declined. There was also a slight decline (less than 5 percent) since 1980 in the reported number of monetary suits filed by the government seeking monetary relief against others (affirmative litigation).

Figure 2: U.S. Attorneys' Civil Workload Change, Fiscal Years 1981-1989

Note: Percentage change calculated using 1980 data as the base.
Source: United States Attorneys' Statistical Report, fiscal years 1980 through 1989.

Nationwide, the U.S. Attorney offices used about 2,650 attorney workyears processing criminal and civil matters and cases in 1989—a 56-percent increase from the 1,694 attorney workyears spent in 1980.

Department of Justice Resource Allocation Process

When additional attorney positions are approved by Congress, officials at the Executive Office of U.S. Attorneys use a two-step process to allocate the new positions among the U.S. Attorney offices. First, preliminary allocations are produced using an allocation model that employs data on seven variables. The seven variables include the numbers of district court judges, civil and criminal cases, grand jury hours, and trials. Second, additional factors bearing on the allocation, including numbers of branch offices in the districts, court-related travel time, and client agency workload projections, are qualitatively assessed. The allocation proposals developed using this two-step process are subject to approval by the Deputy Attorney General.

At 17 of the 28 U.S. Attorney offices we visited, officials questioned this decisionmaking process. They maintained that the process does not adequately account for the complexity of caseloads and thus does not accurately measure the work done by different offices. Officials from the majority of the U.S. Attorney offices we visited supported the need for a system that takes case complexity into account by quantitatively weighting cases by their complexity or difficulty. For example, an official from one large office we visited pointed out that a case weighting system would show that his district's work includes an unusually high percentage of large, complex, multidefendant cases, a fact that is not reflected in raw caseload statistics.

We assessed Justice's resource allocation model and found two weaknesses. First, as asserted by many officials with whom we spoke, Justice's model does not quantitatively measure case complexity, omitting such aspects as type of litigation and number of defendants. Since the type of litigation is omitted, Justice's model cannot provide guidance in allocating positions that are targeted to specific litigation areas, such as drugs.

Second, Justice's model uses the number of judges as a predictor of the number of attorney positions. Treating the number of judges as a predictor suggests that the number of judges causes the number of attorneys. However, it is more appropriate to think that the association between these two variables arises from their mutual dependence upon a common cause, namely the dependence of both number of judges and number of attorneys upon the need for legal services. Justice officials acknowledge that the use of judges as a predictor could have the adverse consequence that any inequities or inefficiencies in the distribution of judges would affect the distribution of attorneys.

Workload Model

To improve the measurement of actual workload, we developed a model that assigns weights to cases on the basis of the average time required to handle cases of the same general type.[5] For example, according to the literature on legal case weighting systems and discussions with officials of the Executive Office for U.S. Attorneys and the 28 U.S. Attorney offices we visited, criminal cases involving multiple defendants, grand jury indictments, and jury trials usually take more time than those that do not, and civil cases involving multiple defendants, motions, and jury

---

[5]Separate workload models were developed for each of 13 civil and criminal litigation areas (see app. I).

trials usually take more time than those that do not. In all, criminal and civil cases were assigned to 12 classes depending on the amount of time they take.

On the criminal side, we grouped cases into the following six litigation areas: public corruption, economic crime, organized crime, drugs, violent crime, and other criminal offenses. On the civil side, we grouped cases into seven areas: affirmative-monetary, affirmative-nonmonetary, defensive-monetary, defensive-other, civil forfeitures, foreclosures, and prisoner petitions.

The workload model uses a statistical technique called multiple regression to estimate the weights, i.e., the average attorney time requirements, of the different types of cases that we identified in our preliminary interviews and literature review.[6] The main products of the statistical analysis are the estimated staffing ratios of U.S. Attorney offices. The staffing ratio of each office compares the model's estimate of the time needed to process the office's workload (the numerator of the staffing ratio) with the time reported by the office (the denominator of the staffing ratio). The staffing ratio of each office gauges its need for additional attorneys relative to other offices.[7]

The workload model assumes that U.S. Attorneys do not control the number or types of criminal and civil case referrals. That is determined by the amount of criminal and civil litigation brought to U.S. Attorney offices. The model also assumes that both the data and the theory of factors affecting attorney time requirements are accurate and complete. The model is designed to assess resource disparities and not to gauge how many cases an attorney can handle before the quality of legal work declines.

Resource Disparities Among the U.S. Attorney Offices

The workload model results suggest staffing disparities among the U.S. Attorney offices. For fiscal year 1989, the model identified 44 offices that spent significantly less attorney time than the model predicted for their criminal workload and 22 offices that used significantly more time than predicted. For the civil workload, there were 35 offices that took

---

[6]As discussed in appendix I, two kinds of data, data on reported attorney time expenditures and data on the numbers of cases of the different types, were used to estimate the average time requirements of different types of cases.

[7]This interpretation of the staffing ratio of an office assumes that the assumptions of the workload model are satisfied. See appendix I for possible violations of the assumptions and for alternative interpretations of the staffing ratios.

less time than predicted for their workload and 37 that took more time than predicted.

Other things being equal, offices using less attorney time than predicted are understaffed relative to other offices, and offices using more attorney time than predicted are overstaffed relative to other offices. Since all inferences are relative to other offices, inferring that an office is overstaffed does not imply that the office does not need additional attorneys.

Allocation Model

The allocation model uses the estimated staffing ratios from our workload model to allocate new attorney positions among offices. The model allocates new positions so that any staffing disparity among offices shown by the workload model is decreased as much as possible without reducing the staff of any U.S. Attorney office. The model allocates any number of new positions on the basis of the total criminal workload, the total civil workload, or the workload of a specialized litigation area, such as drugs or violent crime. The following discussion compares Justice's 1989 allocation of 423 new attorney positions to prosecute drug cases with an allocation of the 423 positions based on our workload and allocation models.

In fiscal year 1989, the U.S. Attorneys spent 566 attorney workyears to handle about 33,360 drug cases and matters. On the basis of the fiscal year 1989 reported workloads and reported time expenditures, the workload model indicates that 31 offices spent significantly less attorney time than expected for their drug workload—indicating a relatively greater need for additional staff. Our workload model analysis further indicates that 25 offices spent significantly more time than expected for their drug workload. For 35 offices, there was no significant difference between the reported and expected times.

Table 1 compares Justice's actual allocation of 423 new attorneys during fiscal year 1989 with our resource allocation model's allocations. The table generally shows a high level of agreement between Justice's and our model's allocations, but there are large differences for a number of individual offices. For example, the model's allocation for 18 U.S. Attorney offices differed by 5 or more positions from Justice's allocation. On the other hand, 52 offices received the same or nearly the same allocation (differed by 1) through Justice's allocation process and the model. Appendix IV shows that offices that were relatively understaffed in the drug crime area before Justice's 1989 allocation, according

to our workload model, tended to remain relatively understaffed in drug crime attorneys after Justice's 1989 allocation.

Table 1: Comparison of GAO Allocation Model Results and Justice's Reported Allocations of 423 Attorneys in 1989

District	Model staff allocation	Reported staff allocation	Difference[a]
Florida south	21	42	21
California central	1	21	20
Texas west	18	6	12
District of Columbia[b]	16	28	12
Tennessee west	13	4	9
Texas south	15	8	7
New York east	23	16	7
Washington west	9	3	6
New Mexico	9	3	6
Texas east	1	7	6
New Jersey	2	7	5
California north	1	6	5
Connecticut	1	6	5
Arizona	9	4	5
California south	11	6	5
Virginia east	3	8	5
New York south	4	9	5
West Virginia south	8	3	5
West Virginia north	5	1	4
Ohio south	7	3	4
Hawaii	7	3	4
Louisiana east	7	3	4
Alabama south	6	2	4
North Carolina east	5	1	4
Massachusetts	0	4	4
Nevada	0	4	4
Texas north	5	9	4
Maryland	3	6	3
Alabama north	2	5	3
North Carolina middle	5	2	3
North Carolina west	4	2	2
Nebraska	3	1	2
Montana	3	1	2
Michigan west	1	3	2
New York north	3	5	2
Missouri west	3	5	2
Puerto Rico	5	3	2
Illinois central	5	3	2

(continued)

District	Model staff allocation	Reported staff allocation	Difference[a]
Tennessee east	4	2	2
Louisiana middle	2	3	1
Oklahoma west	4	3	1
Tennessee middle	3	2	1
Georgia middle	4	3	1
Pennsylvania middle	5	4	1
Iowa south	3	2	1
Iowa north	4	3	1
Indiana north	4	3	1
Wisconsin east	5	4	1
Kentucky east	3	2	1
Kentucky west	3	4	1
Maine	3	2.	1
Michigan east	7	8	1
Pennsylvania east	5	6	1
Colorado	1	2	1
Oregon	4	5	1
Pennsylvania west	2	3	1
Indiana south	2	3	1
South Carolina	2	1	1
Alabama middle	2	3	1
Mississippi north	2	3	1
Mississippi south	2	3	1
Alaska	2	3	1
Hawaii	4	3	1
Rhode Island	2	1	1
New Hampshire	2	1	1
Virginia west	3	2	1
Georgia south	1	2	1
Illinois south	3	4	1
Vermont	1	2	1
New York west	4	3	1
Georgia north	6	5	1
Florida middle	13	12	1
Florida north	6	5	1
Minnesota	4	3	1
Illinois north	10	9	1
Oklahoma north	3	2	1
Utah	3	2	1

(continued)

District	Model staff allocation	Reported staff allocation	Difference[a]
Wyoming	2	2	0
Delaware	2	2	0
Kansas	4	4	0
Louisiana west	2	2	0
Arkansas east	2	2	0
Missouri east	4	4	0
California east	5	5	0
Idaho	2	2	0
Washington east	2	2	0
Wisconsin west	2	2	0
Oklahoma east	2	2	0
North Dakota	2	2	0
Arkansas west	2	2	0
South Dakota	3	3	0
Ohio north	4	4	0

Note: There were 91 U.S. Attorney offices included in our workload and allocation models. Guam, Northern Mariana Islands, and Virgin Islands were excluded because of data limitations.

[a]Differences are in absolute values. In other words, they reflect the difference between the higher and lower allocation, irrespective of which one is greater.

[b]Justice's allocation for Washington, D.C., includes both the district and superior courts, whereas the model's allocation is for district court only. Some of the additional 28 attorneys may have been allocated to the superior court, accounting for some of the difference.

Conclusion

The Department of Justice does not have a systematic way to assess the complexity of U.S. Attorney workloads. Justice's quantitative model treats all cases as equal in complexity. Since the amount of attorney time needed to investigate and prosecute cases varies greatly according to the various factors that reflect case complexity (e.g., type of case), weighting cases by their expected time requirements can be a helpful tool in accurately gauging the workloads and needs for additional attorneys of the U.S. Attorney offices. Once those workloads and needs are determined, an allocation model such as ours will help distribute new attorney positions to reduce staffing disparities.

The models we developed are meant to aid decisionmaking by identifying offices that appear to deviate significantly from typical resource usage patterns. They can also aid in allocating new Assistant U.S. Attorney positions to increase the efficiency of U.S. Attorney effort overall. They are not substitutes for professional judgment. Our models can only provide information that enables decisionmakers to compare

offices in what we believe is a useful way to support their effort to operate in the most efficient manner. Professional judgment of responsible Justice Department officials is obviously critical to the proper interpretation and use of the models.

Recommendations

We recommend that the Attorney General direct the Executive Office for U.S. Attorneys to take the following steps to improve the resource allocation process for U.S. Attorney offices:

- Implement a quantitative workload model for gauging the workloads of U.S. Attorney offices that is sensitive to variations in the time requirements of different types of cases.
- Adopt a resource allocation model that responds to resource disparities among U.S. Attorney offices identified by the workload model and can be used in conjunction with judgmental factors to allocate additional attorney positions.

Agency Comments

The Department of Justice provided written comments on a draft of this report. Justice said that our model will serve as a useful guide in conjunction with other factors to be considered when making attorney allocation decisions and that it was looking forward to using our model when future allocations are made. Justice also discussed what it felt were some relevant factors omitted from our model that go into Justice's allocation decisionmaking process. We agree that there are some significant factors not measured by our workload model that management would want to consider in the overall decisionmaking process. Some of the omitted factors Justice mentioned, such as trial length and grand jury time, could be incorporated in our workload model; however, they would first have to be gathered and reported in sufficient detail by the U.S. Attorneys in their database systems. Justice's comments and our responses are presented in appendix VII.

As arranged with your office, unless you publicly announce the contents of the report earlier, we plan no further distribution until 7 days from the date of this letter. At that time, we will send copies to the Attorney General, the 93 U.S. Attorneys, and other interested parties.

Major contributors to this report are listed in appendix VIII. If you have any questions about this report, please contact me on (202) 275-8389.

Sincerely yours,

Lowell Dodge
Director, Administration
   of Justice Issues

Contents

Contents

Figures

Abbreviations

OT	other criminal offenses
AM	affirmative-monetary
AN	affirmative-nonmonetary
AO	Administrative Office of the United States Courts
AUSA	Assistant United States Attorney
CF	civil forfeitures
DM	defensive-monetary
DO	defensive-other
DR	drugs
DSE	dual system estimate
EC	economic crime
EO	Executive Office for United States Attorneys
FTE	full time equivalent
FO	foreclosures
OC	organized crime
PC	public corruption
PP	prisoner petitions
USAO	United States Attorney Office
VC	violent crime

Criminal and Civil Workload Weighting Models

The purpose of the workload weighting models is to estimate the time required to process different kinds of cases and defendants through the steps of adjudication. Our approach features the empirical weighting of cases of different kinds on the basis of a statistical model of attorney time requirements.[1] This appendix (1) documents the sources of data used to build the models, (2) describes the variables and modeling assumptions of our criminal and civil workload weighting models, (3) assesses how well these models describe the data, and (4) discusses some alternative interpretations of the models.

Sources of Data

We used two kinds of data to build the models. Both are collected on a continuing basis by the U.S. Attorneys.

1. Workload data, i.e., the numbers of legal proceedings (matters and cases) of different kinds that were litigated by each of the attorney offices during the fiscal year.

2. Time expenditure data, i.e., data on the amounts of regular time (excluding uncompensated overtime), conventionally measured in person-years of work called "full time equivalents" (FTE), that were spent by each of the offices during the fiscal year.

For our evaluation, workload data for fiscal years 1987, 1988, and 1989 were obtained from the Executive Office for U.S. Attorneys. In the criminal weighting analysis, although not in the civil weighting analysis, workload data from the Administrative Office of U.S. Courts were also used. We refer to the Executive Office for U.S. Attorneys information as the EO data and the Administrative Office of U.S. Courts information as the AO data.

Time expenditure data were obtained from the U.S. Attorneys' resource utilization reports (USA-5s) for fiscal years 1987, 1988, and 1989. The reporting system does not report the amount of time attorneys spend on individual cases but rather the numbers of FTEs spent by U.S. Attorney offices in 10 criminal and 11 civil litigation areas. The criminal areas include official corruption, organized crime, economic crime, narcotics, and violent crime. The civil areas include defensive and affirmative litigation (monetary and other), bankruptcy, and foreclosures. U.S. Attorneys report these data monthly to the Executive Office for use in

---

[1] Our approach is "empirical" in that weights are estimated using quantitative data rather than postulated on the basis of expert opinions, as in some previous approaches.

preparing the budget, responding to inquiries about the allocation of
U.S. Attorney resources, and monitoring the allocation of resources.

Assessing Staffing Disparities by Comparing Reported and Predicted Time Expenditures

In our model, the staffing needs of U.S. Attorney offices are assessed
relative to each other. For example, suppose office A handles 100 cases
and spends 2 FTEs, and office B handles 500 cases and spends 25 FTEs.
We can say that A is understaffed relative to B since attorneys in A
handle 50 cases each (i.e., 100 cases divided by 2 FTEs) whereas attor-
neys in B handle 20 cases each (i.e., 500 cases divided by 25 FTEs).
Equivalently, B is overstaffed relative to A. In this kind of analysis,
staffing needs are assessed by comparing each office's ratio of total
caseload to total time expenditures. Our models compute ratios similar
to these.

The simple model of offices A and B illustrates the two main properties
of the more complex models that we used to estimate weighted work-
loads and to infer staffing disparities among offices:

1. Staffing needs are measured by comparing reported workloads with
reported time expenditures.

2. Inferences about staffing needs of any U.S. Attorney office (USAO) are
relative to other offices. This is because there is no absolute criterion of
the appropriate size of an attorney's caseload. Although B is inferred to
be overstaffed relative to A in the simplest model, B might still be
understaffed.

One problem with the simplified model of office A and office B is that
cases vary in their time requirements. A second problem is that time
requirements depend not only on the number of cases but also on the
number of defendants. That is, a case with multiple defendants will gen-
erally require more attorney time. These problems imply that a simple
model is likely to give misleading results: two offices with identical num-
bers of staff and cases but with different mixes of cases or numbers of
defendants, requiring different amounts of time, might be incorrectly
inferred to have the same degree of staffing need.

For these reasons, we used weighted workloads, with different kinds of
cases receiving different weights, rather than unweighted caseloads to
gauge the work performed by USAOs. The estimated weights for different
types of cases depend on a theory of the factors affecting attorney time
requirements, e.g., factors such as the litigation area and whether or not

the case had a trial during the fiscal year. These factors are the predictor (independent) variables of the model, and the attorney time required is the predicted (dependent) variable, which the model predicts.

Our theory is based upon a review of the literature on legal case weighting systems and discussions with officials of the Executive Office and of the 28 USAOs that we visited. We hypothesize that criminal cases involving multiple defendants, grand jury indictments, and jury trials generally require greater attorney time expenditures than criminal cases that do not. Civil cases involving multiple defendants, motions, and jury trials take more time than those that do not.

Based upon the theory of time requirements and appropriate modeling assumptions discussed in the following section, our approach allows case and defendant weights, weighted workloads, and staffing ratios to be estimated from available USAO data. The staffing ratios can be interpreted analogously to the ratios of cases and time of the simplified model of offices A and B.

We estimated the weights, weighted workloads, and staffing ratios using multiple regression. The predicted variable of the regression equation was the reported number of FTEs of the USAO in a specified litigation area during the fiscal year. The predictor variables were the numbers of cases of different kinds that were processed by the USAO in the litigation area during the fiscal year. For each USAO in each litigation area, the technique yields a predicted number of FTEs. The predicted number is the number of FTEs that, according to the model, would have been expended to process the USAO's workload based on the average USAO's performance.

Summing the predicted FTEs of a USAO across criminal litigation areas yields the total predicted criminal time expenditure of the USAO. Summing the reported FTEs of a USAO across criminal litigation areas yields the total reported criminal time expenditure of the USAO. Dividing the former sum by the latter yields the estimated criminal staffing ratio of the USAO. The analogous procedure for civil FTEs and case data yields the estimated civil staffing ratios. USAOs for which this ratio is high (low) are inferred to be understaffed (overstaffed) relative to other USAOs.

In any litigation area or set of such areas, the weighted caseload and the weighted defendant load of a USAO are computed by multiplying the predicted time expenditure of the USAO by specified constants. For the

weighted caseload the constant is chosen so that the sum across offices of the weighted caseloads equals the sum across offices of unweighted cases. For the weighted defendant load, the constant is chosen so that the sum across offices of the weighted defendant loads equals the sum across offices of the unweighted defendants.

Since both weighted cases and weighted defendants are constant multiples of the predicted time (just as weight measured in pounds is a constant multiple of weight measured in ounces), all three statistics measure the same concept, i.e., the weighted workload. Staffing inferences can be based on the (1) ratio of weighted cases and reported time, (2) the ratio of weighted defendants and reported time, or (3) the ratio of predicted time and reported time. Inferences about understaffing/overstaffing will be exactly the same with all three approaches.

Yet, the third approach, using the ratio of predicted time and reported time, seems preferable since unity, one (i.e., predicted time = reported time), corresponds to a natural point on the scale, like the freezing point of water on a temperature scale. Offices with ratios that are significantly greater than 1 are inferred to be understaffed relative to the other offices.[2] Offices with ratios that are significantly less than 1 are inferred to be overstaffed relative to the other offices. In this report, we refer to the ratio of predicted time and reported time of an office as the staffing ratio of the office.

The weighting models we used to estimate the staffing ratios and to generate allocations of new attorney positions resulted from a complex model selection procedure. In this procedure, a large number of alternative specifications of the models, distinguished by different numbers and combinations of predictors, were compared with respect to their overall fit to the data and other statistical properties. In brief, the final models are those that appeared to fit the data most closely, according to statistical criteria, without capitalizing on chance variations in the small sample of 91 USAOs (i.e., without "overfitting" the data).[3]

The case and defendant weights in the selected models were estimated, using combined data for fiscal years 1987, 1988, and 1989, under the assumption that the amounts of time required to process different kinds

[2]By "significantly greater than 1," we mean greater than 1 by too large an amount to be attributed to chance.

[3]There were 91 USAOs included in our weighting model. Guam, Northern Mariana islands, and Virgin Islands were excluded because of data limitations.

of cases were constant between 1987 and 1989. The estimation technique allowed for possible dependence between the observations for the same office in different years, such as dependence caused by the continuation of cases from one year to the next (see app. V). The case weights are estimates of the average numbers of FTEs per case expended by AUSAs for each kind of case. The defendant weights are estimates of the average numbers of FTEs per defendant expended by AUSAs for each kind of defendant.

The civil weights were estimated separately for each of seven civil litigation areas by regressing reported FTEs in the litigation area on the numbers of cases and defendants of different kinds. The workload data for the civil analysis were obtained strictly from the EO caseload data. The criminal weights were estimated separately for each of six criminal litigation areas by regressing reported FTEs in the litigation area on estimated numbers of cases and defendants of different kinds. The workload data for the criminal analysis were combined EO-AO best estimates (i.e., "dual system estimates"; see app. II).

In each civil litigation area, the civil weighting model used a total of 18 predictors ("independent variables" or "regressors"). In each criminal litigation area, the criminal weighting model used a total of 16 predictors. (See app. V for technical discussions of the criteria for model selection and the estimation technique.)

Appeals are not included in either of the criminal or civil workload models since the FTE reporting categories did not distinguish criminal and civil appeals in 1987. (Criminal and civil appeals accounted for approximately 5 percent of total reported AUSA FTEs in 1989.) AUSA FTEs classified as management/administration (accounting for approximately 9 percent of total reported AUSA FTEs in 1989) were not modelled because these FTEs could not be related to cases recorded in the workload data. In interpreting both the criminal and civil workload models, we therefore assume that appeal and management/administration attorney time requirements are roughly proportional to predicted workloads for the remaining FTE reporting categories.

Criminal Weighting Model

On the basis of U.S. Attorneys' reporting of FTEs, we grouped criminal cases and matters into six litigation areas: public corruption, economic crime, organized crime, drugs, violent crime, and all other crimes. The types of offenses included in the six criminal litigation areas are as follows:

- Public corruption (PC): bribery of public officials and embezzlement of public money and/or property.
- Economic crime (EC): fraud and embezzlement (procurement, bank, income tax, etc.) and forgery and counterfeiting.
- Organized crime (OC): extortion, racketeering, and gambling.
- Drugs (DR): manufacturing and distribution of narcotics as well as possession of illegal drugs.
- Violent crimes (VC): bank robbery, assaults, murder, and sex offenses.
- Other criminal offenses (OT): regulatory offenses (food and drug act, maritime and shipping laws, postal violations, etc.), violations of national defense laws, theft, immigration (illegal entry and fraud), civil rights, and other miscellaneous offenses.

In each criminal litigation area, attorney FTEs were predicted using (1) total single-defendant cases, (2) total multiple-defendant cases, and (3) total defendants in each of six kinds of legal proceedings during the fiscal year: (1) matters, (2) misdemeanors, (3) pending felony cases (i.e., with filing dates before the referenced fiscal year), (4) new felony cases (i.e., with filing dates during the referenced fiscal year), (5) cases with grand jury indictments during the fiscal year, and (6) cases with trials during the fiscal year.[4] This yielded a total of 18 predictors for each criminal litigation area. Because of small frequencies, single-defendant and multiple-defendant misdemeanors were combined and single-defendant and multiple-defendant new felony cases were combined, yielding 16 predictors in the final model.

Table I.1 shows, for fiscal year 1989, the distributions by criminal litigation area of criminal FTEs, estimated cases/matters, and estimated defendants (combined EO-AO workload estimates).[5] Table I.2 shows, for fiscal year 1989, the distributions of criminal single-defendant cases, multiple-defendant cases, and total defendants by type of legal activity (combined EO-AO workload estimates).

---

[4]The six kinds of legal proceedings are listed, from lowest to highest, in the order of their hypothesized time requirements. That is, we hypothesized that matters are the least time-intensive, and trials the most time-intensive, of the six kinds of proceedings. Cases that were subject to more than one kind of proceeding during the fiscal year were assigned to the most time-intensive kind of proceeding that was reported. For example, a new felony case with an indictment and a trial during the fiscal year was classified as a trial. The hierarchy of six types of legal proceedings was used only as a means of organizing the data for analysis. The results of the workload weighting model do not depend on the validity of the hypothesis that time requirements increase with position in this hierarchy. Indeed, a distinguishing feature of our approach is that case and defendant weights are estimated from the data rather than assigned on the basis of prior information.

[5]Appendix II describes the statistical technique, "dual system estimation," used to develop combined EO and AO workload estimates.

Table I.1: Fiscal Year 1989 Criminal Data—FTEs, Cases/Matters, and Defendants, by Litigation Area (Dual EO-AO Estimates)

	FTEs		Cases/matters		Defendants	
Litigation area	Number	Percent	Number	Percent	Number	Percent
PC	99	6.4	3,236	2.2	4,392	2.1
EC	456	29.6	45,198	30.7	60,844	28.9
OC	42	2.7	1,338	.9	2,767	1.3
DR	566	36.8	33,364	22.7	58,077	27.6
VC	101	6.6	6,474	4.4	7,544	3.6
OT	274	17.8	57,458	39.1	76,747	36.5
Total	**1,538**	**99.9**	**147,068**	**100.0**	**210,371**	**100.0**

Note: Percentages do not always add to 100 because of rounding.

Table I.2: Fiscal Year 1989 Criminal Data—Single-Defendant Cases, Multiple-Defendant Cases, and Total Defendants, by Type of Legal Activity (Dual EO-AO Estimates)

	Single-defendant cases		Multiple-defendant cases		Total defendants	
Type of legal activity	Number	Percent	Number	Percent	Number	Percent
Matter	64,524	54.0	14,251	51.5	105,551	50.2
Misdemeanor	6,737	5.6	219	.8	11,419	5.4
Pending felony case	20,136	16.9	4,823	17.4	36,281	17.2
New felony case	9,078	7.6	541	2.0	11,078	5.3
Indictment	17,099	14.3	6,246	22.6	37,724	17.9
Trial	1,834	1.5	1,579	5.7	8,318	3.9
Total	**119,408**	**99.9**	**27,659**	**100.0**	**210,371**	**99.9**

Note: Percentages do not always add to 100 because of rounding.

Civil Weighting Model

On the basis of the U.S. Attorneys' reporting of FTEs, we grouped civil cases and matters into seven litigation areas: affirmative-monetary, affirmative-other, defensive-monetary, defensive-other, civil forfeitures, foreclosures, and prisoner petitions. The types of offenses included in the seven civil litigation areas are as follows:

- Affirmative-monetary (AM): bankruptcy, civil fraud, and tax refund and collection cases.
- Affirmative-nonmonetary (AN): immigration proceedings, environmental enforcement proceedings, foreign laws, land/real property suits, and other affirmative nonmonetary litigation.
- Defensive-monetary (DM): social security and medicare/medicaid claims, commercial litigation, and torts.

- Defensive-other (DO): suits against government agencies challenging a government policy, civil rights litigation, employment discrimination, and Freedom of Information Act suits.
- Civil forfeitures (CF): forfeitures and seizures of property.
- Foreclosures (FO): tax lien and mortgage foreclosures.
- Prisoner petitions (PP): post-conviction civil remedies such as habeas corpus and parole.

In each of the seven civil litigation areas, attorney FTEs were predicted using total single-defendant cases, total multiple-defendant cases, and total defendants in cases with six types of legal activity during the fiscal year: (1) matters, (2) pending cases, (3) new cases (i.e., with filing dates during the fiscal year), (4) cases with legal events (such as requests for discovery), but no motions or trials, during the fiscal year, (5) cases with one or more motions during the fiscal year, and (6) cases with one or more trials during the fiscal year.[6] This yielded 18 predictors for each civil litigation area.

Table I.3 shows, for fiscal year 1989, the distributions by civil litigation area of civil FTEs, estimated cases/matters, and estimated defendants (EO workload estimates). Table I.4 shows, for fiscal year 1989, the distributions of civil single-defendant cases, multiple-defendant cases, and total defendants by type of legal activity (EO workload estimates).

Because the civil caseloads reported by U.S. Attorneys include a large number of "paper cases," which are usually handled by paralegals and legal technicians, we excluded these from the estimation of the civil model. The excluded cases included claims related to federal assistance programs, student loan defaults, and recovery of outstanding travel advances.

---

[6]As in the discussion of the criminal model, the types of legal proceedings are listed, from lowest to highest, in the order of their hypothesized time-intensiveness. Cases with one or more types of proceeding during the fiscal year were assigned to the highest type that was reported.

Table I.3: Fiscal Year 1989 Civil Data—FTEs, Cases/Matters, and Defendants, by Litigation Area

Litigation area	FTEs		Cases/matters		Defendants	
	Number	Percent	Number	Percent	Number	Percent
AM	140	20.0	46,553	20.2	56,212	17.6
AN	56	8.0	9,888	4.3	12,145	3.8
DM	263	37.6	69,167	30.1	84,196	26.4
DO	81	11.6	16,820	7.3	22,432	7.0
CF	115	16.4	13,070	5.7	20,833	6.5
FO	29	4.1	67,272	29.2	115,915	36.3
PP	15	2.1	7,299	3.2	7,684	2.4
Total	**699**	**98.8**	**230,069**	**100.0**	**319,417**	**100.0**

Note: Percentages do not always add to 100 because of rounding.

Table I.4: Fiscal Year 1989 Civil Data—Single-Defendant Cases, Multiple-Defendant Cases, and Total Defendants, by Type of Legal Activity

Type of legal activity	Single-defendant cases		Multiple-defendant cases		Total defendants	
	Number	Percent	Number	Percent	Number	Percent
Matter	9,475	5.4	3,976	7.2	19,993	6.3
Pending case	61,157	35.0	20,305	36.6	115,024	36.0
New case	10,103	5.8	6,729	12.1	27,000	8.4
Event	63,445	36.4	10,541	19.0	91,079	28.5
Motion	25,223	14.5	10,187	18.3	51,401	16.1
Trial	5,121	2.9	3,807	6.8	14,920	4.7
Total	**174,524**	**100.0**	**55,545**	**100.0**	**319,417**	**100.0**

Weighting Model Assumptions

The validity of staffing ratios and resource allocations estimated using the weighting (and allocation) model rests on four assumptions:

1. USAOs do not control their workloads. The models assume that workloads are largely beyond the control of the U.S. Attorney offices. In particular, USAOs do not control the numbers of crimes of different kinds that are committed nor the numbers of cases that are referred to them by investigative agencies. Furthermore, the law, the seriousness of offenses that are presented, and the quality of the evidence available for prosecution, rather than discretionary actions taken by the offices themselves, primarily determine the numbers and kinds of cases that must be processed.

To the extent that U.S. Attorneys control the workloads of their offices, by means of declination policies[7] or discretionary allocation of the office's resources to one litigation area rather than another, workloads are not strictly beyond their control. Workloads may be both causes and consequences of time expenditures. If so, estimates of weights, predicted time expenditures, and staffing ratios will be biased. The magnitude and direction of the bias is difficult to infer. This kind of bias is called "simultaneous equations bias."

2. There are no major omitted variables. The models assume that the classifications by litigation area, number of defendants, and type of legal activity during the fiscal year capture all important differences among the workloads of offices that affect time requirements. There might be additional factors, including both unknown factors and factors that are currently unmeasured. For example, grand jury and trial hours, appellate hours, and travel time among local jurisdictions are factors affecting time requirements that are omitted from the current model.

Failure to include such omitted variables in the weighting model can induce biases of undetermined direction and magnitude in the staffing estimates. However, this bias will result only if the omitted variables vary across offices in a way that cannot be predicted using the variables included in the model.

3. There are no significant differential coverage errors in the data. The civil weighting model assumes that the FTE data and EO civil workload data are equally complete for all offices. Differential coverage error, i.e., larger proportions of missing cases for some offices than for others, can induce biases in the estimated staffing ratios and allocations. In particular, offices with unusually large proportions of civil cases that are not recorded in the EO civil database will not receive credit for these cases in the staffing inferences. The staffing ratios of such offices will be biased downward. That is, the offices with poor coverage will be inferred to be less severely understaffed than they really are.

Similar comments pertain to the effects of differential coverage error on the results of the criminal weighting model, except that in the criminal analysis, best EO-AO caseload and defendant load estimates (dual system estimates), rather than EO estimates alone, were used to gauge workloads. If the assumptions of dual system estimation (see app. II) are

---

[7]Declination policies are guidelines that cover the types and severity of crimes that U.S. attorneys will not normally prosecute.

valid in this application, then the estimated criminal staffing ratios and recommended criminal attorney allocations should be correct even if differential coverage error exists in one or both data sources.

4. There are no significant differential measurement errors among offices. That is, FTEs, cases, and defendants are assumed to be classified with equal accuracy in all offices. For example, if 20 percent of trials are misclassified as cases without trials in office A, we assume that 20 percent of trials in office B are also misclassified as cases without trials. Differential measurement error can cause bias in the staffing ratios and recommended allocations. The direction of the bias depends on the particular category of the data (e.g., litigation area, number of defendants, trial, indictment) that is measured erroneously. Appendix II presents estimates of both differential coverage error and differential measurement error for the criminal workload data.

Goodness-of-fit of the Weighting Models

By goodness-of-fit we mean the extent to which a model agrees with the data.[8] Table I.5 shows the most commonly used measure of goodness-of-fit, the proportions of variance explained, or R^2 (Pearson correlation). R^2 is displayed separately for each of the seven civil weighting models. Table I.6 shows the same statistics for each of the six criminal litigation areas. Technically, R^2 answers the question: "What proportion of the total variation among USAOs in FTEs for the litigation area is statistically accounted for by the workload predictors?" R^2 cannot be greater than

[8]See, e.g., G. Snedecor and W. Cochran, Statistical Methods, 7th ed. (Ames, Ia.: Iowa State U., 1990), ch. 5. Goodness-of-fit is only one criterion used to evaluate models. Theoretical cogency and explanatory relevance are thought to be more important criteria in many applications. That is, one generally asks: "How theoretically reasonable are the predictors as causal antecedents of the predicted variable?" and "Are the theoretical concepts measured by the predictors distinct?"

However, the purpose of the weighting models is to predict the average time expenditures required to process entire workloads rather than to estimate the individual contributions of specified kinds of cases and defendants to time expenditures. That is, the weighting models are "predictive" rather than "explanatory" (or "structural") models. Fortunately, the theoretical requirements are much less stringent for predictive models than for explanatory models. Correspondingly, goodness-of-fit is more important in evaluating the weighting models and other predictive models than in evaluating explanatory models.

unity or less than zero. Except in research using randomized experiments, R^2s greater than .50 are indicative of unusually good fit of the model to the data.[9]

Table I.5: Criminal Weighting Models' Proportions of Variance Explained (R^2s), by Criminal Litigation Area and Fiscal Year (n=91)

Criminal litigation area	Percent of FTEs (1989)	R^2 by fiscal year		
		1987	1988	1989
Public corruption	6.4	.76	.82	.84
Economic crime	29.6	.83	.86	.85
Organized crime	2.7	.70	.62	.54
Drug crime	36.8	.93	.95	.91
Violent crime	6.6	.69	.70	.79
Other crime	17.8	.69	.71	.69

Table I.6: Civil Weighting Models' Proportions of Variance Explained (R^2s), by Civil Litigation Area and Fiscal Year (n=91)

Civil litigation area	Percent of FTEs (1989)	R^2 by fiscal year		
		1987	1988	1989
Affirmative-monetary	20.0	.70	.64	.69
Affirmative-other	8.0	.75	.84	.85
Defensive-monetary	37.6	.81	.80	.82
Defensive-other	11.6	.96	.97	.94
Civil forfeitures	16.4	.90	.87	.91
Foreclosures	4.1	.73	.79	.68
Prisoner petitions	2.1	.47	.71	.68

By this criterion, tables I.5 and I.6 show that each of the six criminal and seven civil weighting models fit the data quite well. Goodness-of-fit (R^2) is especially high for litigation areas containing large proportions of reported FTES, such as drug crime in the criminal area and defensive-monetary in the civil area.

The R^2s in tables I.5 and I.6 are presented separately for 1987, 1988, 1989, even though, as discussed previously, the final weights were estimated by pooling the data for the 3 years and assuming that the time requirements of different kinds of cases did not change between 1987 and 1989.

---

[9]For appropriately defined variables x and y, R^2 is the square of the "Pearson correlation coefficient" R, which gauges the degree of association (correlation or dependence) between the two variables x and y. By definition, R cannot be less than -1 or greater than +1. Positive values indicate that the two variables are positively ("directly") associated (e.g., like annual rainfall and umbrella sales). See Snedecor and Cochran for further discussion. For understanding the sequel of this report, it is helpful to keep in mind that strongly positive values of R (say, between .5 and 1) are indicative of a substantial positive correlation between two variables.

Note that R^2 is imperfect as a measure of goodness-of-fit because, like all global measures of fit, it may conceal systematic departures of particular data points from the model. For this reason, we inspected "residual plots," i.e., plots of the deviations of predicted times from reported times for particular offices, before deciding upon the final weighting models. The residual plots revealed no patterns indicative of systematic lack-of-fit of the weighting models.

Alternative Interpretations of the Staffing Ratios

We interpret the ratio of predicted time and reported time of a USAO as gauging the degree of understaffing/overstaffing of that USAO relative to other USAOs. However, other interpretations of the staffing ratios are plausible:

1. Efficiency. Offices with high staffing ratios may be more efficient in processing cases than offices with low staffing ratios. That is, by working efficiently, offices with high ratios are able to process more cases than offices with low ratios in the same amount of time. Even so, the principle that new positions should be allocated disproportionately to offices with high staffing ratios seems advisable. It is reasonable to allocate more positions to relatively efficient offices than to relatively inefficient offices.

2. Quality of work. Offices with high staffing ratios might be doing poorer quality legal work because they are handling more cases and defendants per staff year. If so, and if quality is an overriding consideration, it may be unreasonable to allocate a disproportionately large number of positions to offices with high ratios. This is difficult to actually observe because the quality of legal work is hard to judge since it depends on detailed knowledge of laws and legal procedures and how they were applied in each case. Even experienced attorneys can disagree about whether a particular case or defendant has been well or poorly handled.

3. Differential bias due to omitted variables, coverage errors, measurement errors. As discussed above, the staffing ratios may be differentially biased by omitted variables, coverage error, and measurement error. These sources of potential bias might also be used as alternative interpretations of the ratios. An office might be incorrectly inferred to be understaffed relative to other offices because large amounts of time were spent on a particularly high profile trial, a variable that was omitted in the model. An office might be incorrectly inferred to be overstaffed because an unusually large proportion of cases or defendants

were not recorded (differential coverage error) or because an unusually large proportion of cases with trials were recorded as cases without trials (differential measurement error).

The alternative interpretations of the weighting model lay the basis for further research and point to the need for discretion in interpreting its results. Future research could seek to identify and incorporate any omitted variables and to improve the measurement and coverage of the data. Discretion should be exercised in interpreting results for offices inferred to have poor coverage or measurement. (See app. II.)

Above all, the results of the weighting model should be regarded as an aid to decisionmaking rather than as a substitute for experienced professional judgment. Clearly, the models take into account only the reported workloads and time expenditures of USAOs, which may differ from the ideal workloads and expenditures of the same offices. The maximum workload that can be handled by any attorney or office without compromising the quality of legal work is an additional question that these models do not address. The advantage of developing explicit models ultimately lies in the public character of the assumptions and data upon which decisions are made and, correlatively, in the prospect for improving decisionmaking.

Accuracy and Completeness of U.S. Attorneys' Criminal Workload Data

In this appendix we evaluate the quality of the U.S. Attorneys' data on the criminal workloads of AUSAs that we used to build our criminal workload model. We also assess the sensitivity of recommended allocations of new positions to errors in the workload data. In addition, we discuss how the effects of differences in data completeness among USAOs (i.e., differential coverage error) on allocations can be reduced by matching legal case records from two sources and applying a statistical technique called dual system estimation.

Concepts of Data Quality Evaluation: Consistency, Completeness, and Sensitivity

To evaluate the quality of the U.S. Attorneys' data, we looked at three statistical properties of the data—consistency, completeness (also called "coverage"), and sensitivity. These measures are commonly used in evaluating data quality.

By "consistency" or "reliability," we mean the extent to which two sources of data measuring the same characteristic or attribute agree. For example, if two attorney workload data files both record that a specified legal case had a trial during fiscal year 1989, we say that the two files yield consistent information for that case about the occurrence of a trial in 1989. If two data sources give the same answer for the vast majority of cases, we say the two sources are highly consistent or reliable for the characteristic being measured.[1]

By "completeness" or "coverage" we mean the extent to which a data set includes all of the elements, observations, or cases in a relevant universe. Evaluation of coverage, like evaluation of measurement, depends upon having two or more data sources for the same population. For example, if one data source can be assumed to be fully complete (i.e., coverage equals 100 percent) and if a second data source contains almost exactly the same set of cases as the first, we infer that coverage is high, albeit not perfect, in the second source. Even if neither data source can be assumed to be fully complete, coverage can still be evaluated using the technique of "dual system estimation" (see discussion later in this appendix), provided that the assumptions of this technique are reasonably well satisfied.

By "sensitivity," we mean the extent to which the results of an analysis are adversely affected by (sensitive to) errors in the data or in the

---

[1]Consistency, however, is not the same as accuracy. Rather, consistency is a necessary but not a sufficient condition for accurate measurement. This is because both data sources can be erroneous. In the absence of an absolute criterion of accuracy, there exists no infallible means of evaluating measurements.

assumptions. If a specified magnitude of error in a particular data item or assumption renders the results of an analysis highly inaccurate, we say the results are sensitive. If the same magnitude of error does not greatly affect the results, we say the results are insensitive, or "robust."

In order to evaluate the consistency, completeness, and sensitivity of the U.S. Attorneys' criminal workload data, we matched case records from two independent data sources: (1) the EO data set, i.e., the file maintained by the Executive Office of U.S. Attorneys, and (2) the AO data set, i.e., the file maintained by the Administrative Office of U.S. Courts.

Matching of AO and EO Case Records

This section presents overviews of (1) the definitions of universes of EO and AO cases that were included in the match and in the analyses of staffing imbalances, (2) the match rules that were used to determine whether an EO and an AO record "matched" (i.e., pertained to the same legal case), and (3) evaluation of the quality of the match. Technical documentation of these topics can be found in appendix VI.

Universes of Cases

For each of civil AUSA and criminal AUSA workloads in fiscal years 1987, 1988, and 1989, we defined the universe of relevant cases as all cases that were both relevant to the workloads of AUSAs in the fiscal year and "active" at some time during the fiscal year. Appendix VI documents the civil and criminal cases that were excluded from our evaluation because these cases were judged to be inactive or irrelevant to the work of AUSAs during the fiscal year.

Civil AUSA workloads in our analysis were estimated using the EO data alone. Criminal AUSA workloads were estimated using matched EO and AO data. Only criminal case records in the EO and AO files were matched as a part of our evaluation. Thus, the data quality evaluations in the following sections are restricted to criminal cases in the relevant universes for fiscal years 1987, 1988, and 1989. Since matters and appeal cases are not included in the AO data, these types of cases are also excluded from the evaluations. In the criminal as well as in the civil analyses of this report, matter and appeal workloads are estimated strictly from the EO data.

Match Rules ·

By a "matching variable" we mean a variable (data item) that is used to determine whether two data records, an EO record and an AO record, "match," i.e., pertain to the same case. In order for a variable to be used

as a matching variable, the variable must be measured in both data sources. Four matching variables were used in GAO's match of EO and AO case records:[2]

1. USAO (i.e., federal district court), 2. satellite court (i.e., local jurisdiction within the district), 3. court docket number, and 4. case filing date.

GAO's match rules used the four matching variables to classify unique pairs of case records, each pair consisting of a single EO record and a single AO record, as AO-EO matches. EO records for which no matching AO record could be located were classified as EO nonmatches. AO records for which no matching EO record could be located were classified as AO nonmatches. The first panel of table II.3 presents the three possible outcomes of the match rules in tabular format.

There are many possible match rules, each using the same four matching variables, that could have been used to classify EO and AO case records as matches and nonmatches. All such "multiple-variable" match rules must incorporate complex judgments about (1) the relative accuracy of alternative matching variables and (2) the relative costs of two kinds of matching errors, called false matches and false nonmatches.[3]

In brief, GAO's match rules are based on (1) detailed inspection of prospective AO and EO matching records in 20 USAOs and (2) hypotheses about the most plausible kinds of coding and keying errors, such as mispunching of the day, month, or year of case filing dates.

GAO's match rules initially classify an EO-AO pair as a match if and only if one or more of the following conditions are met:[4]

Rule 1. The EO record and the AO record exactly agree with respect to district court, satellite court (if any), and court docket number (suitably

---

[2]In order to use court docket number as a matching variable, it was necessary to transform, or recode, the docket numbers to have the same numerical format in all districts and both data files. This is because court docket numbers are recorded in different formats by the EO and the AO and by district courts within each data system. Unfortunately, even after transforming the data, it was not possible to uniquely identify cases on the basis of court docket number alone. This is the reason why four matching variables were needed to match EO and AO case records.

[3]A recent article by M. Jaro discusses the statistical properties of multiple-variable match rules. See M. Jaro, "Advances in Record-Linking Methodology as Applied to Matching the 1985 Census of Tampa, Florida," Journal of the American Statistical Association (June 1989), pp. 414-420.

[4]Duplicate records on each of the EO and AO files, defined as records with exactly the same values on all four matching variables as a record previously encountered in the sorted file, were deleted from the file prior to matching.

transformed), and no other record on either file has the same values of these three variables.

Rule 2. The EO record and the AO record exactly agree with respect to district court, satellite court (if any), and court docket number (suitably transformed), and the filing dates of the EO record and the AO record differ by no more than 31 days.

Rule 3. The EO record and the AO record exactly agree with respect to district court, satellite court (if any), and court docket number (suitably transformed), and the day and month of the EO and AO case filing dates are exactly the same.

Rule 4. The EO record and the AO record exactly agree with respect to district court, satellite court (if any), and court docket number (suitably transformed), and the day and year of the EO and AO case filing dates are exactly the same.

For each of fiscal years 1987, 1988, and 1989, we found that more than 90 percent of all EO-AO matches were identified using rule 1. If two or more EO-AO pairs involving the same EO (or AO) record were classified as a match using rules 2 through 4, the EO-AO pair with case filing dates that were most proximate in time was designated as the match. No single EO (or AO) record was allowed to form a match with more than one AO (or EO) record.

Quality of the Match

Two kinds of matching errors, called false matches and false nonmatches, affect the data quality evaluations of this chapter and the staffing inferences and allocations presented elsewhere in this report. A false match is an EO-AO pair that was designated as a match even though the EO record and the AO record actually pertain to different cases. A false nonmatch is an EO (or AO) record that was not matched to an AO (EO) record even though a record pertaining to the same case existed on the AO (EO) file.

False matches spuriously inflate (i.e., bias upward) estimates of measurement inconsistency and spuriously deflate (i.e., bias downward) estimates of data coverage. False nonmatches deflate estimates of inconsistency and inflate estimates of coverage. Both types of matching error can distort (bias) combined AO-EO estimates (dual system estimates) of legal workloads. Yet, as discussed in appendix I, such biases will affect

allocations of new positions to USAOs only to the extent that the magnitude of matching error varies among USAOs.

Since matching variables are always measured with error, match rules always engender false matches and nonmatches. The risk of a false match can be decreased only at the cost of increasing the risk of a false nonmatch, and vice versa. Thus, match rules must be designed to optimize, in some sense, the trade-off between the two kinds of error.[5] The design of match rules 1 through 4 reflected our expectation that, in matching EO and AO records, the risk of false nonmatches is greater than the risk of false matches. This would be true, for example, if coding or keying errors caused many cases to be included more than once, with different identifying information, in one or both files.

Even after substantially liberalizing the match rules, however, by adding rules 2 through 4 to rule 1, we found that more than 90 percent of the matches in each fiscal year were identified on the basis of rule 1 alone. Hence, the vast majority of matches were "preferred matches" in the sense that district court, satellite court, and docket number agreed exactly and no other record on either file had the same values as these three variables.

The adequacy of the match rules depends on the use to be made of the matched data file. The main use of the EO-AO match is to make improved estimates of the legal workloads of USAOs. It is therefore instructive that matching cases have a higher prevalence of time-intensive legal activities than nonmatching cases. For example, based on fiscal 1989 data, 52 percent of matched EO-AO cases had either a trial or an indictment during the fiscal year (according to the EO data), whereas only 13 percent of EO nonmatches had either a trial or an indictment during the fiscal year. Similar results were obtained using fiscal 1988 and fiscal 1987 data and using AO rather than EO indicators of trial and indictment. These results suggest that cases with large work requirements were especially likely to have been recorded on both data systems and to have been correctly identified as matches by GAO's match rules.

---

[5]M. Jaro, "Advances in Record-Linking Methodology as Applied to Matching the 1985 Census of Tampa, Florida."

Evaluation of Measurement Consistency

Table II.1 explains the technique that we used to evaluate measurement consistency. The top panel shows how data are organized for the evaluation, the cross-classification of two measurements of the same characteristic of the same unit. The bottom panel shows how the data format was used to evaluate data on whether a trial occurred during fiscal 1989. Comparing the top and bottom panels, we have a = 1,305 cases, b = 1,674 cases, and so on. The total number of EO-AO matches in fiscal 1989 equals n = a + b + c + d = 37,787 cases.

Table II.1: Evaluation of Measurement Inconsistency—Data Format (Cross-Classification of Two Measurements) And Application to EO and AO Indicators of a Trial During Fiscal 1989 (Matched EO-AO Nonappeal Cases, Excluding Matters)

Data format	First measurement		
	Yes	No	Total
Second measurement			
Yes	a	b	a + b
No	c	d	c + d
Total	**a + c**	**b + d**	**n = a + b + c + d**

Application	EO measurement		
	Trial	No trial	Total
AO measurement			
Trial	1,305 (3.5%)	1,674 (4.4%)	2,979 (7.9%)
No trial	606 (1.6%)	34,202 (90.5%)	34,808 (92.1%)
Total	**1,911 (5.1%)**	**35,876 (94.9%)**	**37,787 (100.0%)**

The evaluation of inconsistency was restricted to AO-EO matches because two measurements of case attributes are available only for matched cases. This implies that evaluations of inconsistency are affected by matching error. For example, false matches would spuriously inflate estimates of measurement inconsistency if two different cases were more likely to differ in their recorded attributes than two different measurements of the same case.

An additional factor affecting the evaluation of measurement inconsistency is discrepancy between the AO and EO definitions of case attributes. Table II.1 shows that 7.9 percent of matched EO-AO cases were trials according to AO, while only 5.1 percent were trials according to EO. The difference might be due to the use of a more restrictive definition of "trial" by EO than by AO rather than to measurement error. (App. VI discusses the EO and AO data file fields that were used to define comparable case attributes for the two data systems.)

The simplest gauge of measurement inconsistency is the percent incon-
sistent. In table II.1 the percent inconsistent for data format ("%") is
computed as % = 100 x (b + c)/n. For this application, it is computed as
% = 100 x (1,674 + 606)/37,787 = 4.4 + 1.6 = 6.0%. In words, about 6
percent of the cases had inconsistent information about the occurrence
of a trial in 1989.

Table II.2 displays the percent inconsistent for 11 data items in each of
1987, 1988, and 1989. All 11 items, with the exception of "plea during
year," were used in the understaffing analysis to gauge the complexity
of AUSA criminal workloads (see app. I). The "plea" variable was
dropped not only because, according to the AUSAs interviewed by GAO,
pleas are not especially time-intensive, but also because, as shown in
table II.2, "plea" is measured more inconsistently than other attributes.
In each year, more than 20 percent of matched cases had inconsistent
information about whether a plea occurred.

Table II.2 also presents an additional statistic called the Index of Incon-
sistency (I).[6] The purpose of I can be discerned from table II.2: compare
the marginal "percents in class" (i.e., percentages of cases having the
attribute in 1989, as estimated using the AO data), shown in the second
column, with the corresponding percentages inconsistent (%) for the
same year, shown in the seventh column. Generally, the closer the per-
cent in class to 50 percent (i.e., the further from 0 and from 100), the
higher the % is. This correlation suggests that % is not a pure gauge of
measurement inconsistency. On the contrary, % depends on the preva-
lence of the attribute in the population as well as upon the accuracy of
the measurement procedure.

The purpose of I is to compare inconsistency across different attributes
while adjusting the comparison for the prevalence of the attributes.
Based on I, table II.2 suggests that "indictment," "filing," "single defen-
dant," and "drug crime" are measured relatively well, while "felony,"
"trial," "public corruption," and "organized crime" are measured rela-
tively poorly.

---

[6]In terms of the symbols of table II.1, under "data format" the Index of Inconsistency is defined as

$$I = \frac{n(b + c)}{(a + b)(b + d) + (a + c)(c + d)}$$

See R. Groves, Survey Errors and Survey Costs (New York: Wiley, 1989), p. 326.

Trends in % and in I between 1987 and 1989 suggest modest improvements in measurement. For example, for "trial during year," % declines from 8 percent in 1987 to 6 percent in 1989, and I declines from 68 in 1987 to 50 in 1989. Yet, for most attributes, the changes in measurement inconsistency between 1987 and 1989 are statistically insignificant.

Table II.2: Percent Inconsistent (%) and Index of Inconsistency (I) for U.S. Attorneys' Criminal Workload Data Items (AO-EO Matched Cases (Nonappeals, Nonmatters), Fiscal Years 1987- 1989)

Data item	Percent in class[a] (AO, 1989)	Inconsistency, by fiscal year					
		1987 (n=29,341)		1988 (n=31,003)		1989 (n=37,787)	
		%	I	%	I	%	I
Plea during year	43	24	49	25	51	21	43
Felony	96	4	59	4	63	3	58
Indictment during year	51	7	14	7	13	5	9
Trial during year	8	8	68	8	54	6	50
Filing during year	67	5	3	2	4	2	1
Single defendant	76	6	15	6	15	5	13
Public corruption case	2	3	80	3	73	2	75
Economic crime case	29	17	39	16	40	14	35
Organized crime case	1	1	81	2	75	1	74
Drug crime case	33	3	7	3	8	3	8
Violent crime case	8	5	42	5	42	5	37

[a]This is the percentage of cases that were recorded as having the specified attribute in the fiscal year. For example, according to AO data, 43 percent of cases had a plea sometime during fiscal year 1989.

Evaluation of Data Completeness

Table II.3 explains the technique that we used to evaluate data coverage. The data format for this evaluation is as follows: on the basis of GAO's match rules, cases are classified as "in" or "out" of the AO data source and as "in" or "out" of the EO data source. Equivalently, each case record in the union of the two files is classified as an AO-EO match, an AO nonmatch, or an EO nonmatch. The frequency in the "in-in" cell is the number of times that two distinct case records, an AO record and an EO record, were inferred to represent the same case. The total number of AO case records is obtained by summing the number of AO-EO matches and the number of AO nonmatches. The total number of EO case records is obtained by summing the number of AO-EO matches and the number of EO nonmatches.

Without additional assumptions, it is impossible to make any evaluation of the coverage of either file. This is because the sum of AO-EO matches,

AO nonmatches, and EO nonmatches does not equal the total number of cases in the universe and because an estimate of total cases is needed to evaluate coverage.

In particular, the AO coverage probability is defined as the ratio of total AO cases and total cases in the universe. The EO coverage probability is defined as the ratio of total EO cases and total cases. The denominators of these ratios cannot be deduced without additional assumptions, because the number of cases that are excluded from both files (i.e., the number of cases in the "out-out" cell) is unknown.

Table II.3: Evaluation of Coverage—Data Format (Cross-Classification of Two Data Sources) And Application to Fiscal 1989 EO-AO Matched-Merged Caseload File (Nonappeal Cases, Excluding Matters)

| Data format | EO data | | | |
	In	Out	Total	
AO data				
In	Match	AO nonmatch	AO total	
Out	EO nonmatch	(unknown)	(unknown)	
Total	**EO total**	**(unknown)**	**(unknown)**	

| Application | EO data | | | Coverage probability |
	In	Out	Total	
AO data				
In	37,787	11,334	49,121	.71[a]
Out	12,771	(7,509)	(20,280)	
Total	**50,558**	**(18,843)**	**(69,401)**	
Coverage probability		.73[a]		

[a]Based on dual system estimation; see text.

Dual System Estimation

Dual system estimation is a technique for estimating the total number of cases in a universe, including cases excluded from both data sources, using data in the form of the data format section of table II.3. If the assumptions of the technique are reasonable, the estimate of total cases (called the dual system estimate (DSE)) can be used as the denominator of estimated coverage probabilities.

To fix the basic idea of DSE, consider an example. Suppose a game warden is interested in knowing the total number of bears in a preserve. On the first day, the warden captures 100 bears, ties a ribbon around the neck of each bear, and releases the bears. On the second day, the warden again captures 100 bears and observes how many of them have ribbons around their necks. The most extreme outcomes are instructive: if all of the bears captured on the second day have ribbons, the total

population is probably small (possibly no more than 100); if none have ribbons, the total population is probably large (at a minimum no less than 200).

Generally, the larger the number of ribboned bears that are recaptured (i.e., the larger the number of "matches") relative to the total number captured on the second day, the smaller is the inferred number of bears in the preserve. To be sure, assumptions are required for the validity of any such inference: if the ribbons did not allow all of the recaptured bears to be identified (say, because some bears ate their ribbons), the inference might be faulty. Alternatively, if ribboned bears were more likely to be captured on the second day than unribboned bears (say, because the ribbons impeded escape), the inference might be faulty.

The example illustrates the two principal assumptions of DSE: (1) the match rules allow all matches and nonmatches to be correctly identified (i.e., no matching errors, as discussed above) and (2) whether an element included in the second data set is not conditioned by whether the element was included in the first data set. (Technically, the second assumption is that inclusion in the second data set is independent of inclusion in the first data set.) Violations of the second assumption, called correlation biases, are especially plausible. If certain kinds of cases, such as cases requiring little time or attention, are likely to be omitted from both data sets, then DSEs of total population size are too low ("biased downward") and estimated coverage probabilities are too high ("biased upward").

The application section of table II.3 shows the computation of the DSE for the 1989 data. Entries whose computation depended upon dual system estimation are marked with note a. There were 37,787 AO-EO matches, 11,334 AO nonmatches, and 12,771 EO nonmatches. The DSE of the total number of cases in 1989 equals 69,401. Hence, the estimated AO coverage probability equals $(37,787 + 11,334)/69,401 = .71$, and the estimated EO coverage probability equals $(37,787 + 12,771)/69,401 = .73$. In other words, each source is inferred to be slightly more than 70 percent complete.

A technical note on the computation of DSEs: Because we expected violations of the second assumption ("independence"), DSEs were first computed separately within refined subclasses of cases, defined on the basis of litigation area, number of defendants, and legal proceedings during the fiscal year (see app. I), and then aggregated across subclasses to

yield the total DSEs shown in table II.3 and subsequent tables of this section. This stratification of the estimation by variables correlated with the probability of inclusion reduces bias because of violations of the second assumption.[7]

Estimated Coverage Probabilities

Table II.4 presents dual system estimates and estimated EO and AO coverage probabilities for fiscal years 1987, 1988, and 1989. Recall that dual system estimates were computed only for nonappeal criminal cases, excluding criminal matters, that were relevant to the work of AUSAS and active during the fiscal year.

Table II.4: EO Cases, AO Cases, EO-AO Matches, Percent Matched, Dual System Estimate, and Estimated EO and AO Case Coverage Probabilities (Criminal Nonappeal Cases Active During the Fiscal Year, Excluding Matters), **Fiscal Years 1987-1989**

Statistic		Fiscal year		
		1987	1988	1989
1.	EO cases	41,257	45,243	50,558
2.	AO cases	39,833	40,244	49,121
3.	EO-AO matched cases	29,339	31,003	37,787
4.	Percent of EO matched	71%	69%	75%
5.	Percent of AO matched	74%	77%	77%
6.	Dual system estimate	59,090	60,847	69,401
7.	EO coverage prob. = (1)/(6)	.70	.74	.73
8.	AO coverage prob. = (2)/(6)	.67	.66	.71

Row 6 shows that the number of nonappeal criminal cases increased modestly between 1987 and 1988, from 59,090 to 60,847, and more sharply between 1988 and 1989, from 60,847 to 69,401. Rows 7 and 8 show that the EO data are inferred to be slightly more complete than the AO data in each year.

The coverage of each data source improved modestly between 1987 and 1989, although neither trend is monotonic.[8] EO coverage increased from 70 percent in 1987 to 73 percent in 1989. AO coverage increased from 67 percent in 1987 to 71 percent in 1989.

Table II.5 presents 1989 DSEs and estimated coverage probabilities for each USAO. Data for Guam, Northern Mariana islands, and Virgin Islands

---

[7]For example, see K. Wolter, "Some Coverage Error Models for Census Data," Journal of the American Statistical Association, pp. 338-346 (June 1986).

[8]By monotonic we mean continually increasing or continually decreasing with the passage of time.

are given in table II.5 even though these districts were not included in our analysis of staffing imbalances (see app. I).

Table II.5 identifies USAOs where the coverage of one or both of the data systems appears to be relatively poor, districts that are candidates for targeted efforts to improve coverage. Excluding U.S. territories, the five lowest EO coverage probabilities are estimated for District of Columbia (.147), western Missouri (.219), Kansas (.337), Hawaii (.521), and middle Pennsylvania (.532). The five lowest AO coverage probabilities are estimated for Hawaii (.228), District of Columbia (.306), Kansas (.314), southern Georgia (.466), and western Missouri (.490). Recall that the poor estimated coverage of an office can result from inaccuracies in matching or other special problems in the data of the office.

Table II.5: Estimated Case Coverage Probabilities, Fiscal Year 1989 Criminal Data

U.S. Attorney office	EO[a]	AO[b]	Matches[c]	DSE[d]	EO coverage probabilities	AO coverage probabilities
Alaska	325	121	109	356	.911	.339
Alabama M	240	212	195	260	.923	.815
Alabama N	340	352	301	394	.862	.892
Alabama S	242	277	201	339	.719	.823
Arkansas E	237	242	221	264	.896	.915
Arkansas W	108	103	96	117	.922	.879
Arizona	1,351	1,172	1,044	1,599	.845	.733
California C	1,598	1,840	1,326	2,268	.704	.811
California E	794	710	593	1,207	.658	.588
California N	727	820	583	1,048	.693	.782
California S	1,442	1,697	1,151	2,186	.659	.776
Colorado	590	462	396	685	.861	.674
Connecticut	342	356	258	462	.740	.770
District of Columbia	384	798	58	2,604	.147	.306
Delaware	141	134	124	152	.928	.882
Florida M	1,162	1,146	976	1,369	.848	.837
Florida N	399	351	277	525	.759	.668
Florida S	2,292	2,021	1,795	2,596	.883	.778
Georgia M	210	218	141	335	.626	.650
Georgia N	881	739	597	1,092	.807	.677
Georgia S	284	159	129	341	.832	.466
Guam	13	155	9	159	.081	.972
Hawaii	560	245	149	1,075	.521	.228
Iowa N	251	217	187	295	.849	.734
Iowa S	159	146	134	172	.922	.846
Idaho	197	133	109	238	.827	.558
Illinois C	488	399	364	535	.912	.746
Illinois N	1,388	1,247	955	1,915	.725	.651
Illinois S	223	207	157	307	.727	.675
Indiana N	351	346	311	394	.891	.879
Indiana S	217	239	169	304	.714	.786
Kansas	416	387	138	1,234	.337	.314
Kentucky E	267	262	245	287	.929	.911
Kentucky W	329	256	224	381	.862	.671
Louisiana E	573	575	481	690	.831	.833
Louisiana M	90	65	55	103	.874	.631
Louisiana W	246	282	200	349	.704	.807
Massachusetts	457	566	361	725	.630	.780

(continued)

U.S. Attorney office	EO[a]	AO[b]	Matches[c]	DSE[d]	EO coverage probabilities	AO coverage probabilities
Maine	186	205	151	250	.742	.818
Maryland	859	645	522	891	.739	.724
Michigan E	927	856	689	1,175	.789	.728
Michigan W	220	234	194	266	.828	.881
Minnesota	434	433	375	501	.866	.864
Missouri E	424	445	385	490	.865	.908
Missouri W	471	1,052	360	2,145	.219	.490
Mississippi N	189	197	159	234	.806	.840
Mississippi S	346	318	273	417	.830	.762
Montana	294	271	242	332	.885	.816
North Carolina E	379	358	316	436	.868	.820
North Carolina M	438	376	338	483	.907	.778
North Carolina W	555	521	466	630	.881	.827
North Dakota	210	216	187	241	.871	.896
Nebraska	221	228	167	298	.742	.766
New Hampshire	66	62	49	87	.759	.713
New Jersey	640	716	513	945	.677	.757
New Mexico	770	690	607	889	.866	.776
Northern Mariana	3	24	3	24	.125	1.000
Nevada	657	549	435	875	.751	.627
New York E	1,265	1,279	1,043	1,568	.807	.816
New York N	300	238	154	436	.688	.546
New York S	1,635	1,755	1,428	2,025	.807	.867
New York W	345	342	267	454	.760	.754
Ohio N[e]	829	705	288	1,221	.678	.577
Ohio S	659	558	490	764	.863	.730
Oklahoma E	87	89	78	100	.873	.893
Oklahoma N	329	228	205	358	.919	.637
Oklahoma W	313	332	267	383	.817	.867
Oregon	639	629	539	749	.853	.840
Pennsylvania E	784	805	672	952	.824	.846
Pennsylvania M	268	265	129	503	.532	.526
Pennsylvania W	395	348	300	467	.845	.745
Puerto Rico	707	587	550	754	.938	.779
Rhode Island	131	125	98	170	.770	.735
South Carolina	669	651	568	769	.869	.846
South Dakota	294	278	256	322	.913	.863
Tennessee E	351	303	269	397	.884	.763
Tennessee M	399	335	284	489	.816	.685

(continued)

U.S. Attorney office	EO[a]	AO[b]	Matches[c]	DSE[d]	EO coverage probabilities	AO coverage probabilities
Tennessee W	716	580	461	905	.791	.641
Texas E	382	348	237	621	.615	.561
Texas N	1,143	1,017	863	1,344	.851	.757
Texas S	2,985	3,036	2,309	4,127	.723	.736
Texas W	1,682	1,662	1,340	2,151	.782	.773
Utah	431	418	299	626	.689	.668
Virginia E	655	818	564	1,029	.636	.795
Virginia W	280	244	219	310	.904	.787
Virgin Islands	679	444	304	921	.737	.482
Vermont	176	157	137	201	.874	.780
Washington E	540	655	426	836	.646	.783
Washington W	601	522	418	777	.774	.672
Wisconsin E	309	288	232	388	.797	.743
Wisconsin W	234	170	136	275	.852	.619
West Virginia N	382	355	329	409	.933	.867
West Virginia S[e]	422	370	299	451	.937	.821
Wyoming	139	132	109	173	.803	.763
Total	**50,558**	**49,121**	**37,787**	**69,401**	**.728**	**.708**

[a]EO equals number of U.S. Attorney case records included in the match.

[b]AO equals number of U.S. district court records included in the match.

[c]Matches equal number of EO and AO cases matching.

[d]DSE equals "dual system estimate" of total cases, using AO information for matches.

[e]1989 workload estimates for Ohio N were computed by multiplying 1988 estimates by 1.28; 1989 workload estimates for West Virginia S were computed by multiplying 1988 estimates by 1.09. These projections were necessary because of inability to match 1989 case records in Ohio N and West Virginia S. The adjustment factors were based upon estimated workload trends for 1987-88 in Ohio N and West Virginia S, respectively.

Allocation Results Are Robust Against Data Errors

In our resource allocation model, allocations of new positions to USAOs depend on the relative staffing levels that are estimated using the workload weighting model. Measurement errors, coverage errors, matching errors, errors due to omitted variables, and any other errors have no effect on the recommended allocations unless the magnitudes of these errors are variable among USAOs. That is, only differential error, error that is greater for some offices than for others, biases the recommended allocations.

It follows that using DSEs to estimate workloads should yield more efficient and cost-effective allocations than using AO data alone or EO data

alone. This is because DSEs adjust the data for differential coverage errors. Irrespective of the coverage of an office's EO or AO data file, the DSE estimates the total workload of the office. Yet, DSEs will give superior results only if the magnitudes of departures from the two principal assumptions of the technique, the assumption of no matching error and the assumption of independence, do not vary widely among USAOs.

We did a sensitivity analysis designed to investigate the effects of differential errors on allocations. First, we estimated the criminal weighting model of appendix I separately using each of three sets of estimated workloads: the AO estimated workloads, the EO estimated workloads, and the DSE (i.e., combined AO-EO) workloads. Second, we applied the allocation algorithm (see app. IV) to each of the three sets of estimated understaffing ratios. Each allocation assumed that 100 new AUSA positions were to be distributed among the 91 offices. Table II.6 shows the results of our analysis.

Table II.6 suggests that, although allocations to a few USAOs are significantly altered, the overall pattern of the allocation is robust against (insensitive to) the choice of workload estimates: the Pearson correlation between the DSE-based allocations and the AO allocations equals .80; the Pearson correlation between the DSE allocations and the EO allocations equals .85.

Table II.6: Comparison of Recommended Allocations Based Upon EO, AO, and DSE Estimates of the Criminal AUSA Workloads of USAOs in Fiscal Year 1989 (Assume That 100 New AUSA Positions Are to Be Allocated Among the 91 Offices)

Statistic	Allocation based upon		
	AO	EO	DSE
FTEs before	2,073	2,073	2,073
Positions allocated	100	100	100
FTEs after	2,173	2,173	2,173
Offices in allocation	91	91	91
Offices receiving one or more new positions	36	37	37
Standard deviation of allocations (n = 91)	1.84	1.93	1.81
Maximum allocation	8	9	7
Correlation with AO allocation	1.00	.77	.80
Correlation with EO allocation	.77	1.00	.85
Correlation with DSE allocation	.80	.85	1.00

Note: Allocations take into account 535 criminal AUSAs allocated so far in fiscal 1990. Guam, Northern Mariana islands, and Virgin Islands excluded because of insufficient data.

Since no match was carried out on the civil side, the allocations of civil AUSAs in appendix IV are based strictly on EO workload data. The high

correlation of EO and DSE allocations in table II.6 suggests that using EO workload estimates yields reasonable, if not optimal, allocations.

Using the Workload Weighting Models to Assess Resource Disparities

In this appendix we apply the weighting model presented in appendix I to the data sources described in appendix II in order to make inferences about resource disparities among the U.S. Attorney offices in fiscal year 1989. We also analyze changes in criminal and civil AUSA staffing levels and changes in the level and pattern of inequality among offices between 1987 and 1989. In the criminal analyses, our inferences are based on reported FTEs and on dual system estimates (see app. II) of criminal AUSA workloads. In the civil analyses, our inferences are based on reported FTEs and on EO estimates of civil AUSA workloads.

Resource Disparities Exist Among the U.S. Attorney Offices

In fiscal year 1989, the U.S. Attorneys handled about 147,000 criminal cases and matters and 230,000 civil cases and matters.[1] To process this workload, they spent about 1,538 attorney workyears on criminal investigations and prosecutions and 700 attorney workyears on civil litigation.[2] Based on the criminal weighting model applied to fiscal year 1989 data, 43 offices spent significantly less attorney time than predicted for their criminal workloads and 22 spent significantly more time than predicted. Based on the civil weighting model applied to fiscal year 1989 data, 35 offices spent significantly less time than predicted for their civil workloads and 37 spent significantly more time than predicted.

When an office spends significantly less attorney time than predicted, we infer that the office is understaffed relative to other offices. Conversely, spending significantly more attorney time than predicted may indicate that an office is overstaffed or inefficient when compared with other offices. (App. I gives alternative interpretations of the ratios of predicted and reported time.)

Inferences about the staffing or efficiency of any office are relative to the performances of other offices. In particular, the inferences do not mean that those offices spending more time than predicted could not effectively utilize additional resources.

Moreover, some offices that are inferred to be relatively overstaffed may be affected by exceptional circumstances that are imperfectly reflected in GAO's workload weighting models. For example, an office

---

[1] Excludes cases that involve little or no involvement by AUSAs (see app. VI for classes of criminal and civil cases dropped from our analysis).

[2] The reported workyears are in full time equivalents (FTE) and exclude FTEs reported for management and administration and appeal litigation.

currently handling unusually complex and time-intensive defense procurement fraud and/or bank fraud cases might be incorrectly inferred to be relatively overstaffed simply because the demands of such unusual cases are not fully captured in the model. Identifying and measuring such "omitted variables" would be one way to improve the model.

Improving the model may require not only improving current theories and data (see app. II) but also collecting additional data. For example, data on the time attorneys spend before grand juries to develop the evidence for criminal indictments are not currently collected. These activities are time consuming, especially in the investigative stage, but are perhaps infrequent relative to other activities. According to Justice officials, bank fraud and procurement fraud investigations have high priority but are especially difficult to prosecute. As a result, these fraud investigations do not result in indictments as frequently as other criminal investigations.

Reported and Predicted Time: Criminal and Civil AUSAs, Fiscal Year 1989

Tables III.1 and III.2 present, for criminal and civil AUSAs, respectively, the reported times, predicted times, and staffing ratios of the 91 USAOs in fiscal year 1989. Also shown are the weighted caseload (see app. I) and inference. Inference pertains to whether, according to the conventional level of statistical significance (i.e., alpha = .05), the predicted time of the office is significantly greater than the reported time (more than), significantly less (less than), or insignificantly different from the reported time (no inference).

The average ratio of predicted and reported times of the 91 offices during the period 1987-89 (i.e., the range of years used to estimate the criminal and civil weighting models; see app. I) equals 1.0. Offices with staffing ratios significantly greater than 1.0 are inferred to be understaffed or efficient relative to this average. Offices with staffing ratios significantly less than 1.0 are inferred to be overstaffed or inefficient relative to this average. More specifically, if the lower bound of the 95-percent confidence interval for the staffing ratio exceeds 1.0, we infer that the office is relatively understaffed or efficient. If the upper bound of this confidence interval is less than 1.0, we infer that the office is relatively overstaffed or inefficient. If the 95-percent confidence interval covers 1.0, no inference is supported.

Table III.1: Analysis of the U.S. Attorneys' Criminal Caseload and Attorney Resource Allocation, Fiscal Year 1989

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Time-expended inference[b]
Alaska	494	6.80	5.09	.749	More than
Alabama M	719	5.35	7.41	1.385	Less than
Alabama N	1,307	15.75	13.46	.855	More than
Alabama S	963	6.99	9.92	1.419	Less than
Arkansas E	726	10.86	7.48	.688	More than
Arkansas W	502	3.17	5.17	1.630	Less than
Arizona	2,982	30.87	30.72	.995	No inference
California C	4,957	65.81	51.07	.776	More than
California E	1,955	17.37	20.14	1.159	No inference
California N	2,892	29.54	29.79	1.009	No inference
California S	2,919	39.30	30.07	.765	More than
Colorado	1,544	20.87	15.91	.762	More than
Connecticut	1,112	20.28	11.46	.565	More than
Dist. of Columbia	3,402	40.17	35.05	.873	More than
Delaware	494	4.26	5.09	1.194	No inference
Florida M	3,676	35.48	37.87	1.067	No inference
Florida N	1,234	10.98	12.71	1.158	Less than
Florida S	6,001	66.57	61.82	.929	More than
Georgia M	981	6.16	10.11	1.641	Less than
Georgia N	2,480	24.27	25.55	1.053	No inference
Georgia S	724	7.37	7.46	1.013	No inference
Hawaii	1,251	9.76	12.89	1.321	Less than
Iowa N	797	4.73	8.21	1.737	Less than
Iowa S	662	4.15	6.82	1.642	Less than
Idaho	630	4.47	6.49	1.451	Less than
Illinois C	1,088	10.98	11.21	1.021	No inference
Illinois N	5,723	59.22	58.96	.996	No inference
Illinois S	810	9.19	8.34	.908	No inference
Indiana N	1,047	9.72	10.79	1.110	Less than
Indiana S	1,026	12.23	10.57	.864	More than
Kansas	1,433	7.29	14.76	2.025	Less than
Kentucky E	922	10.29	9.50	.923	No inference
Kentucky W	1,085	8.42	11.18	1.328	Less than
Louisiana E	1,744	17.26	17.97	1.041	No inference
Louisiana M	591	5.43	6.09	1.122	No inference
Louisiana W	1,165	9.38	12.00	1.279	Less than
Massachusetts	1,903	25.87	19.60	.757	More than

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Time-expended inference[b]
Maine	711	8.03	7.32	.911	No inference
Maryland	2,004	28.78	20.64	.717	More than
Michigan E	3,095	35.42	31.88	.900	More than
Michigan W	748	9.98	7.71	.773	More than
Minnesota	1,410	15.61	14.53	.931	No inference
Missouri E	1,486	20.42	15.31	.750	More than
Missouri W	1,474	15.40	15.19	.987	No inference
Mississippi N	711	5.08	7.32	1.441	Less than
Mississippi S	982	8.07	10.12	1.254	Less than
Montana	673	4.59	6.93	1.510	Less than
North Carolina E	1,077	9.07	11.10	1.224	Less than
North Carolina M	1,042	5.81	10.73	1.847	Less than
North Carolina W	1,207	6.72	12.43	1.850	Less than
North Dakota	7,681	4.28	7.02	1.640	Less than
Nebraska	785	6.99	8.09	1.158	Less than
New Hampshire	472	3.96	4.86	1.228	Less than
New Jersey	3,530	44.50	36.37	.817	More than
New Mexico	1,526	10.44	15.72	1.505	Less than
Nevada	1,362	15.92	14.03	.881	More than
New York E	4,041	39.73	41.63	1.048	No inference
New York N	1,346	11.79	13.87	1.176	Less than
New York S	4,959	91.06	51.09	.561	More than
New York W	1,598	13.90	16.46	1.184	Less than
Ohio N	2,597	24.46	26.75	1.094	Less than
Ohio S	2,029	16.14	20.90	1.295	Less than
Oklahoma E	478	2.87	4.92	1.715	Less than
Oklahoma N	724	8.43	7.46	.884	No inference
Oklahoma W	1,244	9.14	12.82	1.403	Less than
Oregon	1,342	15.55	13.83	.890	More than
Pennsylvania E	2,811	46.36	28.96	.625	More than
Pennsylvania M	1,326	10.27	13.66	1.331	Less than
Pennsylvania W	1,446	18.02	14.90	.827	More than
Puerto Rico	1,239	10.57	12.76	1.207	Less than
Rhode Island	516	6.11	5.32	.871	No inference
South Carolina	1,618	12.31	16.67	1.354	Less than
South Dakota	750	5.54	7.73	1.395	Less than
Tennessee E	1,127	11.41	11.61	1.017	No inference
Tennessee M	965	7.48	9.94	1.329	Less than
Tennessee W	1,892	11.78	19.49	1.655	Less than

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Time-expended inference[b]
Texas E	1,278	11.68	13.17	1.128	Less than
Texas N	2,531	25.76	26.07	1.012	No inference
Texas S	4,043	41.79	41.65	.997	No inference
Texas W	3,334	26.43	34.35	1.300	Less than
Utah	1,057	9.80	10.89	1.111	Less than
Virginia E	2,560	31.57	26.37	.835	More than
Virginia W	727	7.03	7.49	1.065	No inference
Vermont	556	6.15	5.73	.932	No inference
Washington E	579	6.30	5.96	.946	No inference
Washington W	2,159	20.43	22.24	1.089	Less than
Wisconsin E	1,363	12.64	14.04	1.111	Less than
Wisconsin W	592	5.15	6.10	1.185	Less than
West Virginia N	978	6.83	10.07	1.474	Less than
West Virginia S	1,794	13.07	18.48	1.414	Less than
Wyoming	555	4.71	5.72	1.215	Less than
Total	**147,067**	**1,537.93**	**1,515.12**		

[a]Ratio of predicted and reported time expenditures.

[b]Time-expended inference determined using 95-percent confidence interval.

Table III.2: Analysis of the U.S. Attorneys' Civil Caseload and Attorney Resource Allocation, Fiscal Year 1989

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Time-expended inference[b]
Alaska	1,383	3.34	3.74	1.119	Less than
Alabama M	1,265	3.68	3.42	.929	No inference
Alabama N	1,997	10.52	5.40	.513	More than
Alabama S	1,657	3.17	4.48	1.413	Less than
Arkansas E	2,160	4.31	5.84	1.355	Less than
Arkansas W	1,317	3.06	3.56	1.163	Less than
Arizona	2,766	14.48	7.48	.517	More than
California C	9,271	33.12	25.07	.757	More than
California E	3,628	14.60	9.81	.672	More than
California N	5,736	20.71	15.51	.749	More than
California S	6,105	14.39	16.51	1.148	Less than
Colorado	2,574	7.37	6.96	.945	No inference
Connecticut	2,015	5.96	5.45	.914	More than
Dist. of Columbia	7,581	22.00	20.50	.932	More than
Delaware	1,191	1.92	3.22	1.680	Less than
Florida M	3,395	8.90	9.18	1.032	No inference
Florida N	1,660	4.48	4.49	1.003	No inference
Florida S	6,693	21.77	18.10	.832	More than
Georgia M	1,668	4.57	4.51	.986	No inference
Georgia N	4,031	11.22	10.90	.972	No inference
Georgia S	1,272	5.45	3.44	.631	More than
Hawaii	1,868	3.99	5.05	1.265	Less than
Iowa N	1,538	3.84	4.16	1.084	No inference
Iowa S	1,557	4.94	4.21	.852	More than
Idaho	1,524	5.40	4.12	.763	More than
Illinois C	1,331	3.61	3.60	.998	No inference
Illinois N	4,696	14.91	12.70	.852	More than
Illinois S	1,694	3.92	4.58	1.169	Less than
Indiana N	1,945	4.11	5.26	1.280	Less than
Indiana S	2,725	5.11	7.37	1.442	Less than
Kansas	2,208	10.35	5.97	.576	More than
Kentucky E	2,100	8.52	5.68	.667	More than
Kentucky W	1,856	5.64	5.02	.891	More than
Louisiana E	2,034	10.24	5.50	.537	More than
Louisiana M	1,084	2.38	2.93	1.233	Less than
Louisiana W	2,977	5.13	8.05	1.569	Less than
Massachusetts	4,796	10.88	12.97	1.192	Less than

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Time-expended inference[b]
Maine	1,357	3.90	3.67	.940	No inference
Maryland	2,193	6.39	5.93	.928	No inference
Michigan E	3,731	13.75	10.09	.734	More than
Michigan W	2,219	5.71	6.00	1.050	No inference
Minnesota	2,027	6.06	5.48	.904	More than
Missouri E	2,271	4.21	6.14	1.458	Less than
Missouri W	1,912	8.18	5.17	.632	More than
Mississippi N	1,217	4.55	3.29	.723	More than
Mississippi S	1,424	5.78	3.85	.666	More than
Montana	1,394	3.94	3.77	.957	No inference
North Carolina E	1,476	5.27	3.99	.757	More than
North Carolina M	1,335	2.30	3.61	1.570	Less than
North Carolina W	1,257	1.89	3.40	1.799	Less than
North Dakota	1,302	2.99	3.52	1.176	Less than
Nebraska	1,535	5.26	4.15	.789	More than
New Hampshire	987	1.94	2.67	1.375	Less than
New Jersey	4,918	12.93	13.30	1.029	No inference
New Mexico	1,916	8.12	5.18	.638	More than
Nevada	1,757	4.68	4.75	1.014	No inference
New York E	7,359	24.92	19.90	.799	More than
New York N	2,038	3.80	5.51	1.450	Less than
New York S	7,008	29.52	18.95	.642	More than
New York W	2,400	5.72	6.49	1.135	Less than
Ohio N	3,750	13.22	10.14	.767	More than
Ohio S	3,339	10.53	9.03	.857	More than
Oklahoma E	1,169	2.06	3.16	1.534	Less than
Oklahoma N	1,424	2.92	3.85	1.317	Less than
Oklahoma W	2,223	8.36	6.01	.718	More than
Oregon	1,757	7.25	4.75	.655	More than
Pennsylvania E	4,574	13.19	12.37	.938	No inference
Pennsylvania M	1,679	3.53	4.54	1.287	Less than
Pennsylvania W	2,762	5.45	7.47	1.371	Less than
Puerto Rico	2,836	8.19	7.67	.936	No inference
Rhode Island	1,246	2.68	3.37	1.258	Less than
South Carolina	2,762	11.14	7.47	.671	More than
South Dakota	1,298	4.75	3.51	.738	More than
Tennessee E	1,908	4.39	5.16	1.176	Less than
Tennessee M	1,457	3.78	3.94	1.041	No inference
Tennessee W	2,322	5.33	6.28	1.177	Less than

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Time-expended inference[b]
Texas E	1,731	7.03	4.68	.666	More than
Texas N	4,508	13.49	12.19	.904	No inference
Texas S	5,447	17.85	14.73	.825	More than
Texas W	2,163	11.41	5.85	.513	More than
Utah	1,538	5.50	4.16	.757	More than
Virginia E	2,463	11.83	6.66	.563	More than
Virginia W	1,598	2.45	4.32	1.763	Less than
Vermont	1,228	2.12	3.32	1.567	Less than
Washington E	1,379	3.55	3.73	1.051	No inference
Washington W	3,284	7.92	8.88	1.122	Less than
Wisconsin E	1,916	4.29	5.18	1.208	Less than
Wisconsin W	1,712	3.65	4.63	1.268	Less than
West Virginia N	1,276	1.63	3.45	2.118	Less than
West Virginia S	1,812	3.73	4.90	1.315	Less than
Wyoming	1,187	2.63	3.21	1.221	Less than
Total	**230,068.9**	**699.64**	**622.15**		

[a]Ratio of predicted and reported time expenditures.

[b]Time-expended inference determined using 95-percent confidence interval.

Trends in the Ratio of Predicted and Reported Time, Fiscal Years 1987 Through 1989

As part of our evaluation of staffing imbalances, we also looked at trends in staffing ratios during fiscal years 1987 through 1989. Trends in the ratio depend upon both the need for attorney services, estimated using predicted time, and the availability of attorney services, estimated using reported time.

Table III.3 shows, separately for civil and criminal litigation, the reported times, predicted times, and ratios of predicted and reported times of the combined 91 U.S. Attorney offices in fiscal years 1987, 1988, and 1989. During this period, the predicted time expenditures of both civil and criminal attorneys increased.[3] Assuming that data quality and performance of the weighting model were constant from 1987 through 1989, these increases in predicted time reflect the increased need for civil and criminal attorney services during 1987 through 1989. During the same period, the reported time expenditures of both civil and

[3]As discussed in appendix I, the predicted times were estimated by applying the weighting models to all 3 years of data while constraining the estimated case weights to be equal across years.

criminal attorneys generally increased more rapidly than the corresponding predicted time expenditures. These increases in reported time reflect the increase in the total number of AUSAS.

Table III.3: Reported Times, Predicted Times, and Staffing Ratios: Criminal and Civil AUSAs, Fiscal Years 1987, 1988, and 1989

Fiscal year	Civil litigation			Criminal litigation		
	Reported	Predicted	Ratio	Reported	Predicted	Ratio
1987	603.2	604.9	1.003	1,367.3	1,401.7	1.025
1988	650.0	615.9	0.947	1,444.6	1,422.0	0.984
1989	699.7	622.2	0.889	1,537.9	1,515.2	0.985

For both criminal and civil AUSAs, the more rapid increase of reported than of predicted time occasioned an improvement in the overall staffing level of the 91 offices. For civil AUSAs, the overall staffing ratio declined from 1.00 in 1987 to 0.89 in 1989. For criminal AUSAs, the ratio declined from 1.03 in 1987 to 0.99 in 1989.

The amount of disparity among offices declined modestly during 1987 through 1989. The estimated standard deviation of the staffing ratios declined from .44 in 1987 to .33 in 1989 for civil AUSAs, and from 1.21 in 1987 to 1.13 in 1989 for criminal AUSAs.

Notwithstanding the improvement in overall staffing and modest reduction in staffing disparities among offices, the allocations of new attorney positions to USAOs during 1987 through 1989 did not significantly alter the pattern of staffing disparities among offices. Offices that were relatively "poor" (or "rich") in legal resources in 1989 tended to be the same offices that had been "poor" (or "rich") in 1987. The R^2, or Pearson correlation (based on 91 U.S. Attorney office observations), between civil ratios in 1987 and civil ratios in 1989 equals .71. The analogous correlation between criminal ratios in 1987 and 1989 equals .80. Both correlations are close to their theoretical maximum of 1.0.

Resource Allocation Model

In this appendix we discuss the criteria we developed for allocating new attorney positions to the USAOs and present several illustrative allocations. By presenting a number of allocations, we demonstrate the flexibility of the proposed technique.

Given the estimated staffing ratios of USAOs (see app. III), a simple mathematical algorithm, called the allocation model, can be used to generate a more equal allocation of attorneys to the USAOs. This algorithm operates to equate, as closely as possible, the post-allocation staffing ratios of the offices, i.e., the ratios of predicted and reported time after the allocation. As discussed below, the allocation model reduces disparities among offices and, if a certain assumption (called "isoelastic shifts in productivity") is satisfied, increases the economic efficiency of the distribution of attorneys among USAOs.

Each allocation presented in this appendix corresponds to an increase in the number of attorney positions funded by Congress in either fiscal year 1989 or fiscal year 1990. For example, in fiscal 1989, Congress appropriated money for 423 new criminal AUSAs for the purpose of fighting the war on drugs. The allocation model's distribution of the 423 drug crime positions is positively correlated with the allocation made by the Department of Justice. However, this appendix shows some important differences between the two allocations.

In contrast to the staffing inferences of appendix III, the allocations of this appendix disregard chance variations in the estimated staffing ratios of offices. That is, even though the estimated staffing ratio of any office may differ insignificantly from one according to the statistical tests of appendix III, the estimated staffing ratio of each office is still taken, for the purpose of the allocation, as the best "point estimate" (i.e., single-number guess) of the staffing need of the office. More advanced allocation models might incorporate information about the degrees of uncertainty of the estimated staffing ratios due to chance variations.

The allocations also assume that any factors affecting the workloads of offices that were omitted, imperfectly measured, or inadequately incorporated in the workload models of appendix I do not bias (i.e., systematically distort) the estimated staffing ratios. As discussed in appendix I, there might be relevant factors that were omitted or imperfectly measured in the workload models and that might be usefully included or gauged with greater precision in more advanced models. More advanced workload models might also test whether it is helpful to (1) treat

attorney time as both cause and consequence of workloads (since attorneys may play a role in generating their own work) and (2) allow for joint dependence (joint endogeneity) of attorney workloads with the workloads of judges, investigators, or other legal agents.

In summary, since the allocation model uses point estimates of the staffing ratios of offices as input, proper use of the allocations depends upon alertness to possible violations of the assumptions of the workload model. More advanced measurement and modelling techniques might yield different estimates of the staffing ratios of offices and hence different allocations.

Economic Theory for Allocation

The marginal product of any input factor is defined in economics as the extra output added by one more unit of that factor, while other input factors are held constant. In our application, the input factor is attorney time, and one unit of the input factor equals one AUSA FTE. The output is "legal production" or "work."

The marginal product of attorney time depends on the supply of other input factors. We think there are two additional input factors that affect the marginal productivity of attorney time:

1. Other legal inputs. These inputs include investigative agents, paralegals, judges, and also physical inputs, such as courtrooms and prisons.

2. Civil and criminal action. How many crimes of various kinds are committed for which evidence is available for prosecution? The number of crimes (and civil suits) does not appear to be subject to control by the government. In economics, such factors are said to be "exogenously determined."

Economics teaches that if an input, in this case new attorneys, is increased, the marginal product of an additional attorney must eventually decline. This "law of diminishing returns" will operate even if other legal inputs, such as judges and investigative agents, are added proportionately, provided that civil and criminal cases grow less than proportionately. If so, each attorney that is added will have fewer cases to work with and will therefore yield a smaller product.

This theory has direct implications for how additional legal resources should be allocated among the USAOs: How should new positions be allocated among the offices in order to increase the overall efficiency of

legal production? In other words, how should the positions be allocated in order to yield the greatest amount of legal product per additional FTE of attorney time?

The economic answer is that the new positions should be allocated so as to equate, as closely as possible, the marginal products of attorney time in the different offices. This implies that new positions should initially be allocated to those offices where the marginal productivity of attorney time is relatively high.[1]

Allocation Model

Given staffing ratios estimated using the workload model, a simple algorithm can generate an allocation of additional attorney positions to U.S. Attorney offices that is as equitable as possible, given the preallocation distribution of attorneys and the number of new positions to be allocated. The allocation model uses as input (1) the estimated staffing ratios of the offices and (2) the total number of new positions that are available to be allocated.

The algorithm allocates the new positions one at a time. At each stage, an additional position is given to the office that has the highest current staffing ratio, based upon (1) the distribution of attorneys among offices before the beginning of the allocation and (2) the distribution of positions that were allocated at previous stages of the allocation. After each successive position is allocated, the model recalculates the staffing ratios for all offices. This is done by adding one FTE to the reported time (i.e., to the denominator of the staffing ratio) of the office receiving the last allocated position before dividing predicted times by reported times. The next position is given to the USAO that has the highest degree of understaffing (efficiency) on the basis of these recalculated ratios. The iterative process continues until all new positions have been allocated.

The algorithm has two important properties: (1) the number of AUSAs in each U.S. Attorney office increases, or at least does not decrease, and (2) postallocation disparities among offices in the staffing ratios are made as small as possible, given the preallocation distribution of attorneys among offices and the number of new positions available to be allocated.

---

[1]See, e.g., P.A. Samuelson, Economics, sixth ed., p. 514.

In general, equating the staffing ratios of offices is not equivalent to equating the marginal products,[2] and, as discussed in the preceding section, marginal product is the quantity that, according to economic theory, must be equated among offices in order to maximize economic efficiency. To establish equivalence, we must assume that, at any possible staffing level, the future product per attorney ("average product") of any office is strictly a function of the future staffing ratio (i.e., not a function of any other characteristic of the office) and that, at any possible staffing level, the same functional relationship between future average product and future staffing ratio holds for every office.[3] If this assumption (called isoelastic shifts) holds, then equating the postallocation staffing ratios of the offices is equivalent to equating the marginal products of the offices.[4] Hence, if the assumption of isoelastic shifts is met (and if the assumptions of the workload model also hold), GAO's allocation model will yield a more efficient distribution of legal resources.

The assumption of isoelastic shifts seems plausible in its application to legal production because the work demands of particular legal cases are often idiosyncratic, and hence the individual attorneys in any office must exercise considerable autonomy in carrying out their assignments. If so, the overall volume of work in an office and the total number of attorneys in the office are likely to have much smaller effects on the future average product than the ratio of the two, i.e., the staffing ratio.

---

[2]P.A. Samuelson, Economics.

[3]Technically, we assume that the functional relationships between the average attorney products and future staffing levels of different offices are isoelastic shifts (or horizontal multiples) of each other. For example, suppose that office A has workload $W_a = 2$, office B has workload $W_b = 1$, and office C has workload $W_c = .5$. (Assume that none of the offices currently have any attorneys assigned to them or, equivalently, that any previous allocations are irrelevant.) The assumption of isoelastic shifts stipulates that if 20 attorneys were assigned to A, 10 attorneys were assigned to B, and 5 attorneys were assigned to C, then the average product of attorneys in the three offices would be the same. Similarly, if 40 attorneys were assigned to A, 20 attorneys were assigned to B, and 10 attorneys were assigned to C, the average products in the offices would be the same. In general, irrespective of the total number of attorneys to be assigned, the future average products of the offices would be equal if, and only if, the number of attorneys assigned to the offices were proportional to (i.e., a constant multiple of) the workloads of the offices.

[4]Let R_i denote the staffing ratio of office i. Then $R_i = W_i/N_i$, where W_i is the "objective" workload of office i (estimated using the workload model of app. I) and N_i is the number of attorneys assigned to office i. By the assumption of isoelastic shifts, the average product of office i, say AP_i, can be written as $AP_i = f(R_i)$, where the function f does not depend on the office. The total product of office i, say TP_i, is given by $TP_i = N_i f(R_i) = N_i f(W_i/N_i)$. The marginal product of office i, say MP_i, is obtained by differentiating TP_i with respect to N_i: $MP_i = f(R_i) - R_i f'(R_i)$.

It follows that equating the R_is of offices also equates their MP_is. For further discussion, see, e.g., J. Robinson, The Economics of Imperfect Competition, MacMillan (1964); R. Leftwich, The Price System and Resource Allocation, 6th ed., Dryden (1976), ch. 9; R. Awh, Microeconomics, Wiley (1976), app. 14.

Yet, we cannot rule out the possibility that economies of scale, or other factors associated with the volume of work or the size of an office (such as the organization of work in the office or opportunities for collaboration with other attorneys), also affect attorney productivity.

Allocations Using GAO's Model

Tables IV.1 and IV.2 present two allocations, one pertaining to criminal AUSAs and the other to civil AUSAs. These allocations assume that preallocation positions are not reallocated. Table IV.1 presents the allocation of 535 new positions to two litigation areas by Congress in 1989—drug crime (423 positions) and bank fraud (112 positions). Table IV.2 illustrates an allocation of 100 new positions to civil litigation. The first allocation is based on the estimated criminal staffing ratios in 1989, and the second allocation is based on the estimated civil staffing ratios in 1989.

The tables show, for each of the 91 offices included in our analysis, the staffing ratios before and after the simulated allocation. Comparing the before and after ratios indicates that the number of new positions in each allocation largely eliminated disparities among offices in the staffing ratios of AUSAs.

Table IV.1: Allocation of 535 Criminal Attorneys to Handle Fiscal Year 1989 Workload Using GAO's Model

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio before allocation[a]	Attorneys added	Staffing ratio after allocation[a]
Alaska	494	6.80	5.09	.749	0	.749
Alabama M	719	5.35	7.41	1.385	5	.716
Alabama N	1,306	15.75	13.46	.855	2	.758
Alabama S	963	6.99	9.92	1.419	7	.709
Arkansas E	726	10.86	7.48	.688	0	.688
Arkansas W	502	3.17	5.17	1.630	4	.721
Arizona	2,982	30.87	30.72	.995	10	.752
California C	4,957	65.81	51.07	.776	2	.753
California E	1,955	17.37	20.14	1.159	10	.736
California N	2,892	29.54	29.79	1.009	10	.753
California S	2,919	39.30	30.07	.765	1	.746
Colorado	1,544	20.87	15.91	.762	1	.727
Connecticut	1,112	20.28	11.46	.565	0	.565
Dist. of Columbia	3,402	40.17	35.05	.873	6	.759
Delaware	494	4.26	5.09	1.194	3	.701
Florida M	3,676	35.48	37.87	1.067	15	.750

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio before allocation[a]	Attorneys added	Staffing ratio after allocation[a]
Florida N	1,234	10.98	12.71	1.158	6	.749
Florida S	6,001	66.57	61.82	.929	15	.758
Georgia M	981	6.16	10.11	1.641	8	.714
Georgia N	2,480	24.27	25.55	1.053	10	.746
Georgia S	724	7.37	7.46	1.013	3	.719
Hawaii	1,251	9.76	12.89	1.321	8	.726
Iowa N	797	4.73	8.21	1.737	7	.700
Iowa S	662	4.15	6.82	1.642	5	.745
Idaho	630	4.47	6.49	1.451	5	.685
Illinois C	1,088	10.98	11.21	1.021	4	.748
Illinois N	5,723	59.22	58.96	.966	19	.754
Illinois S	809	9.19	8.34	.908	2	.745
Indiana N	1,047	9.72	10.79	1.110	5	.733
Indiana S	1,026	12.23	10.57	.864	2	.743
Kansas	1,433	7.29	14.76	2.025	13	.727
Kentucky E	922	10.29	9.50	.923	3	.715
Kentucky W	1,085	8.42	11.18	1.328	7	.725
Louisiana E	1,744	17.26	17.97	1.041	7	.741
Louisiana M	591	5.43	6.09	1.122	3	.722
Louisiana W	1,165	9.38	12.00	1.279	7	.733
Massachusetts	1,902	25.87	19.60	.757	0	.757
Maine	710	8.03	7.32	.911	2	.730
Maryland	2,003	28.78	20.64	.717	0	.717
Michigan E	3,094	35.42	31.88	.900	7	.752
Michigan W	748	9.98	7.71	.773	1	.702
Minnesota	1,410	15.61	14.53	.931	4	.741
Missouri E	1,486	20.42	15.31	.750	0	.750
Missouri W	1,474	15.40	15.19	.987	5	.745
Mississippi N	710	5.08	7.32	1.441	5	.726
Mississippi S	982	8.07	10.12	1.254	6	.719
Montana	673	4.59	6.93	1.510	5	.723
North Carolina E	1,077	9.07	11.10	1.224	6	.737
North Carolina M	1,041	5.81	10.73	1.847	9	.725
North Carolina W	1,206	6.72	12.43	1.850	10	.743
North Dakota	681	4.28	7.02	1.640	5	.756
Nebraska	785	6.99	8.09	1.158	4	.736
New Hampshire	472	3.96	4.86	1.228	3	.698
New Jersey	3,530	44.50	36.37	.817	4	.750
New Mexico	1,526	10.44	15.72	1.505	11	.733

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio before allocation[a]	Attorneys added	Staffing ratio after allocation[a]
Nevada	1,362	15.92	14.03	.881	3	.742
New York E	4,041	39.73	41.63	1.048	15	.761
New York N	1,346	11.79	13.87	1.176	7	.738
New York S	4,959	91.06	51.09	.561	0	.561
New York W	1,598	13.90	16.46	1.184	8	.752
Ohio N	2,596	24.46	26.75	1.094	11	.754
Ohio S	2,029	16.14	20.90	1.295	12	.743
Oklahoma E	478	2.87	4.92	1.715	4	.716
Oklahoma N	724	8.43	7.46	.884	2	.715
Oklahoma W	1,244	9.14	12.82	1.403	8	.748
Oregon	1,342	15.55	13.83	.890	3	.746
Pennsylvania E	2,811	46.36	28.96	.625	0	.625
Pennsylvania M	1,326	10.27	13.66	1.331	8	.748
Pennsylvania W	1,446	18.02	14.90	.827	2	.744
Puerto Rico	1,239	10.57	12.76	1.207	7	.726
Rhode Island	516	6.11	5.32	.871	1	.748
South Carolina	1,618	12.31	16.67	1.354	10	.747
South Dakota	750	5.54	7.73	1.395	5	.733
Tennessee E	1,127	11.41	11.61	1.017	4	.753
Tennessee M	965	7.48	9.94	1.329	6	.737
Tennessee W	1,892	11.78	19.49	1.655	14	.756
Texas E	1,278	11.68	13.17	1.128	6	.745
Texas N	2,530	25.76	26.07	1.012	9	.750
Texas S	4,043	41.79	41.65	.997	13	.760
Texas W	3,334	26.43	34.35	1.300	19	.756
Utah	1,057	9.80	10.89	1.111	5	.736
Virginia E	2,560	31.57	26.37	.835	4	.741
Virginia W	727	7.03	7.49	1.065	3	.747
Vermont	556	6.15	5.73	.932	2	.703
Washington E	578	6.30	5.96	.946	2	.718
Washington W	2,159	20.43	22.24	1.089	9	.756
Wisconsin E	1,363	12.64	14.04	1.111	6	.753
Wisconsin W	592	5.15	6.10	1.185	3	.748
West Virginia N	977	6.83	10.07	1.474	7	.728
West Virginia S	1,794	13.07	18.48	1.414	12	.737
Wyoming	555	4.71	5.72	1.215	3	.742
Total	**147,067**	**1,537.93**	**1,515.12**		**535**	

[a]Ratio of predicted and reported time expenditures.

Table IV.2: Allocation of 100 Civil Attorneys to Handle Fiscal Year 1989 Workload Using GAO's Model

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio before allocation[a]	Attorneys added	Staffing ratio after allocation[a]
Alaska	1,383	3.34	3.74	1.119	1	.862
Alabama M	1,265	3.68	3.42	.929	1	.731
Alabama N	1,997	10.52	5.40	.513	0	.513
Alabama S	1,657	3.17	4.48	1.413	2	.867
Arkansas E	2,160	4.31	5.84	1.355	3	.799
Arkansas W	1,316	3.06	3.56	1.163	1	.877
Arizona	2,766	14.48	7.48	.517	0	.517
California C	9,271	33.12	25.07	.757	0	.757
California E	3,628	14.60	9.81	.672	0	.672
California N	5,735	20.71	15.51	.749	0	.749
California S	6,105	14.39	16.51	1.148	4	.898
Colorado	2,574	7.37	6.96	.945	1	.832
Connecticut	2,015	5.96	5.45	.914	1	.783
Dist. of Columbia	7,581	22.00	20.50	.932	1	.891
Delaware	1,191	1.92	3.22	1.680	2	.821
Florida M	3,395	8.90	9.18	1.032	2	.842
Florida N	1,660	4.48	4.49	1.003	1	.819
Florida S	6,693	21.77	18.10	.832	0	.832
Georgia M	1,668	4.57	4.51	.986	1	.810
Georgia N	4,031	11.22	10.90	.972	1	.892
Georgia S	1,272	5.45	3.44	.631	0	.631
Hawaii	1,867	3.99	5.05	1.265	2	.843
Iowa N	1,538	3.84	4.16	1.084	1	.860
Iowa S	1,557	4.94	4.21	.852	0	.852
Idaho	1,524	5.40	4.12	.763	0	.763
Illinois C	1,331	3.61	3.60	.998	1	.781
Illinois N	4,696	14.91	12.70	.852	0	.852
Illinois S	1,694	3.92	4.58	1.169	2	.774
Indiana N	1,945	4.11	5.26	1.280	2	.861
Indiana S	2,725	5.11	7.37	1.442	4	.809
Kansas	2,208	10.35	5.97	.576	0	.576
Kentucky E	2,100	8.52	5.68	.667	0	.667
Kentucky W	1,856	5.64	5.02	.891	0	.891
Louisiana E	2,034	10.24	5.50	.537	0	.537
Louisiana M	1,083	2.38	2.93	1.233	1	.867
Louisiana W	2,977	5.13	8.05	1.569	4	.882
Massachusetts	4,796	10.88	12.97	1.192	4	.872

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio before allocation[a]	Attorneys added	Staffing ratio after allocation[a]
Maine	1,357	3.90	3.67	.940	1	.749
Maryland	2,193	6.39	5.93	.928	1	.802
Michigan E	3,731	13.75	10.09	.734	0	.734
Michigan W	2,219	5.71	6.00	1.050	1	.894
Minnesota	2,026	6.06	5.48	.904	0	.904
Missouri E	2,270	4.21	6.14	1.458	3	.852
Missouri W	1,912	8.18	5.17	.632	0	.632
Mississippi N	1,217	4.55	3.29	.723	0	.723
Mississippi S	1,424	5.78	3.85	.666	0	.666
Montana	1,394	3.94	3.77	.957	1	.763
North Carolina E	1,475	5.27	3.99	.757	0	.757
North Carolina M	1,335	2.30	3.61	1.570	2	.840
North Carolina W	1,257	1.89	3.40	1.799	2	.874
North Dakota	1,302	2.99	3.52	1.176	1	.882
Nebraska	1,535	5.26	4.15	.789	0	.789
New Hampshire	987	1.94	2.67	1.375	1	.908
New Jersey	4,918	12.93	13.30	1.029	2	.891
New Mexico	1,915	8.12	5.18	.638	0	.638
Nevada	1,756	4.68	4.75	1.014	1	.836
New York E	7,859	24.92	19.90	.799	0	.799
New York N	2,038	3.80	5.51	1.450	3	.810
New York S	7,008	29.52	18.95	.642	0	.642
New York W	2,400	5.72	6.49	1.135	2	.841
Ohio N	3,750	13.22	10.14	.767	0	.767
Ohio S	3,339	10.53	9.03	.857	0	.857
Oklahoma E	1,169	2.06	3.16	1.534	2	.778
Oklahoma N	1,424	2.92	3.85	1.317	2	.783
Oklahoma W	2,222	8.36	6.01	.718	0	.718
Oregon	1,756	7.25	4.75	.655	0	.655
Pennsylvania E	4,574	13.19	12.37	.938	1	.872
Pennsylvania M	1,679	3.53	4.54	1.287	2	.821
Pennsylvania W	2,762	5.45	7.47	1.371	3	.884
Puerto Rico	2,836	8.19	7.67	.936	1	.835
Rhode Island	1,246	2.68	3.37	1.258	2	.720
South Carolina	2,762	11.14	7.47	.671	0	.671
South Dakota	1,298	4.75	3.51	.738	0	.738
Tennessee E	1,908	4.39	5.16	1.176	2	.808
Tennessee M	1,457	3.78	3.94	1.041	1	.824
Tennessee W	2,322	5.33	6.28	1.177	2	.857

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio before allocation[a]	Attorneys added	Staffing ratio after allocation[a]
Texas E	1,731	7.03	4.68	.666	0	.666
Texas N	4,508	13.49	12.19	.904	0	.904
Texas S	5,447	17.85	14.73	.825	0	.825
Texas W	2,163	11.41	5.85	.513	0	.513
Utah	1,538	5.50	4.16	.757	0	.757
Virginia E	2,463	11.83	6.66	.563	0	.563
Virginia W	1,597	2.45	4.32	1.763	3	.793
Vermont	1,228	2.12	3.32	1.567	2	.806
Washington E	1,379	3.55	3.73	1.051	1	.820
Washington W	3,284	7.92	8.88	1.122	2	.895
Wisconsin E	1,915	4.29	5.18	1.208	2	.824
Wisconsin W	1,712	3.65	4.63	1.268	2	.819
West Virginia N	1,276	1.63	3.45	2.118	3	.745
West Virginia S	1,812	3.73	4.90	1.315	2	.855
Wyoming	1,187	2.63	3.21	1.221	1	.844
Total	**230,069**	**699.64**	**622.15**		**100**	

[a]Ratio of predicted and reported time expenditures.

The third example, presented in table IV.3, compares Justice's 1989 allocation of 423 new attorney positions to prosecute drug cases with the model's allocation. GAO's allocation is based strictly on the reported and predicted times for the drug crime litigation area. The Justice and GAO allocations are highly correlated (Pearson correlation = .57). Yet selected offices fare very differently under the two allocations.

Table IV.3: Comparison of GAO's and Justice's Allocation of 423 Criminal Attorneys Targeted in Fiscal Year 1989 to Prosecute Drug Crimes

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	GAO allocation	Justice allocation
Alaska	89	1.11	1.51	1.361	2	3
Alabama M	131	1.89	2.22	1.172	2	3
Alabama N	214	4.98	3.63	0.730	2	5
Alabama S	277	2.76	4.70	1.703	6	2
Arkansas E	149	2.29	2.53	1.103	2	2
Arkansas W	92	0.73	1.56	2.144	2	2
Arizona	749	12.70	12.70	1.000	9	4
California C	761	20.70	12.89	0.623	1	21
California E	407	7.02	6.90	0.982	5	5
California N	552	15.47	9.35	0.605	1	6

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	GAO allocation	Justice allocation
California S	948	16.56	16.06	0.970	11	6
Colorado	274	7.06	4.64	0.657	1	2
Connecticut	206	5.67	3.49	0.616	1	6
Dist. of Columbia	1,136	16.74	19.25	1.150	16	28
Delaware	114	1.54	1.94	1.261	2	2
Florida M	1,037	16.76	17.58	1.049	13	12
Florida N	412	6.55	6.98	1.066	6	5
Florida S	2,122	39.18	35.97	0.918	21	42
Georgia M	247	3.25	4.18	1.285	4	3
Georgia N	555	9.93	9.40	0.947	6	5
Georgia S	167	3.98	2.83	0.712	1	2
Hawaii	358	3.64	6.07	1.667	7	3
Iowa N	192	2.26	3.25	1.437	4	3
Iowa S	159	2.03	2.69	1.327	3	2
Idaho	86	1.07	1.45	1.351	2	2
Illinois C	336	4.96	5.69	1.147	5	3
Illinois N	817	13.69	13.84	1.011	10	9
Illinois S	228	3.55	3.87	1.091	3	4
Indiana N	222	2.94	3.77	1.283	4	3
Indiana S	174	3.38	2.95	0.874	2	3
Kansas	217	2.21	3.67	1.659	4	4
Kentucky E	231	4.42	3.92	0.887	3	2
Kentucky W	232	3.68	3.94	1.072	3	4
Louisiana E	413	5.29	7.00	1.324	7	3
Louisiana M	64	0.12	1.09	9.067	2	3
Louisiana W	134	2.30	2.27	0.985	2	2
Massachusetts	251	8.10	4.26	0.526	0	4
Maine	248	4.95	4.20	0.849	3	2
Maryland	450	9.81	7.63	0.778	3	6
Michigan E	767	15.28	13.00	0.851	7	8
Michigan W	165	4.59	2.79	0.607	1	3
Minnesota	340	5.65	5.77	1.022	4	3
Missouri E	341	5.94	5.78	0.974	4	4
Missouri W	309	6.50	5.24	0.806	3	5
Mississippi N	139	2.07	2.35	1.134	2	3
Mississippi S	155	2.52	2.62	1.040	2	3
Montana	133	1.52	2.26	1.486	3	1
North Carolina E	287	3.62	4.86	1.342	5	1
North Carolina M	261	2.45	4.42	1.806	5	2
North Carolina W	209	2.41	3.54	1.470	4	2

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	GAO allocation	Justice allocation
North Dakota	102	1.27	1.73	1.362	2	2
Nebraska	158	2.00	2.67	1.337	3	1
New Hampshire	105	1.83	1.78	0.970	2	1
New Jersey	419	10.76	7.10	0.660	2	7
New Mexico	455	4.74	7.71	1.627	9	3
Nevada	171	5.82	2.90	0.498	0	4
New York E	1,097	8.42	18.59	2.208	23	16
New York N	266	5.06	4.51	0.891	3	5
New York S	1,121	28.46	18.99	0.667	4	9
New York W	301	4.64	5.10	1.100	4	3
Ohio N	451	8.78	7.65	0.871	4	4
Ohio S	480	6.85	8.14	1.188	7	3
Oklahoma E	103	1.04	1.75	1.682	2	2
Oklahoma N	197	2.78	3.34	1.202	3	2
Oklahoma W	250	3.56	4.23	1.189	4	3
Oregon	358	6.86	6.06	0.884	4	5
Pennsylvania E	653	14.06	11.06	0.786	5	6
Pennsylvania M	272	3.40	4.61	1.356	5	4
Pennsylvania W	241	5.51	4.09	0.741	2	3
Puerto Rico	318	4.38	5.39	1.231	5	3
Rhode Island	136	2.35	2.31	0.983	2	1
South Carolina	238	4.80	4.04	0.842	2	1
South Dakota	117	1.25	1.98	1.586	3	3
Tennessee E	278	4.16	4.71	1.132	4	2
Tennessee M	156	1.92	2.65	1.381	3	2
Tennessee W	555	3.21	9.40	2.928	13	4
Texas E	192	4.58	3.26	0.712	1	7
Texas N	438	7.66	7.43	0.970	5	9
Texas S	1,165	18.93	19.75	1.043	15	8
Texas W	1,100	13.18	18.64	1.414	18	6
Utah	174	2.76	2.95	1.068	3	2
Virginia E	431	9.84	7.30	0.742	3	8
Virginia W	165	2.59	2.80	1.083	3	2
Vermont	118	2.35	2.00	0.850	1	2
Washington E	145	2.76	2.46	0.892	2	2
Washington W	564	7.51	9.56	1.273	9	3
Wisconsin E	330	4.84	5.59	1.154	5	4
Wisconsin W	139	2.49	2.35	0.945	2	2

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	GAO allocation	Justice allocation
West Virginia N	278	3.82	4.72	1.235	5	1
West Virginia S	461	5.62	7.82	1.391	8	3
Wyoming	105	1.46	1.78	1.217	2	2
Total	33,364	566.12	565.40		423	423

[a]Ratio of predicted and reported time expenditures.

Figure IV.1 compares the drug crime staffing ratios before the allocation, the drug crime staffing ratios after the Justice allocation, and the drug crime staffing ratios after the GAO model's allocation. Both allocations substantially reduce disparities among offices, but the GAO allocation is more effective in reducing disparities than the Justice allocation.

Figure IV.1: Staffing Ratios of Attorneys for Drug Crimes, Before and After Allocation of New Attorney Positions

U.S. Attorney offices

——— Ratio before allocation
■ ■ ■ ■ Ratio after Justice's allocation
▬▬▬ Ratio after GAO's allocation
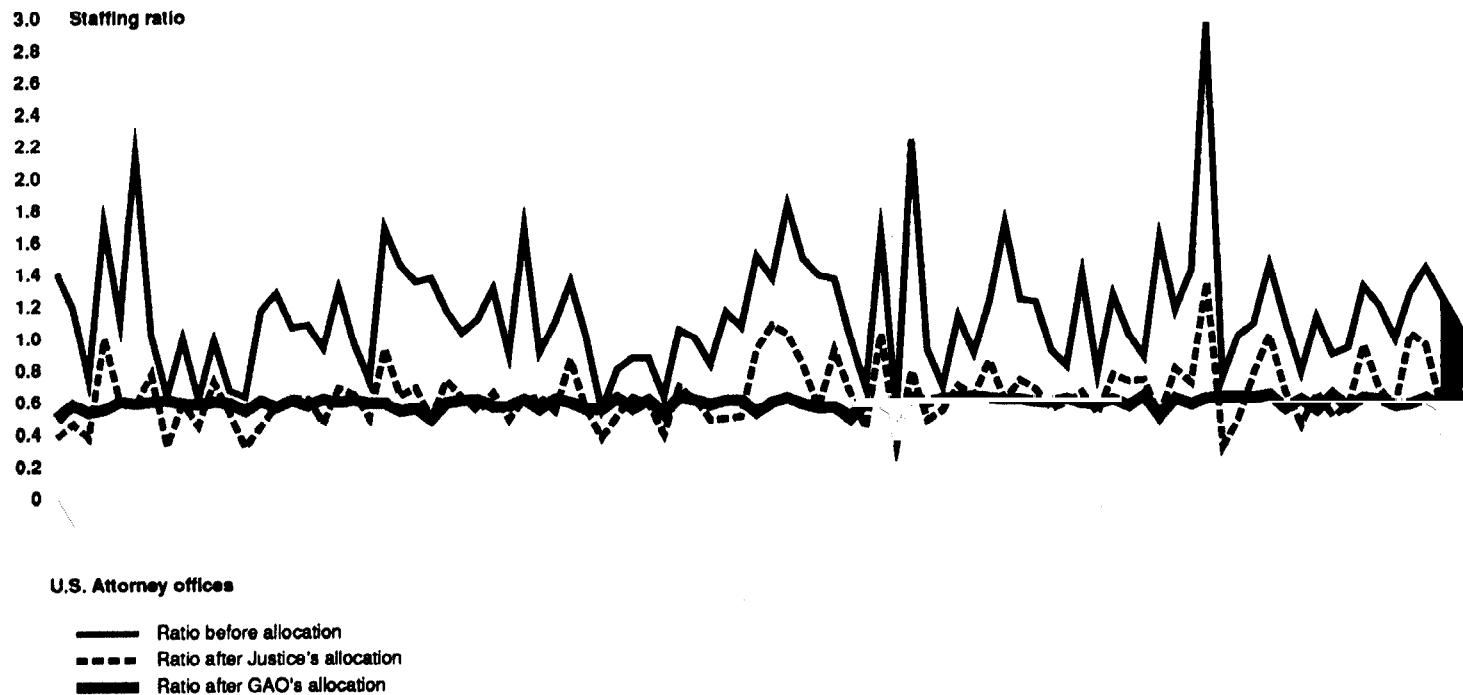
Note: 423 new attorney positions allocated in fiscal year 1989.

The final two examples pertain to fiscal year 1990. During fiscal year 1990, Congress appropriated funds for 614 new AUSAs to prosecute drug

crime and violent crime. Of these 614 positions, Justice has already allocated 557 to specified USAOs. Justice is currently considering whether and how the remaining 57 positions should be allocated. Of the 557 already allocated, 340 positions have been allocated to drug crime: 122 to USAOs in areas designated as "high intensity" by the administration and 218 to other USAOs that Justice believes to be especially in need of drug crime legal resources. The balance of the positions already allocated by Justice, a total of 217 (557-340), have been targeted to violent crime litigation. Further details on these allocations are not currently available.

Table IV.4 shows GAO's allocation of the 340 drug crime AUSA positions. This fiscal 1990 allocation incorporates two adjustments of the fiscal year 1989 predicted and reported times presented in table IV.3. First, before starting the allocation, we added a total of 423 previously allocated drug positions to the reported times of offices receiving new AUSAs in fiscal 1989. Second, to attempt to account for changes in the predicted time (workload) that will result from the 423 newly added positions, we multiplied the increase in reported time (i.e., the number of newly added positions) of each office by a factor of .899 and then added this product to the predicted time of the office before starting the allocation.

The factor .899 was estimated from a trend analysis of drug crime predicted and reported times during 1987 through 1989 (see the next section). In brief, between 1987 and 1989, there were 99 drug crime attorneys added and predicted time increased by 89. The adjustment factor of .899 was estimated by dividing the increase in predicted time (89) during 1987 through 1989 by the increase in reported time (99) during the same period. Technically, the adjustment factor of .899 represents the estimated marginal product of an additional drug crime attorney.

Table IV.4: Allocation of 340 Criminal Attorneys Targeted in Fiscal Year 1990 to Prosecute Drug Crimes (Using GAO's Resource Model)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Attorneys added
Alaska	148	4.11	4.21	1.024	2
Alabama M '	173	4.89	4.91	1.005	2
Alabama N	287	9.98	8.13	.814	2
Alabama S	229	4.76	6.50	1.365	5
Arkansas E	152	4.29	4.32	1.008	2

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Attorneys added
Arkansas W	118	2.73	3.36	1.232	2
Arizona	575	16.70	16.30	.976	6
California C	1,121	41.70	31.77	.762	2
California E	402	12.02	11.39	.948	4
California N	520	21.47	14.75	.687	0
California S	757	22.56	21.46	.951	7
Colorado	227	9.06	6.44	.710	0
Connecticut	314	11.67	8.89	.762	1
Dist. of Columbia	1,567	44.74	44.42	.993	16
Delaware	132	3.54	3.74	1.057	2
Florida M	1,001	28.76	28.37	.986	10
Florida N	405	11.55	11.47	.993	4
Florida S	2,601	81.18	73.73	.908	19
Georgia M	242	6.25	6.87	1.100	4
Georgia N	490	14.93	13.90	.931	4
Georgia S	163	5.98	4.63	.774	1
Hawaii	309	6.64	8.76	1.320	6
Iowa N	210	5.26	5.94	1.130	3
Iowa S	158	4.03	4.49	1.115	3
Idaho	114	3.07	3.24	1.057	2
Illinois C	296	7.96	8.39	1.054	4
Illinois N	774	22.69	21.93	.967	8
Illinois S	263	7.55	7.47	.989	3
Indiana N	228	5.94	6.47	1.089	3
Indiana S	199	6.38	5.65	.886	2
Kansas	256	6.21	7.26	1.169	4
Kentucky E	202	6.42	5.72	.891	2
Kentucky W	266	7.68	7.54	.982	3
Louisiana E	342	8.29	9.70	1.170	5
Louisiana M	134	3.12	3.79	1.213	3
Louisiana W	143	4.30	4.06	.945	2
Massachusetts	277	12.10	7.85	.649	0
Maine	212	6.95	6.00	.864	2
Maryland	460	15.81	13.03	.824	2
Michigan E	712	23.28	20.19	.867	5
Michigan W	193	7.59	5.48	.722	0
Minnesota	299	8.65	8.47	.979	3
Missouri E	331	9.94	9.38	.944	3
Missouri W	343	11.50	9.73	.846	2
Mississippi N	178	5.07	5.04	.995	2

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Attorneys added
Mississippi S	188	5.52	5.32	.963	2
Montana	111	2.52	3.16	1.253	2
North Carolina E	203	4.62	5.76	1.247	4
North Carolina M	219	4.45	6.22	1.398	4
North Carolina W	188	4.41	5.34	1.211	3
North Dakota	124	3.27	3.53	1.079	2
Nebraska	126	3.00	3.57	1.191	2
New Hampshire	94	2.83	2.67	.945	1
New Jersey	473	17.76	13.40	.754	1
New Mexico	367	7.74	10.41	1.345	7
Nevada	229	9.82	6.49	.661	0
New York E	1,163	24.42	32.98	1.350	21
New York N	318	10.06	9.01	.895	3
New York S	955	37.46	27.08	.723	0
New York W	275	7.64	7.80	1.021	3
Ohio N	396	12.78	11.24	.880	3
Ohio S	382	9.85	10.83	1.100	5
Oklahoma E	125	3.04	3.55	1.167	2
Oklahoma N	181	4.78	5.14	1.075	3
Oklahoma W	244	6.56	6.93	1.056	3
Oregon	373	11.86	10.56	.890	3
Pennsylvania E	580	20.06	16.45	.820	3
Pennsylvania M	290	7.40	8.21	1.109	4
Pennsylvania W	239	8.51	6.78	.797	1
Puerto Rico	285	7.38	8.09	1.096	4
Rhode Island	113	3.35	3.21	.958	2
South Carolina	174	5.80	4.94	.852	1
South Dakota	165	4.25	4.68	1.101	3
Tennessee E	230	6.16	6.51	1.057	3
Tennessee M	157	3.92	4.45	1.135	3
Tennessee W	458	7.21	12.99	1.802	11
Texas E	337	11.58	9.55	.825	2
Texas N	547	16.66	15.52	.932	5
Texas S	950	26.93	26.94	1.001	10
Texas W	848	19.18	24.03	1.253	14
Utah	168	4.76	4.75	.997	2
Virginia E	512	17.84	14.50	.813	2
Virginia W	162	4.59	4.60	1.003	2
Vermont	134	4.35	3.80	.873	1
Washington E	150	4.76	4.26	.895	2

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Attorneys added
Washington W	432	10.51	12.25	1.166	7
Wisconsin E	324	8.84	9.18	1.039	4
Wisconsin W	146	4.49	4.15	.924	2
West Virginia N	198	4.82	5.62	1.165	3
West Virginia S	371	8.62	10.52	1.220	6
Wyoming	126	3.46	3.57	1.033	2
Total	**33,363**	**989.11**	**945.67**		**340**

[a]Ratio of predicted and reported time expenditures.

Finally, table IV.5 shows how GAO's model would have allocated the 217 violent crime attorney positions. This allocation is based on the estimated 1989 staffing ratios in violent crime. The violent crime allocation in table IV.5 , as well as the drug crime allocations in tables IV.3 and IV.4, illustrate how the GAO model can be used to make specialized allocations for litigation areas that are targeted by Congress.

Table IV.5: Allocation of 217 Criminal Attorneys Targeted in Fiscal Year 1990 to Prosecute Violent Crimes (Using GAO's Resource Model)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Attorneys added
Alaska	31	.30	.49	1.623	2
Alabama M	32	.10	.50	4.977	2
Alabama N	43	.50	.67	1.347	2
Alabama S	33	.11	.52	4.705	2
Arkansas E	34	.68	.53	.779	1
Arkansas W	30	.07	.46	6.600	2
Arizona	307	7.24	4.78	.660	6
California C	507	8.92	7.89	.885	13
California E	107	1.56	1.67	1.068	3
California N	128	1.95	2.00	1.023	4
California S	143	1.40	2.22	1.584	5
Colorado	74	.47	1.16	2.474	3
Connecticut	69	4.77	1.07	.225	0
Dist. of Columbia	60	1.42	.93	.656	2
Delaware	28	.17	.44	2.603	2
Florida M	139	1.79	2.17	1.211	5
Florida N	46	.17	.72	4.246	2
Florida S	175	3.50	2.73	.779	4
Georgia M	33	.19	.51	2.709	2

(continued)

U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Attorneys added
Georgia N	109	2.30	1.69	.737	3
Georgia S	44	.58	.68	1.174	2
Hawaii	59	.57	.92	1.619	2
Iowa N	28	.30	.43	1.423	1
Iowa S	24	.07	.38	5.463	1
Idaho	39	.34	.61	1.796	2
Illinois C	25	.26	.39	1.518	1
Illinois N	80	2.98	1.25	.421	1
Illinois S	34	.45	.53	1.171	1
Indiana N	30	.35	.46	1.312	1
Indiana S	39	1.15	.60	.522	1
Kansas	125	.22	1.95	8.865	6
Kentucky E	39	.94	.60	.641	1
Kentucky W	53	.83	.82	.984	2
Louisiana E	53	.47	.82	1.753	2
Louisiana M	19	.00	.29	.000	0
Louisiana W	34	.05	.53	10.616	2
Massachusetts	89	3.50	1.38	.395	1
Maine	28	.07	.43	6.207	2
Maryland	116	2.12	1.80	.847	3
Michigan E	72	.53	1.12	2.112	3
Michigan W	49	.56	.76	1.361	2
Minnesota	51	.92	.79	.860	2
Missouri E	54	.23	.84	3.634	3
Missouri W	68	.69	1.06	1.537	3
Mississippi N	31	.49	.49	1.003	1
Mississippi S	44	.47	.69	1.463	2
Montana	75	1.27	1.17	.924	2
North Carolina E	80	.45	1.25	2.772	3
North Carolina M	92	.58	1.44	2.489	4
North Carolina W	139	1.15	2.17	1.885	5
North Dakota	94	1.15	1.46	1.266	3
Nebraska	47	.51	.73	1.429	2
New Hampshire	25	.19	.39	2.028	1
New Jersey	69	1.37	1.08	.785	2
New Mexico	128	1.44	2.00	1.392	5
Nevada	92	1.01	1.44	1.425	3
New York E	86	3.88	1.34	.345	0
New York N	34	.60	.53	.880	1
New York S	76	2.76	1.18	.428	1

(continued)

| U.S. Attorney office	Weighted cases	Reported time	Predicted time	Staffing ratio[a]	Attorneys added
New York W	59	.56	.92	1.635	2
Ohio N	121	1.98	1.88	.948	4
Ohio S	57	.65	.89	1.370	2
Oklahoma E	21	.13	.33	2.537	1
Oklahoma N	40	1.00	.62	.618	1
Oklahoma W	56	.78	.88	1.130	2
Oregon	172	2.99	2.68	.895	5
Pennsylvania E	76	1.70	1.19	.698	2
Pennsylvania M	55	.22	.85	3.860	3
Pennsylvania W	62	.67	.96	1.437	2
Puerto Rico	52	1.05	.81	.775	2
Rhode Island	21	.12	.33	2.740	1
South Carolina	83	.21	1.29	6.126	4
South Dakota	157	2.28	2.45	1.077	5
Tennessee E	54	.60	.84	1.404	2
Tennessee M	44	.38	.69	1.806	2
Tennessee W	54	.40	.84	2.099	2
Texas E	42	.66	.65	.977	2
Texas N	78	1.42	1.21	.853	2
Texas S	85	1.21	1.32	1.090	3
Texas W	58	1.43	.91	.637	2
Utah	66	1.74	1.03	.591	2
Virginia E	99	1.73	1.54	.888	3
Virginia W	27	.27	.42	1.571	1
Vermont	25	.26	.39	1.513	1
Washington E	75	.05	1.17	23.379	4
Washington W	69	1.38	1.08	.781	2
Wisconsin E	47	.70	.74	1.052	2
Wisconsin W	30	.05	.47	9.333	2
West Virginia N	21	.13	.33	2.527	1
West Virginia S	27	.49	.42	.860	1
Wyoming	44	.34	.68	2.005	2
Total	**6,474**	**100.69**	**100.76**		**217**

[a]Ratio of predicted and reported time expenditures.

National Staffing Trends, Fiscal Years 1987-1989

To estimate trends in the marginal productivity of attorney time, we analyzed USAO national staffing trends in the ratio of predicted and reported times for three broad litigation areas: drug crime, other criminal litigation, and civil litigation for 1987 through 1989. Figure IV.2 shows that the staffing increases in civil litigation and in other criminal

litigation during this period, as gauged using reported time, did not yield equal increases in legal work, as gauged using predicted time. Specifically, during 1987 through 1989, the reported times for other criminal and civil litigation increased by 75 FTES and 97 FTES, respectively, and the predicted times increased by 25 and 17 FTES, respectively. In short, figure IV.2 suggests declining marginal productivity for both civil and other criminal litigation during 1987.

Figure IV.2: Trends in Staffing for 91 USAOs
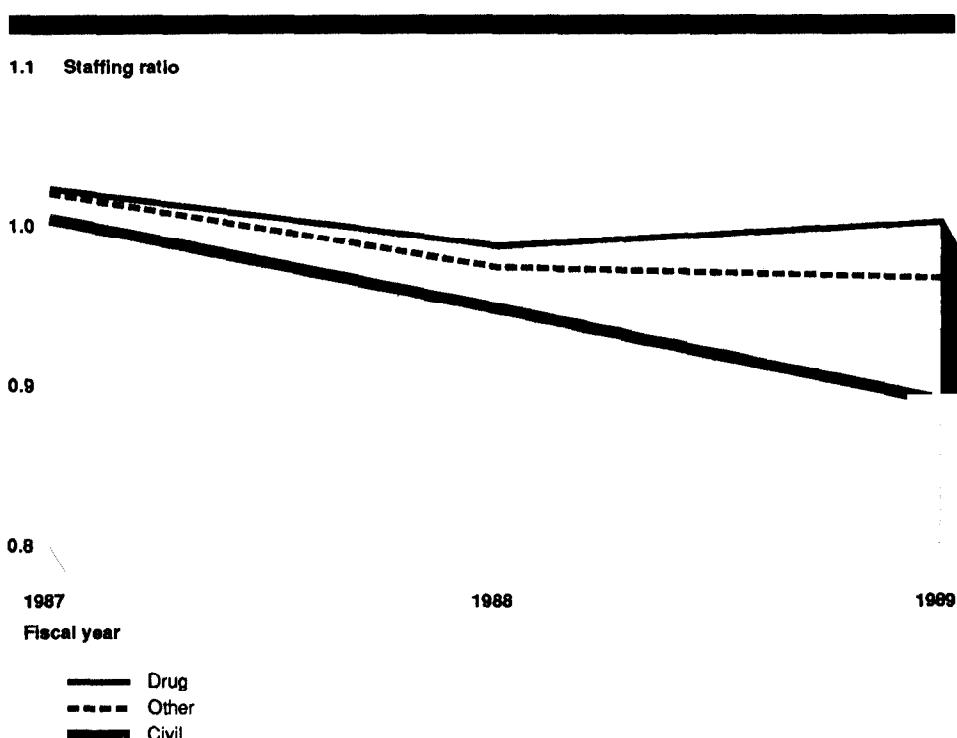
1.1   Staffing ratio

Note: Change in staffing ratio for drug is not statistically significant.

We do not know whether the declining slopes of the other criminal and civil curves in figure IV.2 are affected more by scarcities of other legal inputs or by scarcities of criminal and civil actions. But the flat curve of estimated drug crime marginal productivity suggests there is no shortage of drug crimes to be prosecuted. Hence, the drug trend line appears to be consistent with many U.S. Attorneys' statements that they lacked attorneys to prosecute drug cases, and as a result they had to decline many prosecutable cases.

In summary, the analyses presented in this section suggest that diminishing marginal productivity better describes civil and other criminal litigation than drug crime litigation. Yet these analyses are not definitive since they do not take into account possible lags between the allocation of new resources and the full realization of increased production from these resources. It may take longer, on average, for new resources in civil litigation or in other criminal litigation to affect production than for new resources in drug crime litigation to affect production.

Technical Documentation on Case and Defendant Weighting Models

This appendix discusses technical properties of the criminal and civil case and defendant weighting models described in appendix I. Appendix V has six sections:

- Section 1 gives an overview of the approach to workload weighting.
- Section 2 discusses alternative functional specifications of the weighting model that correspond to different theories of the time requirements of legal work.
- Section 3 compares the overall goodness-of-fit of alternative models and shows that the selected models (app. I) perform well relative to other weighting models that were considered.
- Section 4 considers the analysis of residuals for the selected weighting models and the treatment of "outliers," i.e., eccentric observations.
- Section 5 describes the statistical estimation technique that was used to estimate the parameters (i.e., the case and defendant weights) of the selected weighting models.
- Section 6 discusses the statistical precision of the estimated staffing ratios.

1. Overview. We used multiple regression to predict FTE expenditures for the USAOs in six civil and seven criminal litigation areas and to estimate predicted times and staffing ratios for the USAOs in each litigation area. In each regression, the dependent variable was the reported FTE expenditure of the USAO in the litigation area during the fiscal year. The independent variables were indicators of the workload of the USAO in the litigation area during the fiscal year, namely, the numbers of cases and defendants of different kinds that were handled by the USAO during the year. Each regression was based on 3 years of data (1987-89) for 91 USAOs, the observations for all USAOs except Northern Mariana islands, Guam, and the Virgin Islands.

A number of alternative regression model specifications, using different predictors and different functional forms, were considered for each criminal and civil litigation area. The selected criminal and civil weighting models (see app. I) were the models that were theoretically reasonable and performed best overall, on the basis of measures of goodness-of-fit, in the criminal and civil litigation areas, respectively.

For each USAO, predicted FTEs from the regressions were summed across criminal litigation areas to yield the total predicted criminal FTE expenditure of the USAO. For each USAO, reported FTEs were summed across criminal areas to yield the total reported criminal FTE expenditure of the

USAO. Total predicted and reported civil FTES were computed
analogously.

In each of the 13 litigation areas and in each of the total criminal and
total civil areas, the staffing ratio of a USAO is calculated by dividing the
total predicted FTE expenditure of the USAO in the area by the total
reported FTE expenditure of the USAO in the area. If the staffing ratio of
a USAO is significantly greater than unity, we infer that the USAO is
understaffed or efficient, relative to other USAOs. If the staffing ratio of
a USAO is significantly less than unity, we infer that the USAO is over-
staffed or inefficient, relative to other USAOs.

2. Alternative model specifications. As noted in appendix I, a case with
multiple defendants often requires a greater investment of attorney time
than a similar case with a single defendant. Moreover, the time required
to prosecute or defend any fixed number of defendants often is greater
when each defendant must be litigated as a separate case than when all
defendants can be litigated in a single case. Therefore, the alternative
weighting models we considered used both caseloads and defendant
loads as predictors of FTES.

We considered two classes of models, each corresponding to different
theories of the time requirements of legal work. These two classes of
models are called "quadratic" and "interaction."

In the quadratic class of weighting models, the predictors of FTE expend-
itures in a litigation area are (1) the numbers of cases of different types
(i.e., levels of legal activity, as discussed in app. I) in the litigation area;
(2) the number of defendants of different types in the litigation area;
and (3) the squares of the numbers of cases and defendants of different
types. The quadratic class includes models that are linear in the case
and defendant loads as special cases. The quadratic terms are inter-
preted as the effects of economies of scale in the processing of cases and
defendants of the same type.

In the interaction class of weighting models, the predictors are the num-
bers of cases and defendants of different types with specified numbers
of defendants (e.g., 1, 2, 3, and 4 or more). Unlike the quadratic model,
which assumes that the effects of numbers of cases and numbers of
defendants on FTE expenditures are additive, the interaction model
allows for interactions in the effects of cases and defendants. That is,
FTE expenditures are assumed to depend upon the detailed distribution

of defendants in cases of each type, not just upon the total numbers of cases and defendants of each type.

Many different weighting models, distinguished by different numbers and sets of predictors, can be specified in each of the quadratic and linear classes. In theory, the elaboration of alternative models is limited only by the number of data points that are available to estimate the parameters, i.e., $3 \times 91 = 271$ data points. To mitigate the risk of overfitting the data, we considered only models that included 30 or fewer predictors.[1] As discussed in appendix I, each of the selected civil models has 18 predictors, and each of the selected criminal models has 16 predictors.

The selected models of appendix I combine the features of the two classes of models. Like interaction models, these models distinguish between cases of different types with one defendant and cases of the same types with two or more defendants. Like linear models (a subclass of the class of quadratic models), the selected models use the total numbers of defendants of different types as predictors. Because of the unique characteristics of different cases, legal work is theorized to entail few economies of scale; hence, the selected models include no quadratic predictors. Yet, like quadratic models and unlike interaction models, the selected models account for every defendant that is reported in the litigation area.

3. Goodness-of-fit. Selection of the final models depended on empirical as well as on theoretical criteria. Tables V.1 and V.2 present two global measures of goodness-of-fit, the adjusted R^2 and the residual mean square. In table V.1, the two measures are presented for seven weighting models applied to each of the six criminal litigation areas. In table V.2, the two measures are presented for the same seven weighting models applied to each of the seven civil litigation areas. In each table, results are shown separately for the fiscal year 1987, 1988, and 1989 universes.

The coefficient of determination, R^2, equals the proportion of the total variance in FTE expenditures attributable to the caseload and/or defendant load predictors. Since the unadjusted R^2 is nondecreasing in the number of predictors, the R^2s in tables V.1 and V.2 are adjusted for the number of predictors in order to make comparisons among models with

---

[1]See, e.g., G. C. Judge, et al., The Theory and Practice of Econometrics (New York: Wiley, 1980), ch. 11.

different numbers of predictors meaningful.[2] Note that the adjusted R^2s reported in tables V.1 and V.2 are smaller than the corresponding unadjusted R^2s reported in appendix I.

As Weisberg stated, "Good models will have large values of the adjusted R^2." Moreover, good models will have small values of the residual mean square. Other popular measures of goodness-of-fit, such as Mallow's C_p, and the total and regression sums of squares can be computed from tables V.1 and V.2.

The selected criminal model is labeled "FIN" in table V.1, and the selected civil model is labeled "FIN" in table V.2. The predictors of FTES in the remaining five models are as follows:

- model C - total caseload (one predictor).
- model D - total defendant load (one predictor).
- model CL (caseload linear) - numbers of cases for the six types of legal activity.
- model DL (defendant load linear) - numbers of defendants for the six types of legal activity.
- model CL-DL - numbers of cases and numbers of defendants for the six types of legal activity.

All models include an intercept term.

---

[2]See, e.g., S. Weisberg, Applied Linear Regression (New York: Wiley, 1980), p. 188.

Table V.1: Adjusted R²s and Residual Mean Squares of Weighting Models Applied to Six Criminal Litigation Areas, Fiscal Years 1987, 1988, and 1989 (Total Observations for Each Year = 91)

Model		C	D	CL	DL	CL-DL	FIN
Number of predictors		1	1	6	6	12	16
Area statistic				**Fiscal year 1987**			
PC	adj. R^2	.46	.40	.68	.61	.68	.70
	resid. MS	1.82	2.00	1.07	1.29	1.08	.99
EC	adj. R^2	.74	.77	.78	.75	.78	.79
	resid. MS	5.22	4.64	5.10	4.54	4.47	4.27
OC	adj. R^2	.30	.28	.42	.39	.65	.64[a]
	resid. MS	1.67	1.72	1.39	1.45	.83	.87
DR	adj. R^2	.80	.83	.87	.87	.89	.91
	resid. MS	6.36	5.34	4.13	4.15	3.59	2.79
VC	adj. R^2	.33	.35	.60	.62	.62	.62
	resid. MS	.81	.78	.49	.46	.46	.46
OT	adj. R^2	.43	.47	.51	.53	.60	.62
	resid. MS	4.12	3.85	3.55	3.39	2.88	2.73
				Fiscal year 1988			
PC	adj. R^2	.44	.40	.60	.62	.75	.78
	resid. MS	2.00	2.15	1.43	1.35	.88	.80
EC	adj. R^2	.76	.79	.79	.80	.82	.83
	resid. MS	5.11	4.37	4.34	4.25	3.72	3.56
OC	adj. R^2	.37	.34	.56	.49	.75	.55[a]
	resid. MS	1.16	1.22	.81	.93	.46	.83
DR	adj. R^2	.77	.87	.88	.91	.93	.94
	resid. MS	8.48	4.74	4.58	3.33	2.75	2.33
VC	adj. R^2	.31	.35	.45	.49	.53	.64
	resid. MS	.91	.86	.72	.68	.62	.48
OT	adj. R^2	.55	.51	.57	.61	.64	.65
	resid. MS	3.59	3.96	3.48	3.10	2.87	2.83
				Fiscal year 1989			
PC	adj. R^2	.55	.46	.67	.60	.69	.80
	resid. MS	1.31	1.56	.96	1.16	.89	.58
EC	adj. R^2	.64	.67	.74	.72	.77	.82
	resid. MS	9.74	9.02	7.19	7.73	6.25	4.94
OC	adj. R^2	.28	.28	.31	.37	.41	.45[a]
	resid. MS	1.72	1.71	1.63	1.50	1.40	1.31
DR	adj. R^2	.57	.63	.77	.84	.87	.89
	resid. MS	16.37	14.21	8.76	6.07	5.03	4.35
VC	adj. R^2	.50	.50	.67	.68	.75	.74
	resid. MS	1.00	1.00	.66	.65	.51	.52
OT	adj. R^2	.50	.38	.59	.54	.64	.62
	resid. MS	3.69	4.64	3.03	3.43	2.70	2.80

[a]To prevent overfitting, the final model for OC was estimated using only 14 rather than 16 predictors: the number of multiple defendant new felonies and defendants in such cases were too small to support estimation. Hence, only total new felony cases (in place of single defendant new felony cases, multiple-defendant new felony cases, and defendants in new felony cases) was used in the predicting OC FTEs. This resulted in a loss of two predictors and an apparent loss of predictive power. Actually, the model was probably improved by virtue of the elimination of predictors that were capitalizing only on chance variation.

Table V.2: Adjusted R²s and Residual Mean Squares of Weighting Models Applied to Seven Civil Litigation Areas, Fiscal Years 1987, 1988, and 1989 (Total Observations for Each Year = 91)

Model		C	D	CL	DL	CL-DL	FIN
Number of predictors		1	1	6	6	12	18
Area statistic		**Fiscal year 1987**					
AM	adj. R^2	.40	.42	.49	.50	.58	.62
	resid. MS	.65	.62	.55	.54	.45	.41
AO	adj. R^2	.53	.54	.56	.60	.59	.69
	resid. MS	.21	.21	.20	.18	.18	.14
DM	adj. R^2	.43	.50	.47	.55	.67	.76
	resid. MS	2.62	2.30	2.45	2.05	1.54	1.09
DO	adj. R^2	.74	.56	.91	.89	.94	.95
	resid. MS	1.22	2.08	.41	.52	.28	.26
CF	adj. R^2	.58	.48	.76	.67	.84	.87
	resid. MS	.22	.27	.13	.17	.08	.07
FO	adj. R^2	.30	.11	.52	.33	.65	.66
	resid. MS	.15	.19	.10	.14	.07	.07
PP	adj. R^2	.13	.14	.27	.25	.35	.35
	resid. MS	.09	.08	.07	.07	.06	.06
		Fiscal year 1988					
AM	adj. R^2	.33	.31	.44	.39	.50	.55
	resid. MS	1.19	1.23	1.00	1.08	.89	.81
AO	adj. R^2	.53	.50	.73	.72	.77	.80
	resid. MS	.32	.34	.18	.19	.16	.14
DM	adj. R^2	.44	.49	.54	.65	.72	.75
	resid. MS	2.96	2.73	2.47	1.87	1.51	1.35
DO	adj. R^2	.78	.60	.90	.89	.95	.96
	resid. MS	.62	1.10	.27	.31	.15	.11
CF	adj. R^2	.61	.42	.66	.49	.76	.84
	resid. MS	.40	.59	.35	.52	.25	.17
FO	adj. R^2	.36	.19	.56	.52	.68	.73
	resid. MS	.13	.16	.09	.09	.06	.05
PP	adj. R^2	.39	.36	.58	.57	.62	.64
	resid. MS	.07	.08	.05	.05	.05	.04
		Fiscal year 1989					
AM	adj. R^2	.41	.38	.49	.45	.54	.61
	resid. MS	1.15	1.22	1.00	1.08	.90	.76
AO	adj. R^2	.51	.50	.74	.62	.79	.82
	resid. MS	.44	.45	.23	.34	.19	.16
DM	adj. R^2	.48	.54	.64	.74	.76	.77
	resid. MS	2.81	2.49	1.92	1.37	1.27	1.21
DO	adj. R^2	.84	.76	.87	.86	.91	.92

(continued)

Model		C	D	CL	DL	CL-DL	FIN
	resid. MS	.34	.51	.28	.29	.18	.16
CF	adj. R²	.69	.52	.69	.61	.85	.88
	resid. MS	.67	1.04	.66	.84	.31	.25
FO	adj. R²	.24	.13	.40	.33	.54	.61
	resid. MS	.15	.18	.12	.13	.09	.08
PP	adj. R²	.50	.47	.52	.49	.57	.60
	resid. MS	.02	.02	.02	.02	.02	.02

Statistical tests, called "partial F tests,"[3] were used to test for the significance of differences between pairs of models in tables V.1 and V.2. The selected models generally provided a significantly better fit to the data than models using fewer predictors. (See tables V.1 and V.2.) On the other hand, extending the selected models to include additional predictors generally did not significantly improve goodness-of-fit, unless the improvement could plausibly be interpreted as the result of overfitting because of the small numbers of cases or defendants that were represented in the additional predictors.

4. Plots of residuals and test for outliners. Global goodness-of-fit measures can give an incomplete picture of the empirical adequacy of a model. Complete evaluation of the selected models required analyses of residuals, i.e., inspection of plots of the differences between the reported and predicted FTES of USAOs.[4]

Except for mild positive heteroscedasticity (i.e., unequal error variance) in a few plots, no patterns indicative of systematic lack-of-fit were discerned in plots of the studentized residuals (i.e., the residuals divided by their estimated standard errors) by the corresponding predicted values under the model.[5]

Few outliers were detected in the residual plots of the selected models of the preceding section. To formally test for outliers, we applied the outlier test presented by Weisberg.[6] We favored retaining the few outlying observations in the regression estimation because the outliers might be due to unusual staffing situations and because omitting an outlying USAO

---

[3]S. Weisberg, Applied Linear Regression, p. 88.

[4]S. Weisberg, Applied Linear Regression, p. 120.

[5]S. Weisberg, Applied Linear Regression, p. 120.

[6]S. Weisberg, Applied Linear Regression, pp. 116 and 226.

from the estimation might misrepresent the cases and defendants in the litigation area in which the outlier was observed and in the total workload of the outlying USAO.

5. Estimation technique. The final estimation of the selected criminal and civil weighting models used data for all three years, 1987 through 1989. That is, the estimated case and defendant weights were estimated using a total of 3 x 91 = 273 observations. In each litigation area, the case and defendant weights were assumed to be constant during these years. Hence, even though 3 years of data were employed, the total numbers of parameters estimated equaled 19 for civil litigation areas (i.e., 18 weights and an intercept term) and 17 for criminal litigation areas (i.e., 16 weights and an intercept term). Each model was estimated with an intercept term because of the fixed costs of litigating cases in the area. In each litigation area, however, the estimated intercepts were close to zero.

The estimation technique assumed that observations of the same USAO in different years were autocorrelated (due, e.g., to cases continuing from one year to the next or to persistent sources of error in the USAO), but that observations of different USAOs were independent. Hence, the technique is a special case of Zellner's "seemingly unrelated regressions."[7] The estimation was implemented using the SAS SYSLIN procedure.[8] However, customized software had to be fashioned to compute consistent standard errors of the predicted FTEs under the Zellner-type model.[9]

6. Precision of estimated staffing ratios. The staffing ratio of USAO i in litigation area j equals the ratio of the predicted time of USAO i in litigation area j, say P_{ij}, and the reported time of the USAO in litigation area j, say R_{ij}. The variance of P_{ij} was estimated following Johnston.[10] Since R_{ij} is constant, the standard error of the estimated staffing ratio equals

$$SE(Ratio_{ij}) = SE(P_{ij})/R_{ij}.$$

[7] A. Zellner, A., Journal of the American Statistical Association (1962), pp. 57:348-364, also called "joint generalized least squares;" and G.C. Judge, et al., The Theory and Practice of Econometrics.

[8] SAS Institute, SAS/ETS User's Guide, Version 5, (Cary, N.C.: 1984).

[9] A. Zellner, Journal of the American Statistical Association; J. Johnston, Econometrics, 2nd ed. (New York: McGraw-Hill, 1972), p. 126.

[10] J. Johnston., Econometrics, p. 126.

For the total criminal (or total civil) area, the variance of predicted time was estimated under the assumption that the predicted values in the different criminal (civil) litigation areas are independent.

Inferences of under- and overstaffing (see app. III) are based on the estimated 95-percent confidence intervals of the understaffing ratios. These confidence intervals assume that prediction error is normally distributed. Given that predicted times are weighted sums of the predictors and that the total criminal (or total civil) predicted times are computed by summing across litigation areas, the assumption of normality will be approximately true by virtue of the central limit theorem (unless the error distributions are severely skewed).[11]

---

[11]See G. Snedecor and W. Cochran, Statistical Methods, 7th edition (Ames, Iowa: Iowa State University, 1980), ch. 4.

Documentation of EO and AO Data Used in Matching Case Records and Modelling Resource Disparities

This appendix provides additional information on the EO and AO data files used in the weighting models. It documents the records excluded from the criminal and civil analysis, the rules that were used to reconcile data from the EO and AO criminal data files and to aggregate information from the defendant to the case level, and the computer program that was developed to match EO and AO criminal case records.

Criminal Data Files

Our criminal workload analyses used information from both the EO and AO case tracking information systems. Records in these two systems affecting U.S. Attorney workload were selected, aggregated from the defendant to the case level, and matched in order that information from both systems could be used in the analyses.

For each of fiscal years 1987, 1988, and 1989, three EO and two AO data files were used to extract the data that were needed to measure workloads for the fiscal year. The three EO files were

- the EO main file, which contains summary case data by defendant, including case status, sentence imposed, last major event, matter date, and case date;
- the EO charge file, which contains information on the status and disposition of all criminal charges in a case by defendant; and
- the EO event file, which has data on the major events in a case by defendant, including event type and date.

Both of the AO files contain summary case data by defendant, including date of filing, major offense, disposition date, and whether a trial was held. The first file has information on all cases that were filed during the fiscal year. The second contains information on cases that were pending at the end of the previous fiscal year. The combined universe of the two files comprises all cases that were open in federal district court during the fiscal year.

The information in each data file was aggregated from the defendant to the case level before matching. Cases in each file that were not relevant to U.S. Attorney workloads in the fiscal year were dropped before matching. The EO case data were then separated into two files. The first file contained cases that were active in federal district court during the fiscal year under study. This file was matched with the AO case data. The second file contained criminal matters and appeal cases that could not be matched to AO data since the latter data source does not include matters and appeals. (Only EO data were available to gauge matter and

appeal workloads in the workload analyses.) Both EO main data files were matched with the EO event file in order to supplement the information in these files with auxiliary information on grand jury, indictment, and trial proceedings.

Table VI.1 summarizes the AO and EO data files and variables that were used in calculating the major input variables used in our criminal resource allocation model.

Table VI.1: AO and EO Files and Variables Used in Calculating Input Variables for GAO Resource Allocation Model

| Model input variable | Computer file and input variables | | | |
	Combined AO master files	EO criminal main file	Criminal charge file	EO criminal event file
Program category	Major offense code	Program category [a]		[a]
Felony/ misdemeanor	Offense level	[a]	Misdemeanor/ felony	[a]
Trial during fiscal year	Trial began date, jury trial code, disposition date	Event code, event date	Disposition reason, disposition date	Event code, event date
Indictment during fiscal year	Indictment date	[a]	Charge type, charge activity date	Event code, event date
Case filed during fiscal year	Date filed	Case date, defendant disposition, disposition date	[a]	[a]
Number of defendants	Court code, docket #, office code	District code, USAO #, defendant #	District code, docket #, defendant #	[a]
Grand jury during fiscal year	Indictment date	Event code, event date	[a]	Event code, event date
Appeal case	[a]	USAO number, court type, court docket #, district code	[a]	[a]

[a]Not applicable to this file.

Selection of Matters and Cases Affecting U.S. Attorney Criminal Workload During Fiscal Years 1987, 1988, and 1989

The data files obtained from the EO contained information on a defendant basis for all criminal matters and cases reported to the EO by the 94 U.S. Attorney offices. The files are not restricted to cases that were actively pursued by AUSAs during the fiscal year. We excluded the following kinds of cases before matching and analysis: minor offense cases; prior-year fugitive, mental institution, and pretrial diversion cases; old inactive cases; cases transferred to other districts in prior years; cases closed in prior years; cases in which the U.S. Attorney's office was not

involved; records opened in error; and special noncase records used to track victims, special witnesses, and certain property items. Table VI.2 presents, for each fiscal year, the number of defendant records received from the EO, the number excluded from analysis by reason, and the number of defendant records included in our analyses. Table VI.3 presents, for each fiscal year, the number of AO defendant records received by fiscal year, the number dropped from our analysis by reason, and the number of records retained for additional analyses.

Table VI.2: Summary of EO Criminal Defendant Records Included in Analysis, by Fiscal Year

	Fiscal year		
	1987	1988	1989
Total number of defendant records	**223,159**	**233,359**	**247,712**
Records dropped from analysis			
Noncourt cases[a]	477	53	7
U.S. Attorney not involved	1,272	1,174	1,158
Cases closed in prior year	6,220	3,250	1,935
Prior-year fugitive, mental institution, or pretrial diversion cases	3,121	6,111	8,553
Special noncase records	278	470	329
Old inactive cases	750	589	477
Prior-year transfer cases	130	150	175
Records opened in error	25,002	26,998	29,935
Total number of records dropped	**37,250**	**38,795**	**42,569**
Total number of defendant records selected for case analysis	**185,909**	**194,564**	**205,143**

[a]Magistrate, petty, and minor offense cases that involve little or no U.S. Attorney time.

Table VI.3: Summary of AO Criminal Defendant Records Included in Analysis, by Fiscal Year

	Fiscal year		
	1987	1988	1989
Total number of defendant records	**91,032**	**92,660**	**107,725**
Records dropped from analysis			
Cases closed in prior year	0	4	4
Prior-year fugitive, mental institutions, or pretrial diversion cases	13,484	13,571	14,270
Old inactive cases	408	440	0
Records opened in error	587	586	196
Total number of records dropped	**14,479**	**14,601**	**14,470**
Total number of defendant records selected for case analysis	**76,553**	**78,059**	**93,255**

Consolidation of AO and EO Defendant Data on Case Basis for Matching and Final Analysis

The EO and AO defendant records appropriate for analyzing U.S. Attorney workload were consolidated on a case basis. Executive Office records were combined using a combination of EO district code and case number, while Administrative Office records were combined using AO district code, district office code, and standard docket number. Both combinations produced a unique identifier for combining all the defendant records pertaining to a particular case. During the consolidation process, each AO and EO case was categorized as a felony, misdemeanor, or petty case. EO cases were additionally categorized as to whether they were an active district court case or in either appeal or matter status. This second categorization was required because the AO records used in our analysis were from the AO's active record file; consequently, only active EO district court cases could be matched with them.

A case was categorized as felony, misdemeanor, or petty as follows:

- a felony case if any of the defendants in that case were charged with a felony,
- a misdemeanor case if none of the defendants were charged with a felony and at least one of them was charged with a misdemeanor, or
- a petty offense case only if none of the defendants had a felony or misdemeanor charge.

All cases identified as petty offenses or active cases before a magistrate were dropped from our analysis because, according to EO officials, they accounted for little or no U.S. Attorney workload.

A similar hierarchy was used to categorize EO cases as being active in district court, on appeal, or in matter status. An EO case was considered an active district court case if any of the case's defendant records were active during the fiscal year; an appeal case if none of the defendant records were active during the year, and at least one of them was in appeal status; or as a matter if all of the defendant records were in matter status. Tables VI.4 and VI.5 summarize the consolidation and categorization of EO and AO case records used in our analyses.

Table VI.4: Summary of Consolidation of EO Criminal Defendant Records Onto a Case Basis for Analysis

	Fiscal year		
	1987	1988	1989
Total number of defendant records selected for case analysis	185,909	194,564	205,143
Number of consolidated case records			
Matters and cases on appeal	102,816	107,136	108,720
Active cases	46,242	48,064	52,837
Total number of cases	**149,058**	**155,200**	**161,557**
Less petty and magistrate cases	0	255	73
Total number of cases used in analysis	**149,058**	**154,945**	**161,484**

Table VI.5: Summary of Consolidation of AO Criminal Defendant Records Onto a Case Basis for Analysis

	Fiscal year		
	1987	1988	1989
Total number of defendant records selected for case analysis	76,553	78,059	93,255
Total number of consolidated cases	55,453	55,865	65,799
Less petty and magistrate cases	14,406	14,765	15,486
Total number of cases used in analysis	**41,047**	**41,100**	**50,313**

Matching of Active EO Cases With AO Case Data

The EO and AO active cases were matched using four variables: AO district code, standard docket number, district court office code (satellite court), and case filing date. Before matching, the EO district code had been converted to the appropriate AO code, and the standard docket number and district court office code were extracted from the court docket number reported in the EO file.

The district court office code was added because the standard docket number alone was not a unique identifier in some districts. Certain judicial districts assign docket numbers on a district court office level rather than on a districtwide basis. Consequently, in these districts, cases adjudicated at different court offices could have the same docket number. Because of time constraints, only court office codes for districts with a significant number of duplicate docket numbers were allowed to be extracted. Court office codes were extracted for the judicial districts of Arizona, Connecticut, middle Florida, Minnesota, southern Mississippi, eastern North Carolina, western North Carolina, southern Ohio, eastern Tennessee, northern Texas, southern Texas, western Texas, and eastern

Virginia. Two additional steps were taken to compensate for the potential of duplicate docket numbers in districts where court office codes could not be extracted. A special variable, indicating nonunique docket numbers within a district, was created and added to every AO and EO case; and additional comparisons using case dates were made in matching AO and EO cases. Together these additional steps greatly reduced the chances of incorrect AO-EO matches being made because of nonunique docket numbers within a district.

The matching process was comprised of seven passes of the AO and EO case record files against each other. The first pass attempted to match only those records with unique docket numbers. These records were then excluded from the match process. Match passes two through seven attempted to match case records in which the docket number was not unique for either the AO or EO file.

During match passes two through seven, only the first case with a given docket number from each file was matched. If no match was found, the case was categorized as a nonmatch. If a match was found, an additional comparison was made between case dates. Matches in which the AO and EO case dates were within 31 days of each other, or the month and day were the same, or the day and year were identical were accepted as proper matches. Those not meeting one of these date criteria were reclassified as nonmatches. Once a determination of match or nonmatch was made for a case record it was excluded from further matching. Table VI.6 summarizes the results of the match process for fiscal years 1987, 1988, and 1989. The first match pass of unique docket numbers accounted for 95 percent of the matches in fiscal years 1987 and 1988 and 92 percent in fiscal year 1989.

Table VI.6: Summary of Results of Match of AO and EO Criminal Record Files, by Fiscal Year and Match Pass

Match pass	Fiscal year		
	1987	1988	1989
EO records entering match	46,240	48,049	52,837
AO records entering match	41,047	41,100	50,313
1st pass matches	28,396	29,865	35,259
2nd pass matches	723	697	1,321
3rd pass matches	93	127	487
4th pass matches	283	281	581
5th pass matches	342	349	438
6th pass matches	30	24	77
7th pass matches	3	4	19
Total records matched	29,870	31,347	38,182
Percentage EO records matched	64.6	65.2	72.3
Percentage AO records matched	72.8	76.3	75.9

Civil Data Files

Our civil workload analyses used only the EO case tracking information systems. As in the criminal analyses, we excluded several classes of civil records from the analysis because they involved little or no involvement by AUSAs. In addition, we excluded a small number of civil cases that were unclassifiable with respect to civil litigation area. Table VI.7 presents, by fiscal year, the number of records received from the EO and the number excluded from our analyses by reason.

Table VI.7: Summary of EO Civil Records Excluded From Analysis, by Fiscal Year

	Fiscal year		
	1987	1988	1989
Total number of EO records before exclusion	293,204	290,572	275,197
Records dropped from analysis			
Noncourt cases	46	27	17
U.S. Attorney not involved	4,831	7,869	7,653
Case closed in prior year	9,430	6,221	3,794
Old inactive cases	6,120	4,358	3,153
Prior-year transfer cases	89	60	78
Cases handled by paralegals	27,961	24,429	24,431
Appeal and collection cases and cases in Virgin Islands, Northern Mariana, & Guam	5,412	5,692	6,208
Cases excluded for more than one reason	−647	−349	−206
Total number of records dropped	53,242	48,307	45,128
Total number of records used in analysis	239,962	242,265	230,069

The aggregation of civil data from the defendant to the case level was carried out using the same procedures that were applied in the criminal EO data management.

Comments From the Department of Justice

U.S. Department of Justice

Washington, D.C. 20530

DEC 2 1 1990

Mr. Richard L. Fogel
Assistant Comptroller General
General Government Operations
General Accounting Office
Washington, D.C.   20548

Dear Mr. Fogel:

This responds to your October 1990 draft report entitled "U.S. Attorneys: Better Models Can Reduce Resource Disparities Among Offices." We appreciate the considerable efforts of the General Accounting Office (GAO) in preparing its draft report reviewing the Department of Justice's (DOJ) process for allocating attorneys among the United States Attorneys' offices (USAOs). It is evident that much time and thought has gone into this project. GAO has produced a statistical model which we believe will be helpful as an additional source to be consulted before allocation decisions are made by DOJ in the future.

The GAO report on DOJ's attorney allocation process was prepared in response to an apparent perception of attorney staffing disparities among USAOs (see page 1.) It is interesting to note however, that when the GAO model was applied to allocations made by the Department in Fiscal Year (FY) 1989 and FY 1990, GAO recommended allocations differed only slightly from actual DOJ allocations. As the report states at page three: "[the] model's allocation showed a high level of agreement with the allocation the Department actually made for most offices, . . . "

See comment 1.

For the 1989 allocation, the report does indicate that ". . . the two allocations differed substantially for some offices." Specifically, in 18 of the 91 offices, the allocations differed by five Assistant United States Attorneys (AUSAs) or more. After a review of the methodology and theory behind the GAO report, we feel that the discrepancies between GAO allocations and actual DOJ allocations are accounted for in large part by the fact that the GAO model does not take into consideration several factors which are of importance to the Department in determining allocations. One such factor is the number of federal agents assigned to certain areas of the country, for example, high intensity drug trafficking areas. Another major factor not considered by GAO is the detailed evaluations completed on each office by the peer review program coordinated by the Evaluation and Review Staff, Executive Office for United States Attorneys

Richard L. Fogel   2

(EOUSA), and the professional judgments as to workload and office
needs that are made during, and as a result of, those
evaluations. Many additional factors considered in the
allocation process are detailed below. These factors are
significant in making final allocation decisions.

Because of the general high level of agreement in the results
produced by the two allocation processes, we do not feel a point
by point response to the GAO report is necessary. There are
nevertheless a few important observations that we would like to
make.

See comment 2.

Defining and Using Workload Information. The GAO report states
that the ". . . Department's allocation process does not
adequately account for differences in complexity of legal
workload among USAOs workloads," such as considering types of
cases, number of defendants, number of indictments and trials
(see page 2.)   On the contrary, the Information Management Staff
of the Executive Office is consulted to determine workload data
applicable to each district in the particular area of crime to
which resources are to be allocated.

For example, in the FY 1990 allocation of AUSAs authorized for
violent crime, workload factors considered included the number of
violent crime cases filed (including weapons and drug cases),
number of defendants and average number of cases per AUSA for
each district. For the FY 1991 allocation of Financial
Institutions Reform, Recovery, and Enforcement Act (FIRREA) and
Organized Crime Drug Enforcement Task Forces (OCDETF) AUSAs,
workload factors considered included matters pending, cases filed
and defendants convicted in the areas of financial institution
fraud and drugs. In addition to these case-related factors, many
other factors, such as number of judges and locations of courts,
total trials handled, length of trials, cases handled per AUSA,
number of branch offices, average AUSA workweek, number of
additional federal agents to be allocated, evaluation reports,
travel time, grand jury time, etc., were considered.

See comment 3.

In arriving at their model, GAO statisticians grouped the
workloads of all the districts together by litigation category
and then determined the average time spent to process cases of
each type. This methodology is based on the premise that all
USAOs are essentially alike, when in reality there are meaningful
differences between the districts, and the way litigation must be
conducted in each district, which materially affect workload and
its resolution. Some of these factors are as follows:

     --In criminal cases the number and type of investigators and
       the investigative resources available has a dramatic
       effect on the amount and quality of the workload. For
       example, with no Postal Service Inspectors readily
       available there will be fewer mail fraud cases.

Richard L. Fogel   3

   --Demographics also affect the workload such that large
     cities, rural areas and different areas of the country
     have totally different types of workloads. For example,
     California, Texas, Arizona and New Mexico have unique
     immigration workloads. In civil matters the presence or
     absence of military health care facilities has a drastic
     impact on workload.

See comment 4.

   --A district's court structure and judicial policy play
     significant roles in how USAOs litigate. Travel between
     USAO headquarters and branch offices, and to multiple
     court locations, is a factor in some districts. A
     judicial decision to enforce a "rocket docket" causes even
     complicated cases to move much more quickly than in other
     districts. In some districts court calendar control
     exercised by judges moves cases to prompt resolution,
     while in other districts a more laissez-faire approach
     results in more time in court.

See comment 5.

   --Historical statistics of cases prosecuted by the USAOs
     indicate that 80% of all cases are handled by 20% of the
     offices. This indicates that an "average" is not a
     meaningful or useful indicator.

See comment 6.

At pages 38-39, the GAO report lists six (6) types of "legal
proceedings" and hypothesizes time requirements for each,
assuming that "matters" are least time consuming and moving up to
"trials" which GAO assumes are most time consuming. The
progression is as follows: Matter; Misdemeanor; Pending Felony;
New Felony; Indictment; and Trial. A "pending felony" has no
meaning for the Department; the activities GAO would include in
this category would be either a "matter" (under investigation) or
an "indictment" (charged). Neither is it clear what a "new
felony" means because, to the Department, these activities are
also either a "matter" or an "indictment." Further, GAO's time
requirement hypothesis is not completely reliable because, in
fact, a "matter," which is a pending investigation, can be
extremely time consuming. Assuming a case with a "trial" is the
most time consuming type is also incorrect, as there are many
one-day trials in simple cases which take much less total AUSA
time than some "matters" which are under investigation. Under

See comment 7.

GAO's hypothesis, a complex bank fraud trial of a bank president
is weighted the same as the much less complex trial of a bank
teller, and districts which shun small cases to work only
complicated cases are not recognized.

See comment 8.

Analyzing and Applying Findings. The report indicates that there
is an uneven distribution of resources among USAOs. (See page
2.) In the model, resources are distributed "unevenly" only in a
relative sense. At page 14 the report states "[s]ince all
inferences are relative to other offices, inferring that an
office is overstaffed does not imply that the office does not

Richard L. Fogel   4

need additional attorneys." The converse then must also be true:
inferring that an office is understaffed does not imply that the
office needs additional attorneys. If this is the case, DOJ must
still rely on other factors to make the final determination of
whether or not to allocate additional AUSAs to an office.

The GAO model is based on "expected" or "predicted" attorney time
expended by each office to process its workload in comparison to
the "average" office. The report concludes that any office
expending more time than the average office may be overstaffed
and any office expending less time than the average may be
understaffed. (See page 83.) The report acknowledges, however,
that an alternative interpretation of this data is possible: An
office which is using less time than expected may not be
understaffed, but rather just be more efficient. An office using
more time than expected may not be overstaffed but just less
efficient. The report concludes that this alternative
interpretation does not really matter because "[i]t is reasonable
to allocate more positions to relatively efficient offices than
to relatively inefficient offices." (See page 52.) This may or
may not be true. In applying GAO's model, DOJ would have to keep
in mind that an indication that an office is understaffed might
instead mean that the office is very efficient. DOJ must then
decide if that is an appropriate reason to give the office
additional AUSAs. In making its decision, DOJ would have to
determine whether an office really needs more AUSAs because it is
efficient; or, in the alternative, whether an office should be
penalized and not receive additional AUSAs because it is not
efficient. Central to DOJ's determination is whether the
decision will accomplish a nationwide improvement in processing
workloads.

Other alternative interpretations are listed in the GAO report at
pages 51-54. Offices using less time than predicted might be
understaffed or efficient and they might be producing a poorer
quality of legal work because they are handling more cases and
defendants per AUSA. When all these possibilities are looked at
together, it is clear that any particular allocation
recommendation based on the application of the GAO model could
indicate a number of different things about a given office. For
example, if GAO's model shows that an office is understaffed it
means the office is understaffed only in relation to other
offices which may mean the office does not really need more
AUSAs. It may also mean that the office is not understaffed but
merely very efficient; or the office is efficient but it is
producing a lower quality of work; or the office is understaffed
and is either producing good work or a lower quality of work.
Judgement based on knowledge of the office and the nature of its
caseload will have to be applied in making allocations.

Richard L. Fogel   5

<u>Conclusion.</u> As noted above, DOJ is aware that workload and types
of cases handled by each district office are important factors to
be considered in making allocations. Available data has been
utilized to achieve that determination. In addition, we are
working on a "case weighting" system and are attempting to
develop different approaches to capture the necessary data which
will produce reliable case weighting indicators. Moreover, as
the report itself states, the model "should be regarded as an aid
to decision-making rather than as a substitute for experienced
professional judgment." (See page 54.) It also states that
"[p]rofessional judgment of responsible Justice Department
officials is obviously critical to the proper interpretation and
use of the models." With the exception of a few points discussed
above, we agree that such a model could be helpful in making
resource allocations when considered in conjunction with other
factors. In fact, EOUSA requested that GAO apply their model to
assist us in making recommendations for the FY 1991 allocations
of FIRREA and OCDETF AUSA positions. Unfortunately, the model
was unable to be applied within our short time frame. We look
forward to using the model when future allocations are made.

We thank you for the opportunity to comment on the draft report.
Please feel free to contact me if you should have any questions
concerning our response.

Sincerely,
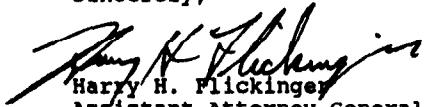
Harry H. Flickinger
Assistant Attorney General
   for Administration

The following are GAO's comments on the Department of Justice's letter dated December 21, 1990.

GAO Comments

1. Justice said that "the GAO model does not take into consideration several factors which are of importance to the Department in determining allocations."

As discussed in comments 3 and 7, there are a number of variables identified by Justice that show considerable promise for incorporation in our workload model. The problem is that their databases (time reporting and case tracking systems) do not capture the required data in sufficient detail for statistical modelling. In appendix I, we pointed to the need to continually improve the model by including additional potentially relevant factors as these factors are identified and measured.

2. Justice disagreed with our statement in the report that its allocation process does not adequately account for differences in complexity of legal workload. It said that "on the contrary, the Information Management Staff of the Executive Office is consulted to determine workload data applicable to each District in the particular area of crime to which resources are to be allocated . . .. For example, in the FY 1990 allocation of AUSAs authorized for violent crime, workload factors considered included the number of violent crime cases filed . . . number of defendants and average number of cases per AUSA for each district."

We recognize in our report that Justice considers qualitatively a number of case-related and other factors in deciding where to allocate additional attorney positions. However, it is unclear how the data are used in decisionmaking. In particular, Justice does not say how relative weights were assigned to the case-related and other factors mentioned in connection with the fiscal year 1990 violent crime allocation. One strength of our model is that relative weights are assigned to the various factors, and therefore they can be quantitatively applied in the model.

3. Justice said that length of trials and grand jury time among other variables are considered in the allocation process. It also stated that "a district's court structure and judicial policy play significant roles" and "travel between USAO headquarters and branch offices, and to multiple court locations, is a factor [to be considered] in some districts."

We agree with Justice and think that these variables are promising choices for improving the specification of the workload. We can think of

no reason why these variables cannot be quantified by criminal categories and explicitly entered and tested out in future workload models.

4. Justice said that "in arriving at [our] model, GAO statisticians grouped the workloads of all the Districts together by litigation category and then determined the average time spent to process cases of each type. This methodology is based on the premise that all USAOs are essentially alike, when in reality there are meaningful differences between the Districts."

In our model, we showed (1) that differences in the staffing requirements of offices could be better explained in terms of the workload requirements of the offices (rather than in terms of more tenuously connected variables, such as the urban population of the district) and (2) that differences in the workloads of offices were explainable using relatively few predictor variables. As documented in our report, the closeness of fit of our models to the data appears to corroborate these assumptions.

Indeed, it is reasonable to think that workloads directly determine staff time requirements. Some of the other variables mentioned in Justice's response seem to be less clearly germane to staffing needs. For example, urban population size is likely to affect staffing requirements only through indicators of workload. If so, adding this variable to factors that already indicate workload would not improve the prediction of attorney time expenditures. This is not to say that management should not consider these and other variables along with the model results as part of the overall decisionmaking process.

In summary, we believe there are advantages in specifying the model strictly in terms of direct determinants of staffing requirements, i.e., legal workloads, rather than in terms of variables that are less closely associated with staffing requirements.

5. Justice said that "historical statistics of cases prosecuted by the USAO's indicate that 80% of all cases are handled by 20% of the offices. This indicates that an 'average' is not a meaningful or useful indicator."

The case and defendant weights estimated in our workload model are, indeed, averages. Specifically, each weight in our model estimates the unweighted average time expenditure, across the universe of USAOs, for a particular case or defendant attribute after other attributes of the case

or defendant are taken into account. <u>Weighted</u> averages, where the contribution of each office is proportional to its total time expenditure, could be more germane to the allocation decision than unweighted averages.

We considered using weighted averages rather than unweighted averages in the model. The main reason we rejected the weighted averages is technical. "Weighted least squares," a standard estimation procedure in econometrics, is known to give more precise results only when the sample size is quite large, a condition not satisfied in our application. In the small-sample case (say, less than 100 observations), it turns out that unweighted estimation often does better.

6. Justice said that it was not clear what a pending felony or new felony means.

In appendix I, we have clarified these terms as cases that had been filed in court either before (pending) or during the referenced year.

7. Justice said that it did not consider "GAO's time requirement hypothesis . . . to be completely reliable because . . . a matter, which is a pending investigation, can be extremely time consuming. Assuming a case with a 'trial' is the most time consuming type is also incorrect, as there are many one-day trials in simple cases which take much less total AUSA time than some 'matters' which are under investigation." Justice further said that "under GAO's hypothesis, a complex bank fraud trial of a bank president is weighted the same as a much less complex trial of a bank teller."

We agree with Justice's statements. These comments point to what we think is, arguably, the most important direction for improving the workload model, namely, classifying cases more finely according to the type of litigation. Our model distinguishes six types of criminal litigation and seven types of civil litigation. Clearly, different kinds of cases subsumed under such broad categories as "economic crime" and "drug crime" cannot reasonably be regarded as equally time-intensive, even when all measured attributes, such as number of defendants and type of disposition during the year, are the same.

The main obstacle to refining the classification of types of litigation is the U.S. Attorneys' time expenditure data rather than their workload data. The U.S. Attorneys' time reporting system does not report the amount of time attorneys spend on individual matters and cases. It

reports only the amounts of time (measured in FTEs) that were spent by U.S. Attorneys' offices in nine broad criminal program categories, such as official corruption, economic crime, and narcotics. A major step forward was made in 1989 when the Executive Office for United States Attorneys began collecting FTE data on bank fraud as a separate category, as distinguished from other kinds of economic crimes. (We tried to develop a separate bank fraud model but found that the 1989 data were too sparse to support statistically defensible inferences.) Still, as Justice's comment indicates, the global category "bank fraud" encompasses widely divergent kinds of cases. Other leading examples are the FTE reporting categories "drug" and "Organized Crime Drug Enforcement Task Forces," which do not permit distinguishing among drug possession, drug marketing, and drug manufacturing.

In summary, Justice's comment identifies an important area in which improving the workload model depends on improving the data, specifically, refining the classification of litigation areas ("program categories") in reports of time expenditures.

8. We agree with Justice's observations on the alternative interpretations that can be made on the results of the workload weighting model. We say in the report that professional judgment should be applied to interpreting the model results, and we recognize and discuss alternative interpretations that could be drawn from the model results.

Major Contributors to This Report

General Government Division, Washington, D.C.	James Burow, Assistant Director, Administration of Justice Issues Barry P. Griffiths, Project Manager Robert A. Johnson, Senior Statistician William J. Kruvant, Advisor Abiud A. Amaro, Evaluator Rebekah K. Locklear, Secretary/Typist
Cincinnati Regional Office	Norman A. Hofmann, Senior Programmer/Analyst

United States
General Accounting Office
Washington, D.C. 20548

Official Business
Penalty for Private Use $300