Making justice count: a review of the issues and obstacles affecting the use of management statistics for improving Court performance in Latin America

Linn A. Hammergren
The World Bank

The Place of Information in Judicial Reforms: As experts in the material are coming to realize, a judicial reform, modernization, or development program without empirical data is a proposal armed in an informational vacuum. When the latest round of Latin American reforms began in the early 1980s, planners had to rely on conventional understandings of system problems and their causes, as no judiciary in the region kept good statistics on its own performance. Twenty years later, we are beginning to see that much of this conventional wisdom was misleading when not completely wrong. Presumably overworked judges turn out to have relatively light caseloads; delays are not as exaggerated as the usual collection of anecdotal evidence indicates and may be less problematic than the large number of cases which either never are resolved, or reach judgments that are not enforced; “underfunded” judiciaries are found to receive a higher than average proportion of the national budgets; and much judicial workload may be composed of cases which arguably should not reach the courts.

These and many other “discoveries” were first advanced by targeted research projects, relying for the most part on privileged, extra-official access to court documents, or on alternative sources of primary data. Courts’ reluctance to cooperate or policies of denying access were sometimes an obstacle, but more often the problem has been the poor state of court record keeping. However, more recent studies have been aided by the increasing availability of better collections of judicial statistics within the region and internationally. Much more could be known if the region’s judicial data bases were further improved and better utilized. Contrary to the view that justice is a quality, not a measurable quantity and that statistics on court operations are of little help, in fact may distort efforts to improve its delivery, numerical data on judicial performance have applications far beyond the basics of judicial productivity (i.e. whether judges are resolving a sufficient number of cases within the legal time limits). They can be used to guide code reform efforts by indicating where the real bottlenecks or irregularities occur, help plan programs to expand access to marginalized groups, help judicial leaders defend requests for more resources and determine where to place them, and indicate where alternative mechanisms (including improved administrative dispute resolution) might be most effective. They can also help courts determine where poor performance by other organizations impedes their own output, as well as evaluating the success of internal reforms.

Despite the large investments in information and communication technology (ICT), funded either by donors or the countries themselves, very few judiciaries have used this equipment to enhance their ability to track and evaluate their own performance. The countries with relatively adequate

---

1 These comments reflect the views of the author and are not meant to represent the official position of the World Bank.
3 Access to court data, both case files and statistics, has been a particular problem in Mexico because of a long standing belief that the former belonged to the parties and the latter to the courts. The recent passage of a national access to information law may be changing the situation, but despite the former Supreme Court President’s promise that the courts would become more transparent, progress has been slow. Also in a federal system, the state judiciaries will be on their own to determine their policy. At the other extreme is Costa Rica whose policy of transparency has produced fears that too much information on parties to sensitive cases is available to all comers. Fortunately, this does not affect the broad availability of Costa Rica’s excellent system of judicial statistics.
management statistics (e.g. Costa Rica, Chile, Colombia, Argentina and Mexico at the federal levels, Brazil, for the federal, labor and some state courts) are a small minority, and most of them do not use their contents as a means of monitoring output, identifying problems, and setting reform goals. What data are collected are rarely checked for accuracy and consistency, usually reach the central offices only in the most abbreviated form (i.e. annual filings and dispositions, at most broken down by major legal areas); and are often used only as inputs to annual reports on judicial activities or are posted on the increasingly popular judicial websites. While website availability may be of some use to researchers, the lack of further explanation of and frequent changes in the presentational formats (and apparently in what is included in individual categories) can make even year-to-year trend analysis nearly impossible for outsiders, and presumably for insiders as well. A few countries do use the yearly submissions to evaluate the performance of individual judges, but virtually no one attempts more sophisticated analysis, for example matching distributions of workload and resources, tracking differences in outputs among types of courts or cases, looking for anomalies or bottlenecks in processing, or comparing their own data with the increasingly available international statistical compilations. It is not unusual to find that Latin American judiciaries are unaware of the potential for comparative analysis, do not in fact know that these other data bases exist.

Thus while Latin America’s courts are beginning to catch up in the areas of automation, budgetary endowments (where they even surpass international averages), ratio of courts and judges to population, career and professional development, and the adoption of access-enhancing and alternative dispute resolution programs, they are far behind the international curve as regards the introduction of mechanisms, including but not limited to management statistics, to monitor, evaluate, and develop methods to improve their performance. Their bilateral and multi-lateral partners share part of the blame, but are now attempting to incorporate this missing element in their own assistance programs. However, resistance to the new trends remains high, and there are also signs that the introduction of better monitoring systems guarantees neither their use nor the additional benefits they presumably would bring. The remainder of this paper explores these ideas in greater detail, starting with a review of the current situation, and continuing to an examination of the remaining sources of resistance and the reasons for their persistence, a discussion of the reasons for adopting quantitative monitoring techniques, and finishing with suggestions as to how advances could be made more rapidly.

Progress in Monitoring Performance – the Role of Research and Performance Statistics: Given the progress made on the basis of targeted research and a review of what statistics are available, one might ask why this is not sufficient. The answer comes in two parts: the limitations of the existing research methodologies, and the desirability that the courts themselves do their own monitoring, rather than relying on the work of others. Here we look at the first issue in some detail and begin to address the second.

Research will always be needed to explore uncharted areas, but it is expensive, sporadic, and necessarily limited in its coverage. Moreover, as judiciaries are not research institutes, a reliance on

---

4 Poor development of judicial statistics is not limited to Latin America. However, the last ten years have seen important advances elsewhere, especially as regards efforts to standardize techniques cross nationally. One of the most interesting is the effort within the European Council to develop judicial performance indicators for 45 countries, as reported on the website for the European Commission for the Efficiency of Justice (CEPEJ) http://www.coe.int/cepej. For the most part the countries included have good quality data bases of their own, but the common effort is also helping to improve their contents. For earlier, related efforts see Blank et al 2004, Contini 2000. Pastor 1993 represents the results of an early effort in Spain. CEJA 2003 is a first effort to compile collections of Latin American statistics, drawing on those submitted by national court systems.
research to assess performance requires either their own development of this capability or a continuing reliance on what others choose to explore. Finally, research itself continues to be hampered by poor statistical systems, forcing investigators to develop other methodologies, and often hindering the latter’s efficacy. A quick review of some of the principal research efforts illustrates these points.

Case file analysis, a method used by the World Bank and gaining currency among the region’s academics and some courts, will always be the preferred method for tracking trends not covered in whatever aggregate statistics a court keeps. Still, it would be greatly aided by better automated data systems, of the same sort needed for good statistical sets. At present in most of the region’s courts, case file analysis requires physical access to paper files and the manual extraction of the data they contain. This is costly and time-consuming, and can generate its own errors. Moreover, the absence of good statistics makes it difficult to draw samples or to assess their representative quality. At some future date, case file analysis will be superseded by the direct “mining” of court data bases, but even then, will not replace the systematic tracking of a lesser number of aggregate statistics as a means of monitoring ordinary performance trends. Managers, whether judicial or otherwise, need to be made aware of new developments (and thus the role of data mining) but their day-to-day job is that of supervising objectives and indicators they have already defined.

Because of the expense, labor-intensiveness, and other logistical difficulties currently inherent in case file analysis, those attempting to evaluate court performance have more often relied on other types of data. Most use opinion polls and related techniques to tap users’ assessments of how the courts are doing and what problems they present. While surveys are also costly and time consuming, they can generate data for entire national systems, and offer apparent advantages in cross-system comparisons. For the time and investment it might take to conduct one moderately sized case file study for one district of one country, those using opinion data can develop broad-brush overviews of 100 national judiciaries. For this reason, these techniques have become especially popular with economists attempting comparative macro analyses of judicial outputs and outcomes, and proposing to tease out their relationship with exogenous variables like economic growth, credit availability, and private investment. Their work has both reinforced the economic importance of judicial reform, and produced some interesting hypotheses as to the specific linkages between judicial quality and growth. While judicial reformers can thus be grateful for the impacts on their programs, the approaches do have other serious shortcomings.

Opinion polls are in the end very subjective, and even when modified to tap respondents’ real experience (and not just their perceptions), are overly reliant on the accuracy of individuals’ memories.

5 This involves drawing a random sample of case files and using data drawn from their contents to explore questions about types of cases and how they are processed. As noted below, a shift to fully electronic case files would eliminate much of the labor involved, and might eventually eliminate the need to draw samples – allowing investigators to work with the entire universe of cases. Some developed countries have found funding to apply the technique massively – see for example Australia Law Reform Commission 2000; Ontario Law Reform Commission 1998, and Kritzer 1983.


7 See for example Brazil, Poder Judiciário, 2004. As the Rio de Janeiro courts’ files are web-based, it appears the study was done with electronic case files. Unfortunately the document does not explain this in any detail.

8 For a discussion of some of the practical impediments, see general discussion and methodological annexes in World Bank, 2002, 2003 a and b.

9 The most famous among such studies are probably those done by Dani Kaufman and others at the World Bank Institute. See www.worldbank.org/wbi/governance/governancedata. Related approaches include those pioneered by Robert Sherwood (Sherwood et al, 2000) based on entrepreneurs’ estimates of their losses due to an “inefficient judiciary” and work by Armando Castelar in Brazil (Castelar,2000) and by a local university (ITAM)and Moody’s investor’s service for Mexico. See ITAM, 1999 and Moody’s 2002.
and their ability to quantify answers (‘How much do you spend on legal fees? How often do you pay bribes?’) they may not automatically manage in this fashion.\textsuperscript{10} The use of expert panels and similar techniques, drawing on the assessments and evaluations of a few key informants, often rely on superficial experience and tend to be tainted by the backgrounds of those asked.\textsuperscript{11} The World Bank’s \textit{Doing Business} reports\textsuperscript{12} rely on a few local lawyers in each country to produce results like “the average time to process a debt collection case” or the “number of steps in an eviction proceeding.” As with experiential surveys it is unlikely that most of the respondents have these figures at hand, and moreover doubtful that their experience (as they are usually chosen from elite firms) represents the typical case. One further criticism of all these techniques (and one commonly used by judiciaries to discount their findings) is that they are done by external (to the court, and sometimes the country) groups, depend on their own interests, and thus may not feature what courts want or need to know to improve their own operations. Certainly, the number of steps in a legal proceeding is not dependent on the courts but on a law enacted by the local congress.\textsuperscript{13} Because all these techniques either focus on a few, often equally unusual indicators, or simply on general impressions, they may tell a court it has problems, but provide little useful guidance as to where or why. Some courts have added opinion surveys to their own monitoring techniques, but more as a rough assessment of how they are doing than as a tool central to developing their reform plans.

As much as courts may object to these outside evaluations, they commonly have had little to offer in return. Hence, they have increasingly seen the need to provide their own evidence and thus to develop their own statistics. The fact that the underlying impetus was so often self defense has further implications for how the process evolved, but the important part is that it has started. Also, to be fair, the leaders in the region have often been motivated by other goals, including a sincere interest in using these systems to track and improve performance, at the individual and sometimes the systemic levels. Progress has been uneven, but throughout Latin America, it is at least evident in the development of such elementary statistical measures as the number of judges and staff, their distribution, and the number of filings received and decided in the course of a year. The accuracy of these figures remains in doubt, even such basic ones as the number of judges, but the past few years have seen considerable improvement.

Court statistics can and have been managed without automation. Honduras recently instituted a largely manual system which provides a reasonably good overview of operations in a large portion of its courts.\textsuperscript{14} However, progress has clearly been aided by the region’s vast investments in court automation, although often as an afterthought. In fact because automation was usually introduced for other ends – commonly the more efficient handling of the individual case – the form it took has sometimes impeded the development of statistical systems. Where an electronic case file is envisioned simply as a more efficient means of storing information featured in a written version, it can be difficult if not impossible to use its contents as a source of aggregate data. Most entries may be in text form,

\textsuperscript{10} The answers are inevitably guesstimates and tend, as critics have noted, to be overly influenced by recent or more dramatic events. See Kritzer 1983 and 1999, for discussions of these problems.
\textsuperscript{11} For example, where those asked have most experience with business law, or with human rights, their scoring will clearly be influenced by their background.
\textsuperscript{12} World Bank (2005a) offers the most recent report, with some efforts to correct the methodological shortcomings. However, using “several” as opposed to one lawyer per country is arguably not a substantial improvement. See Pistor (1995) for criticism from a different angle, focusing on the arguments about the “superiority” of common law systems in defending economic rights.
\textsuperscript{13} It also warrants mention that the “number of steps in a proceeding” is not a standardized concept, and consequently subject to considerable variation depending on how each respondent interprets it. Notification might be regarded as one step, or as a stage with a dozen or more steps within it.
\textsuperscript{14} As reported in World Bank documents related to the development of a project in Honduras. See World Bank 2005b.
categories may be inconsistent across and even within single courts, and they may be selected from the standpoint of increasing detail, not facilitating generalization. Where one data enterer writes in homicide, and another inserts the relevant article of the Criminal Code, entries have to be converted to a standard code to be statistically useful. Additionally, there has been minimal attention to accuracy of entries as is easily apparent from what aggregate results are derived, and many of the aggregates are still calculated and submitted manually from individual courts, even in relatively highly automated systems. In effect, many of the “automated” systems are still not better than what other courts manage manually either in detail captured or quality of data.

The underlying and still unresolved obstacle to further advances is the courts’ lack of demand for anything better. Where the purpose of a statistical system is largely to demonstrate that one is working, all that is needed is a gross estimate of workloads. Accuracy may not be a problem if no one is in a position to question it, and certain systematic errors (doubling counting, or other factors inflating the results) can be an advantage. In countries where either judicial leadership or members of the public have begun to look at statistics to assess productivity, there are clear signs that some of those submitting them have sought to inflate their numbers. Interestingly, the opposite has occurred with figures on judicial budgets, for similar reasons. In the early 1990s, there was a sort of informal competition among the region’s courts to see who laid claim to the largest proportion of the national budget. Once observers began to comment on the lack of connection between performance and budget size, courts started to insist that the real figures are much lower, must be shared with other agencies (judicial councils, training, etc), and to find other ways to indicate their relative poverty.

The courts’ narrowly instrumental approach to statistics – as a means of defending attacks against them – is demonstrated by the usual lack of staff qualified to do analysis and develop better systems. Commonly staff working with statistics is either a part of the informatics department, or performs and is qualified for the purely clerical job of entering submissions from lower levels into a central data base. Neither group has any reason to understand what a management information system should contain, or to foresee to what uses it might be put. Absent a demand for anything better on the part of judicial leadership, they have no reason to provide it. There are important exceptions. However, individuals or groups working to provide better analysis or a better system for entering and collecting data usually lack higher up support, on either the administrative or jurisdictional side. Lack of support means not only limited funding – it also translates into missed opportunities for networking with similarly minded colleagues, tapping into efforts underway in other countries, and access to journals and other educational material that might help them develop their own projects.

A recent World Bank sponsored study in Brazil, looking at the quality of management statistics in courts, public ministries, and government attorney’s offices, found that although there were several groups working independently on these issues, few of them had any knowledge that the others existed. This was even true of groups working within the same city, but in different organizations or jurisdictions. The study’s authors thus concluded that if nothing else, they had at least helped to create awareness of what was already occurring in the country and facilitated communication among the participants.

15 Commonly efforts to track annual trends encounter sudden, inexplicable changes, most probably the result of alterations in how records are kept or what is submitted to the central offices.
16 Calculating budgets can be still more difficult than caseloads. For a discussion of the difficulties in doing this for a number of European countries, see Douat 2001.
17 See Hernández Breña 2003 for an example of an NGO’s analysis of Peruvian court data, apparently not informed by familiarity with more standardized indicators and measures. While the analysis provides many insights it is thus difficult to compare the results with international data bases. Discussions with the author suggest he did not know they existed.
18 World Bank, 2005b.
Brazil is an interesting case both because of its high level of automation and the isolated advances already occurring. Following publication of a report by the federal Ministry of Justice, based on the courts’ own statistics, there has been a move to advance system development further. The report came to some fairly negative conclusions about court efficiency, against which the courts’ primary defense was to admit that their own numbers are inaccurate. A similar story in Argentina from several years earlier revolved around a local think tank that, again based on statistics published by the federal courts, concluded that the cost of litigating a civil case in Argentina was twice that of Spain. Again the judges rallied by admitting their own records were not complete. In countries like Brazil and Argentina, where a good deal of the base has been laid, but is just not utilized, an event like this can promote some rapid corrections of course. The results may still not be what an advocate of good internal monitoring would desire, but they constitute several steps in the right direction. A second response to the Brazilian study was the Rio de Janeiro state court’s production of its own research on repeat users of civil courts. The results themselves are important – suggesting that both large private enterprises and the government contribute to court congestion first, by denying services to the public, thus forcing them into court, and second, by appealing every decision, thereby stringing out the final judgments and congesting the court dockets. However, the experience also demonstrates that the level of automation in Rio allows court staff to go directly to the raw data (a web-based collection on all filings) and use this to explore an infinite variety of trends. The state high court has already used this capability to improve its monitoring of performance, although it clearly has not done as much as is possible. Reasons for that failure, like that in other more advanced countries (Chile, Costa Rica, Colombia, and some other Brazilian, Mexican, and Argentine state and federal courts) are discussed below. In these countries, the lack of more progress clearly has reasons other than technical obstacles.

For much of Latin America, however, the road ahead will be a bit longer, and the obstacles, if not entirely technical, are in part rooted there, often including incomplete automation, a multiplicity of case tracking software, a corresponding multitude of formats for entering data, and minimal attention to the quality of what is entered. Submission of even basic data to central offices is often not enforced, meaning that central records capture only a part of it. Enforced or not, it is often done manually, as the systems installed do not generate statistics automatically. Brazil’s multiple systems for collecting statistics may not be surprising for a federal country of that size. In smaller, unitary nations, however, introduction of automation has often been similarly chaotic, and never has been guided by a thought to statistical generation. Countries with better systems have often arrived there slowly, sometimes because of an extra-sector interest in performance data. The two countries, Argentina and Colombia, where Ministries of Justice had until recently managed judicial administration, may owe their better statistics to that fact. Even today in Argentina, the Ministry has the best collection of statistics on its website and in its periodic publications. Had this task been left to the courts or the country’s new judicial council, it seems doubtful that the results would have been as good. Colombia’s judicial council does manage its own statistics, but again because of the earlier impetus lent by the ministry. While Peru’s courts have long been under judicial management, one consequence of the otherwise lamentable executive-driven (Fujimori) judicial reforms of the 1990s was the development of a good data base. Unfortunately, now that the courts have reasserted control, they have done little to preserve the system. Costa Rica’s collection of statistics, and their use for court planning, is a wholly judicial creation, but a unique example in the region. In other countries, so little attention has been

20 FIEL, 1996.
23 For a discussion of these politically driven reforms which arguably left the courts more demoralized and more corrupt than when they started, see Hammergren, 1998.
given to the matter that even a policy change will take some time to produce real effects.

Where judicial leaders’ (in high courts or councils) do not have a managerial outlook, they will not think to request managerial tools. Many Brazilian courts do use their statistics to “grade” individual first instance judges, but Rio is nearly unique in using them to analyze the composition of demand and identify the causes of congestion. It is also unusual, but not unique in Brazil, in using data on individual judges as the basis for a program to help them catch up and stay ahead of their caseload. This is not the most sophisticated use of performance statistics, but it is an important step forward. The Rio courts are also keeping track of the collection of court fees and have increased the returns here, using them to finance still further automation. Other Brazilian courts may or may not capture the necessary data, but they are not using it as well.

Examples like Rio, and the growing pressure from outside the judiciary may hasten the necessary transformation in the rest of the country – from a caretaker outlook to a managerial one, and thus to an understanding of the need for monitoring tools. The question is how to do this in the rest of the region. Before turning to some tentative answers, we look at some other sources of resistance to both kinds of change.

The Sources of Resistance to Change: Resistance to measuring justice or more accurately, the performance of organizations in the sector, tends to be of two types: practical/cultural and ideological/political. The practical and cultural obstacles have been summed up most succinctly by the frequent observation that if lawyers had wanted to study math, they would have chosen another discipline. While perhaps overstated, sector members’ unfamiliarity with quantitative methodologies, and the resultant fears of not being able to work with them, doubtless account for a good deal of resistance. Added to this is the emphasis put on judicial independence in its individual form (thus reinforcing the normal resistance to evaluation of any employee, anywhere), the general lack of management skills in the sector, and the traditional absence of attention to any measurement of output. None of these barriers is unprecedented – they are common in any organizational environment where performance measurement is first introduced. All take time to overcome, but can be treated through technical assistance, training, and demonstrations of the benefits of the new approaches.

Some judiciaries have begun to respond, largely on their own, understanding that if only to better fight for their budgets, they need to begin to quantify their justifications. The results are often far from perfect (more on that below), but they do demonstrate a willingness to change. Moreover, the presumed innumeracy of lawyers is proving to be its own myth. In Brazil, the Constitutional Court’s (Supremo Tribunal Federal, STF) promotion of better statistical systems for all the nation’s judiciaries has been taken up by many younger judges, who have shown themselves quick learners in more sophisticated techniques. Despite comments like that of a Peruvian judge, that he didn’t understand why donors were so fascinated with flow charts, judges in his own country, Costa Rica, and Bolivia have proved quite capable of using this technique to analyze procedural bottlenecks and unnecessary steps. In short, a little education and encouragement can often have a very positive effect.

It is true that as a whole, court leadership, sometimes, although not always, has been most recalcitrant in this area, combining a fear of quantification with a lack of management skills. As high court members, often the governing body for the court system as a whole, are usually older and are selected, at best, for their legal expertise, they may be understandably reluctant to learn new skills this late in the game, and fearful of demonstrating their shortcomings. This is not always true. Both the past and present presidents of Brazil’s STF have promoted the development of judicial statistical systems; other state court presidents in Brazil have done the same, and the current Costa Rican president has been a
proponent both of quantitative monitoring and of better management as a whole. However these individuals are in a minority, and while their peers sometimes tolerate the new approaches, most have done so without any enthusiasm or real conviction. As their general disdain also extends to their view of their own administrative systems and managers, the latter are also not selected for openness to new approaches, and when they have them, often are forced to keep them in check. In short, on the practical side, there are a series of vicious circles to be broken. Where this has occurred it may be because of visionary leadership, pressure from outside, or the provision of capabilities someone decides to use. The question for those interested in promoting change is how to identify the most useful entry point in any given situation – it is doubtful that there is one best rule for all.

The more difficult sources of resistance lie in another area – the ideological and political leanings of those within the courts and sometimes among the lawyers and others using them. Here two stand out in particular, the notion that measuring justice or judicial performance somehow conflicts with the essence of the subject, is an attempt to commodify a basic value, and the fear that any such move will also undercut the institutional independence of the courts. The conflict with the nature of the subject matter, justice, is on the one hand linked to judges’ definition of their own work as resolving individual conflicts and the idea that each of these conflicts merits exhaustive attention because it is unique. Here efforts to produce aggregate measures of productivity are viewed as a misapplication of techniques more suitable to the private sector. The Brazilians, despite or perhaps because of the strength of this notion among their own judges, have given considerable thought to the phenomenon often explaining it as the problem of transcending from an era where the demand for service was limited to one of the “massification” of justice. Much like any other modern professional facing a similar growth in demand, most judges can simply not envision working differently.

On the other hand the emphasis on quantity and on measurement has been linked in the region to the neo-liberal agenda, and the efforts to advance globalization. This increasingly does not extend to tracking of judges’ individual output, but does affect any efforts to use monitoring to other ends – for example to identify and eliminate bottlenecks created by multiple appeals (regarded by many as a guaranteed right), to prioritize cases, or to give more attention to some of them. The partial acceptance coincides with judges’ interest in using statistics to demonstrate they are working (and even to identifying the slackers), but does not cover analyses which might recommend altering the contents of their caseloads or otherwise modifying the uses to which courts are put. This more sophisticated rejection of the larger package is for the most part limited to the most advanced systems, unfortunately also those where judges are already overtaxed and where the crush of demand may only be resolved by adding more personnel or radically altering what they are expected to resolve. In the majority of the region’s countries, the opposition does not take this form, is more closely tied to lack of basic skills and installed capacity, and probably can be best addressed at that level. For both sets of countries, a better understanding of how performance monitoring can address their problems may help move the project ahead. Thus we turn next to those issues.

The Further Advantages of a Quantitative Approach to Performance: In answer to the argument that justice is different and measuring court outputs is indeed a misapplication of methods more appropriate elsewhere the reasons given in the introduction merit further consideration and expansion. Performance monitoring is more than a way of grading a justice system. It also is a means to improve output without fundamental changes, and for selecting the most important changes to promote.

Resource use: The strongest reason for performance monitoring is doubtless the need to utilize available resources as efficiently as possible. Related to this is the goal of justifying the provision of additional funding. In a world of finite resources and apparently infinite demand, courts like other
public and private institutions, cannot afford to ignore how they are using what they have and with what results. This is true whether they have 0.03 percent of the national budget or 30 percent (both unprecedented extremes it might be noted).

For judiciaries, like many other public entities, rationalization of resource use is complicated by the multiple objectives assigned to them. The efficient resolution of disputes is one thing, but to that can be added: the provision of access to all citizens, the use of the courts to establish a government presence in outlying regions, and the demands for higher priority attention to certain types of cases. In short, the judicial map, pioneered by France, does not provide a simple solution for where resources should be placed. (And it didn’t in France either, provoking still raging controversies about the single minded emphasis on leveling caseloads.) Still knowing where resources are placed, the population, and the raw demand served is a start. If courts or governments want to prioritize other values, most of these can also be tracked, and this kind of analysis may be the best defense against the common desire to “put a court in the congressman’s home constituency,” perhaps the prime cause of inequitable distribution of work.

Most courts already have the basic information required to make some of these determinations. That they have not used it is either because leaders don’t feel responsible for where resources go or because other priorities make resource rationalization irrelevant. Additionally, the notion of resource monitoring is so new that while courts may know which jurisdictions are over or underloaded, they have no idea what a reasonable workload is. This explains why the least taxed judiciaries in the region, those with average workloads far below any kind of international norm, still request additional funding for more judges and courthouses. To be fair, sometimes political considerations (the congressman who wants his hometown to have a judge) or the law itself can make rearrangements difficult. Colombia’s Constitution stipulates a judge for every municipality, and its secondary law makes it difficult to use the Brazilian practice of temporarily assigning judges to help clean up a large backlog. In São Paulo, Brazil, a former president of the state judiciary noted that the state judicial organic law required new judges be added when average workload exceeded 300 cases per judge. Courts may also be legally prohibited from redistributing specialized jurisdictions (e.g. making civil courts into criminal courts or vice versa), even when the number of courts and judges remains the same. However, if these other restrictions are to change, a first essential step will be demonstrating the negative results – and thus having adequate statistics on what courts are currently processing.

Identification of problematic performers: ironically the use of statistics to measure individual performance has been the least resisted application. True, a majority of the region’s judiciaries may not have accepted this development, but as they move toward more objective systems for evaluating judges, the quantity of individual output has become an obvious candidate. In many instances it is preferred to qualitative evaluations because of the latter’s susceptibility to a variety of abuses. The larger problem is what to do once problematic judges are identified – those who don’t keep up with

---

25 Just to give some idea, many judiciaries in Central America, the less developed countries of South America, and the less developed states of Mexico, Brazil, and Argentina, have an annual average of 100 to 300 new filings per first instance judge. As the vast majority is usually very simple cases, many of which are not further pursued by the plaintiffs or not contested by the defendants, it would be hard to consider this as an overwhelming amount of work.
26 This is because of legal stipulations that a case must be resolved by the judge in the jurisdiction to which it corresponds. Interviews in Colombia suggested this was commonly interpreted as prohibiting the temporary assignment of judges to help out the one who had fallen behind.
27 The point was largely rhetorical. São Paulo state judges currently receive nearly 2,000 new filings annually and there is no indication the President’s suggestion that their numbers be multiplied to reduce the average to 300 will ever be met.
caseloads, or whose pattern of judgments deviates suspiciously from the norm. When the provincial high court in Córdoba province (Argentina) first fired a judge for excessive delays in resolving cases, she protested, unsuccessfully, that this was a violation of judicial independence. However, few courts have attempted to do this, either because it is not legally permitted, or because it has not occurred to them that it might be necessary. In Brazil, judges cannot be dismissed for not keeping up with their caseloads, but high courts can prevent their promotion, a power recently reinforced via a constitutional amendment.\(^{28}\) Absent such legal provisions, those managing judicial careers (where they exist) may be limited as to what steps they can take in the face of markedly poor performance or may confront other impediments. Peru’s judicial council can sanction judges for this reason, but the Supreme Court has yet to provide them with the data they would need to act. Colombia’s judicial council manages the statistics and does keep track of levels of output, but has chosen to take a softer approach to encouraging increases. Here, the problem is the on-going drive, backed by many judges, to eliminate the council, and thus the council’s desire not to add fuel to the fire. Although Costa Rica’s common data base allows not only brute measures of productivity, but also calculation of average times to case resolution, interviews with members of the Supreme Court suggest they have been similarly reluctant to take an aggressive approach. Statistics on backlogs and delays may be considered in promotion decisions, but are otherwise not explicitly discussed.

Other types of poor performance are more difficult to establish statistically. However, the civil appellate court in the national district of Buenos Aires city has begun publishing the awards in tort cases. Its members believe that broad access to this information may be sufficient to discourage judges from succumbing to monetary enticements.\(^{29}\) Unfortunately, there is no way to validate that hypothesis. For other courts to use this method, they will have to develop systems allowing the separation of this type of detail, for subsequent publication as is or in aggregates for individual judges and types of cases. This type of exercise is best done through data mining, because of the potential for crossing variables and so controlling for a number of factors influencing results.

Following the patterns in Colombia, Brazil, and Argentina, it is likely that most judiciaries will move quickly to the use of statistics to track individual performance. The simultaneous emphasis on improving mechanisms for more objective evaluations makes it likely, and in the interest of most of the judges affected. The process should hardly stop here, but the concept is easier for management to understand, and also consistent with the efforts to prove they are collectively working hard.

**Identification of structural problem:** were variations in judicial productivity the only factors affecting performance, this would in fact be enough. However, there are clearly other contributors, some endogenous to the courts, some having external origins. It is already evident that statistical analysis will be important in identifying these additional causes although here the other sources of resistance will come into play. Using statistics in this fashion requires a change in mindset – from the caretaker to managerial outlook, and from the belief that the current system simply needs finetuning to acceptance of the possible need for broader change. Although many of their leaders don’t recognize it, Latin America’s courts, after over a decade of steadily increasing budgets, are facing the end of the fat years. As one court president has noted, the logic can no longer be how to do more with more, but how to do more with the same funding levels.\(^{30}\) Given the hard eye being turned on courts by many finance ministries, the challenge may even be how to do more with less. The caretaker mentality, focused only on how to distribute what one has and how to get more, will no longer suffice.

\(^{28}\) Amendment 45 of November 17, 2004.  
\(^{29}\) Highton et al, 2000.  
\(^{30}\) Luis Paulino Mora Mora, President of the Costa Rican Supreme Court.
Most courts could clearly increase the efficiency of the resources they now have. Courts which finance social clubs, large representational allowances, and even piano tuners, will have to start cutting back on extraneous expenses and using what they have more effectively. As the current President of the Brazilian STF noted in a recent speech to representatives of all the nation’s court systems, only 60 percent of the STF’s budget goes to finance its principal function (ação fim) while the rest is used for other purposes. Geographic distribution of services is thus only part of the general problem; the rest, as with the STF, refers to what budgets are being used to produce.

Nonetheless, the Brazilian courts have been extremely creative in increasing the efficiency of the existing system, automating what they can, grouping cases into categories and resolving them with single opinions (what the Brazilians call sentenças padronizadas or template opinions), and putting an emphasis on judicial output. In this they are aided by the fact that so much of the judicial caseload, especially in the higher and lowest level (small claims) courts, is extremely redundant. As the STF president has also noted, the reason he can participate in conferences, despite an annual responsibility for 10,000 opinions, is that the latter represent only a handful of different conflicts. Other court systems, not blessed, or cursed, by this Brazilian characteristic, can still use automation and improve organization at the courtroom level to increase capacity to resolve the cases now entering. If some have not already done more, one suspects one reason is that an effective rationalization of working procedures would leave many judges with an enormous amount of time on their hands. In any event, a good statistical system can help, not only by identifying the poor performers, but also by locating the areas where cases bog down for reasons other than judicial sloth. This is where the aforementioned statistically-based flow charts can also help out – by demonstrating where the bottlenecks appear and so helping locate areas for more attention. For example, if cases get stuck at the notification stage, than more notifiers could be added, or those already in place can be better trained, better supervised, or given improved ways of carrying out their tasks. The Costa Rican Court has introduced an automated system, scheduled for system wide replication, which uses maps and other computerized tools to increase the efficiency of the notification process. If certain types of cases seem to constitute an unusual proportion of the backlog, more attention, again with statistical aids, can be directed to determining why this is so. On the basis of initial analysis, Brazilians now know that tax collection cases have this status and are doing additional work to seek the reasons. A better statistical system or a unified, well constructed data base might have allowed this to be done more automatically – for example, following up on one hypothesis that the inactive cases are those for minimal amounts, cases which in effect state lawyers might be wasting their time pursuing.

Increases in efficiency within the current system will clearly help some courts provide more and better services. Court leaders who see themselves as managers and not just caretakers can push the process along. However, in more congested systems, like Brazil, the limits to an efficiency strategy will soon be encountered. Judges who are already handling over 3,000 new filings a year (over 10,000 for each member of the STF) are not going to be able to do much more. Moreover, there are indications that the cases most benefitting from the efficiency strategies are the simplest ones, those where batch processing is most easily accomplished. The accompanying fear is that more complex cases, what many regard as the judiciary’s most important work, may not be handled any better, going to the bottom of what the Brazilians call the “deus me livre” (god save me) pile. This is still another hypothesis current statistical systems do not allow us to explore adequately, as they do not permit sufficient distinction among types of cases (for example not letting Brazilians determine whether the neglected tax cases are really those

for inconsequential amounts.\textsuperscript{32}) The conclusions many observers are beginning to draw is that the traditional practice of courts’ simply accepting whatever cases come their way is no longer adequate to an era of “massified” justice and that once reasonable levels of efficiency are achieved, additional steps will be needed to differentiate the level of attention given to different types of cases and eventually to control demand itself. The first step, differential case management, is already being adopted by individual judges (and lies behind many of the automation innovations made by the Brazilian courts and others.) It is controversial, and resisted by some judges, but the fascination with automation, and the dramatic increases in numbers of case attended, have helped move it along.

The resistance is not entirely misplaced, as the results are indeed a kind of McJustice, and the cases usually subjected to this approach (mass individual claims all revolving around a single simple controversy – for example, the amount of adjustments to be made to state pensions as a result of currency devaluations or inflation) arguably receive a second best, and not very efficient solution. Although the resisters may believe that such cases deserve more individualized treatment, the real question is whether courts should be seeing them at all. That they continue to do so is partly due to the constraints of the legal framework (limited ability for high courts to set widely binding precedent, the fact that any such precedent does not necessarily bind non judicial actors, lack of sanctions for recalcitrant administrators) and partly to the interest of the major defendants (often government agencies or large private firms, including banks) in postponing payments on amounts legitimately due to their clients. These cases also experience, along with many others, what many consider an excessive and abusive rate of appeals. What we can glean from existing statistics suggests that the appeals rate, even for fairly simple matters, is unusually high in Latin America, adding to court congestion and representing a kind of use of justice to avoid justice. However, this is only an inferential conclusion, based on the number of filings at the first and higher instance, and an occasional ability to identify the types of cases most often seen on appeal. As many legal frameworks allow multiple appeals, we cannot for example tell whether the high rates are most closely linked to a lower number of first instance cases – most not appealed, but those that are, appealed many times – or are fairly uniform across the board. We also can’t tell whether some parties use the appeal system more frequently, and how this affects final outcomes. A better statistical system could not answer all these questions, but it could address some of them.\textsuperscript{33} Moreover, once other types of research (data mining or case file analysis) provide a better understanding of the problems, statistical systems can be used to monitor development over time and the impact of reform measures.

The call for more radical changes to the existing system, and especially those that might decongest courts by reducing certain types of uses, is not widely accepted, but it is implicit in some reform proposals not at all linked to statistical monitoring. Many recommended legal and organizational changes, while not couched in terms of restructuring demand, really aim at that end – limiting appeals, giving judges more discretion in accepting cases or motions, and attempting to channel some disputes into alternative forums. That it could be done more effectively with statistics is obvious; that it is not is a function of the poor development of statistics systems, their lack of wider acceptance as a judicial

\textsuperscript{32} It is not just judicial authorities who don’t know this. Commonly the head of government legal services does not know what the lawyers under his or her supervision are actually litigating – sometimes not even the number of cases, and virtually never the amounts at stake or what is being prioritized. No matter what the level of automation elsewhere in the sector, these offices usually lag far behind, lacking even the simple kinds of case management software used by many private legal firms. For a discussion drawn from the Brazilian experience, see World Bank, 2005b.

\textsuperscript{33} Another gap in the systems is clearly information on what parties do out of court. The suspicion that many cases not coming to judgment might be settled by the parties or that unrecorded payments might be made extrajudicially is an area where other types of investigation will be needed. See Kritzer 1986 for an interesting attempt to measure settlement rates in the U.S.
planning tool, and fears of proponents that their entrance into the debate might provoke other types of resistance. A more explicit strategy, one that would use statistics to identify reforms rather than to support proposals developed by other means, is a ways off, but is gradually gaining acceptance among those who see the problems as exceeding the efficiency-based remedies.

**Use to Develop and Monitor Reform Proposals:** as the foregoing suggests, performance statistics can be a useful tool in understanding sources of poor performance and in developing more effective remedies. They are also essential in monitoring the results of those that are adopted. Latin American judiciaries have undergone a series of reforms intended to improve their performance, starting in the 1980s with macro institutional changes (greater independence, higher budgets, new forms of appointments) and increasingly moving into more targeted remedies intended to address more specific performance problems. As has been documented elsewhere, complaints over their performance have not markedly declined, raising questions as to the efficacy of what has been attempted. While there are many explanations for this phenomenon, some not directly related to actual output, it is evident that the earlier reforms suffered from a lack of empirical base, and that the later ones have all too often been a product of what Brazilians term “achismo” (from “acho,” I think, and thus meaning a basis only in opinion). Efforts to track the impact of some recent changes more directly, indeed suggest they have not had the desired effect, in some cases compounded the problem. What is more noteworthy, however, is that the judiciaries introducing these efforts have not attempted to track them themselves, and that donors’ monitoring and evaluation of their own programs has also fallen short. In both cases, the lack of good performance statistics is the major immediate impediment, but that they do not exist is a function of the priorities of both national leaders and collaborating donors. There are indications that the pressure is on all parties to change these practices, courts because their ministries of finance have begun to ask questions about the returns on investments, and donors because those controlling their funds are also requiring results.

**The Path Ahead:** There is an old saying that we should be careful of what we wish for as we just might get it. It may apply to the desires of many students of Latin American judicial reforms to have a more empirical base for their work. Twenty years ago, when attention turned to the state of the sector, it was not uncommon to find court authorities did not know how many judges were seated, let alone how many cases they handled, or with what budget. It has taken nearly that long for the situation to change, but today, judicial statistics abound and are used by courts to demonstrate their output and justify requests for higher budgets, by reformers to back their proposals, and by academics to cast more light on sector situations. That much is all to the good, but as this brief article seeks to demonstrate, the underlying problem has hardly been resolved. The quality of the data used and the uses to which they are put leave much to be desired, and before we launch forward confident of what we know, more attention is due to how to rectify that situation.

This is obviously an area where greater exchange of information on past developments and new ventures is a high priority, within countries, within the region, and on a global level. Recent experience in Europe through the Council of Europe’s Commission for Judicial Efficiency may help interested Latin Americans and their external partners as they begin to design or redesign their own statistical systems. While the European effort worked with countries with fairly developed systems, all of them had shortcomings and the continuing discussions as to what to measure and how are probably as

---

34 Galindo (2003) for example compiles the results of recent opinion and experiential surveys on judicial operations and output. One notable finding is that post 1996, many countries, despite on-going reforms, have more negative public assessments.

35 See especially World Bank 2002 and 2003a on the impacts of Mexican reforms to the debt collection proceedings and World Bank 2003b on the Brazilian introduction of the “processo monitoreo” intended to speed up debt collection as well.
valuable to extra-regional parties as are the resulting compilations of comparative performance indicators. Experience within the region is also valuable, as there are systems and supporting software that might be easily adopted by late starters. However, overcoming technical shortcomings must go hand in hand with an emphasis on effective use and analysis. Whereas the former exchanges might be largely limited to technicians, the latter must include high level administrators and governing bodies. Judicial leaders will have to be convinced of the larger value of the effort if they are to give it their full support. Ideally, those regional judicial leaders with an interest in and understanding of the needs should reach out to their colleagues to encourage them to do more. Given the new budgetary restrictions on the courts, this is an obvious area for donors to finance, and one where they should share an interest in fomenting change in order to improve their own assistance efforts.

References


Brazil, Ministério de Justiça, (2004), Diagnóstio do Poder Judiciário.


Instituto Tecnológico Autónomo de México (ITAM) and Gaxiola Morailia and Asociados, S.C. (1999),
“La administración de justicia de las entidades federativas mexicanas a partir del caso de la cartera bancaria,” Mexico, October.


Biography

Linn Hammergren is a Senior Public Sector Management Specialist in the World Bank Latin America Regional Department, working in the areas of judicial reform, anti-corruption, and general governance issues. Before coming to the Bank she spent twelve years managing Administration of Justice Projects for USAID in Peru, El Salvador, and Costa Rica, headed USAID’s Regional Administration of Justice Project, and held a USAID Democracy Fellowship. She has a PhD in Political Science (University of Wisconsin) and taught Latin American and Comparative Politics at Vanderbilt University from 1974 to 1981. Her research and publications focus on local politics and decentralization, administrative reform, judicial politics and reform, judicial corruption, and the politics of foreign assistance. They include four manuals for practitioners on the lessons learned from Latin American judicial reform programs, *The Politics of Justice and Justice Reform in Latin American: Peru in Comparative Perspective* (Westview, 1998), studies done for the World Bank on “users and uses of justice” in Argentina, Brazil, and Mexico, and a monograph on judicial councils in Latin America.