The Price of Judicial Economy in the US

BRUCE GREEN*


Abstract

In the US, courts widely perceive that judicial scarcity is a common problem threatening the fair and timely resolution of disputes. Courts cite the attendant interest in judicial economy to justify interpreting the procedural and substantive law to reduce the judicial workload or accelerate the resolution of cases. But courts’ assumption that there are too few judges to handle the current caseload is hard to substantiate. First, it may not be possible to infer from excessive judicial backlogs or other perceived judicial deficiencies that a shortfall of judges is to blame. Second, even when one confidently perceives that a judicial backlog or other deficiency in a particular US court is attributable to a dearth of judges, one cannot fairly generalize from that example to other US courts and jurisdictions. And third, judgments about judicial deficiencies popularly attributed to the inadequacy of judicial resources may turn on contestable assumptions about judges and adjudication. Given these challenges to measuring the adequacy of judicial resources, one might be skeptical whether judicial economies are worth the costs they impose.

Key words

Judges; judiciaries; judicial scarcity; judicial economy

Resumen

En EEUU, los juzgados perciben que la escasez judicial es un problema frecuente que amenaza la resolución justa y puntual de disputas. Los juzgados mencionan el interés intrínseco a la economía judicial para justificar la interpretación del procedimiento y del derecho sustantivo para reducir la carga de trabajo de los jueces o acelerar la resolución de los casos. Pero es difícil justificar la creencia de los juzgados de que no hay suficientes jueces. Para empezar, quizá no sea posible deducir, de los excesivos casos pendientes de atender, que de ello se pueda culpar a una escasez judicial. En segundo lugar, no se puede generalizar a partir de un caso puntual a otros juzgados o jurisdicciones de EEUU. Y en tercer lugar, los juicios de valor sobre deficiencias judiciales pueden dar lugar a creencias contestables sobre los jueces y el arbitraje. Con estas dificultades para medir el grado de adecuación de los recursos judiciales, hay razones para ser escépticos sobre si la economía judicial merece la pena de los costes que supone.

* Prof. Green is the Louis Stein Chair at Fordham Law School, where he directs the Stein Center for Law and Ethics. He can be contacted at: Fordham Law School. 150 West 62nd Street, New York, NY 10023. USA. bgreen@fordham.edu.
Palabras clave
Jueces; poderes judiciales; escasez judicial; economía judicial
Table of contents

1. Introduction ........................................................................................... 793
2. Problems of Causality ............................................................................. 794
3. Problems of Generalization .................................................................... 795
4. Problems of Values .................................................................................. 798
   4.1. Judicial efficiencies, or denials of fair process? ................................. 798
   4.2. Too few judges, or too much litigation and too many rights? ............... 799
5. Judicial Economy and its Costs ............................................................... 800
   5.1. The judicial assumption of judicial scarcity ...................................... 800
   5.2. The influence of perceived judicial scarcity on judges’ conduct............ 801
   5.3. The cost of judicial economy ............................................................ 803
6. Conclusion .............................................................................................. 804
References .................................................................................................... 804
Case Law ....................................................................................................... 807
1. Introduction

The question of whether there are "too few judges" was posed by a 2016 workshop of the International Institute for the Sociology of Law, for which this article was written. Predictably, responses varied with the country studied. This article focuses on the United States, where hundreds of different courts are administered by the federal, state and local governments, and where judicial scarcity is a recurring theme. It explores the question of whether federal, state and local governments in the US unduly economize on judges – that is, whether they fail to establish and fund as many judgships as are needed to fairly and timely resolve all the disputes that are a legitimate subject of adjudication. Additionally, it considers the question of how US courts should respond to the perceived problem of judicial scarcity, including whether US judges should interpret the procedural and substantive law in ways that may reduce the judicial workload or accelerate the resolution of cases.

US judges, lawyers and public officials often complain that there are not enough judges, illustrate the problem based on their perceptions and experience, and urge that funding for judges be expanded. A recent civil rights lawsuit in New York provides one example. In May 2016, several individuals who had been arrested for misdemeanors in the Bronx, the northernmost county of New York City, filed a class action lawsuit in federal district court against representatives of New York State (Trowbridge v. Cuomo 2016). The lawsuit alleged that because of the immense backlog of cases in Bronx Criminal Court, criminal defendants charged with misdemeanors had excessively long waits before their cases could be tried. One of the former criminal defendants bringing the lawsuit claimed that he appeared in court 38 times over the course of 1,258 days before he had a trial, at which he was ultimately acquitted. Faced with such long delays, few defendants waited for a trial; most simply pled guilty, often to reduced charges, or accepted some other resolution offered by the prosecution. The conditions described in the complaint were not surprising and were not seriously contested. Indeed, the City Council held hearings on this very subject earlier in the year, and the Councilman who presided over those hearings flatly declared in response to the lawsuit, "There are just not enough judges" (McKinley Jr. and Weiser 2016). But this declaration reflected presupposition more than study and analysis.

This article does not attempt to study and resolve the question of "too few judges" but instead explores the difficulties in analyzing the question from the US perspective. The article also brings a skeptical approach to courts’ measures to achieve judicial economy in response to the popular perception of judicial scarcity. With respect to these inquiries, the article returns several times to the example of the Bronx Criminal Court backlog described above or to the criminal context generally.

The first three parts of the article identify and describe distinct challenges that one would encounter in attempting to ascertain whether there are too few US judges. Part 2 describes the problem of causality – that it may not be possible to infer from excessive judicial backlogs or other perceived judicial deficiencies that a shortfall of judges is to blame. Part 3 describes the problem of generalization – that even when one confidently perceives that a judicial backlog or other deficiency in a particular US court is attributable to a dearth of judges, one cannot fairly generalize from that example to other US courts and jurisdictions. Part 4 describes the problem of values -- that judgments about judicial deficiencies popularly attributed to the inadequacy of judicial resources may turn on contestable assumptions about judges and adjudication. For example, what one regards as a judge’s regrettable failure to give enough attention to a case because of an excessive caseload, another may regard as...
laudable judicial efficiency or innovation. Further, some might argue that judicial delay and inattention are not problems of judicial scarcity at all, but problems of excess: too much litigation or too many procedural or substantive rights.

As Part 5 describes, notwithstanding the difficulty of proving the inadequacy of judicial resources, US courts assume that judicial scarcity is an ongoing problem, and in part for this reason, they develop and employ measures to promote the institutional interest in “judicial economy.” This Part concludes by questioning whether, particularly given the difficulties of measuring the adequacy of judicial resources, judicial economies are worth the costs they impose.

2. Problems of Causality

What does it mean for a court to have “too few judges”? Ordinarily, it is taken to mean that the workload of the court is not being handled fairly and efficiently because there are not enough judges to handle the work. Determining whether there is judicial scarcity requires ascertaining, first, that the particular court is dysfunctional in these or other ways and, second, that the dysfunction is attributable to the scarcity of judges rather than some other cause. But it may be hard to tell whether a court is dysfunctional and, if so, whether judicial scarcity is the cause.

One can assess judicial scarcity only inferentially. For example, one may examine how long it takes judges on a court to resolve cases and infer that delay in resolving litigation is an undesirable symptom of judicial scarcity. US courts’ assumption that delay is unfair to litigants is reflected in their frequent invocation of the adage, attributed to British Prime Minister William Gladstone, that “justice delayed is justice denied.” (Geo. Walter Brewing Co. v. Henseleit 1911, Silica Prods. Liab. Litig. 2005).

Protracted litigation may have causes other than a scarcity of judges, however, with the result that adding judges will not necessarily reduce backlogs (Feeley 1978, 1983, Church 1985). For example, judges may fail to work efficiently. That was the implication several years ago when the New York City court administration, responding to accounts of backlogs in Bronx criminal cases, temporarily reassigned a judge renowned for her efficiency from Brooklyn to the Bronx. The delays in Bronx criminal cases were the subject of a four-part series in the New York Times in 2013 (Glaberson 2013a, 2013b, 2013c, 2013d). In her six months in the Bronx trial court, the judge reassigned from Brooklyn reportedly “churned through more than 500 cases, slashing by 40 percent the backlog of those over two years old” (Rivera 2013). Other accounts similarly suggesting that delay is sometimes attributable to judicial inefficiency not scarcity, include the account of a federal district judge in New York who reviewed 500 habeas (post-conviction) petitions that had been “languishing for years” on other judges’ dockets, and who then recommended how judges in the future could review habeas cases more efficiently (Perrotta 2003).

Further, it should be remembered that judges are not the only judicial resource. The problem of scarcity may lie elsewhere in the judiciary. Regardless of whether there are enough judges and whether they are sufficiently expeditious, backlogs and delays may result from an inadequacy of judicial resources aside from judges – e.g., court reporters, courtroom deputies, or ministerial judicial employees – or from court administrators’ inability to employ its other resources efficiently (Goode 2013). That is largely the theory of the above-noted civil rights complaint, which asserts that in Bronx misdemeanor cases the state “failed to provide adequate resources, failed to allocate and manage existing resources properly, and failed to train, supervise, and monitor the staff and judges adequately” (Trowbridge v. Cuomo 2016).

2 A year later, The New Yorker published a poignant story about a teenager, Kalief Browder, who was confined to prison for three years, part of that time in solitary confinement, while awaiting trial in the Bronx for allegedly stealing a backpack (Gonnerman 2014). In the end, Browder’s case was dismissed, but, unable to overcome the psychological debilitation of prison and readjust to life at home, he took his own life a year later (Gonnerman 2015).
The problems attributed to high caseloads may also have other causes unrelated to the judiciary and its resources. For example, delays in civil litigation may be caused by abuses of the pretrial discovery process by one or more lawyers in a lawsuit. In criminal cases, delays may be attributable to inadequate defense or prosecutorial resources, to defense strategies, or to prosecutorial policies (Feeley 1983, Glaberson 2013a, Rivera 2013). Either side or both may request continuances or adjournments, perceiving delay to be advantageous. Defense lawyers may prolong cases hoping that prosecution witnesses will die or disappear, or that their memories may dim, so that the prosecution is eventually compelled to offer a more favorable plea offer or dismiss the charges. Prosecutors may seek to make defendants return to court repeatedly, leaving work to appear in court, to pressure defendants to accept plea bargains rather than insist on a trial.

At the same time, the absence of judicial delays and backlogs does not necessarily signify the adequacy of judicial resources. Judicial scarcity may have other symptoms, as judges try to cope with too much work in inappropriate ways. Judges may take shortcuts to move their dockets, failing to give cases the time they deserve, resulting in poorly considered decisions or unfair pressures on parties to settle their cases or to forego procedural options. For example, the late Justice Scalia referred to the possibility that “because of time constraints [or] heavy caseloads,” some state judges would “fail to read the briefs but leave that to their assistants, whose recommendations they rarely reject” or that “only one judge on the state-court appellate panel reads the briefs and considers all the claims, and the others simply join the drafted order” (Johnson v. Williams 2013). But symptoms of judicial scarcity such as judicial inattention or carelessness may be difficult to discern, as may other symptoms, since judges may not acknowledge taking inappropriate procedural shortcuts. And, of course, judges’ propensity to take procedural shortcuts may not be attributable to excessive caseloads but may be the product of individual judges’ laziness or of judicial culture.

In sum, one significant challenge to resolving whether there are enough judges is that of causality: It is difficult to ascertain whether the judiciary is functioning poorly, but even when one concludes that the judiciary takes too long to resolve cases or gives cases too little attention, it may be impossible to say with confidence that judicial scarcity is the cause.

3. Problems of Generalization

At a level of generality, the question of whether there are too few US judges is probably unanswerable for a further reason: US courts differ by level of government, locale and function; their caseloads and resources vary; and how they utilize their resources may vary with local culture (Church 1985). There are no reliable guidelines or benchmarks for how many cases of any particular type can be handled by any particular type of judge in the US, and in any event, high caseloads do not necessarily equate with delay, since individual judges and court systems deal with caseloads in different ways (Feeley 1983). Even if one were persuaded that there are too few judges in one court, it does not necessarily follow that there are too few judges in another – not even if the other court’s work load and resources seem similar on the surface.

Returning to the example of the Bronx Criminal Court: If one were to conclude, as some have, that there are too few judges in this court, one might be tempted to assume that this illustrates a broader problem – for example, that there are too few judges in all Bronx courts, or in the Criminal Courts of all five New York City counties, or in all the courts of New York state, or throughout the US. For example, a senior public defender maintained that: “The Bronx is a window into the problems in the criminal justice system across the city” (Glaberson 2013a). But experience shows that one cannot reliably generalize from the problems of a single court at a single point in time. For example, the problem of trial delays in the Criminal Court, where
misdemeanor cases are heard, was reportedly greater in the Bronx, the state’s poorest county, than in New York City’s other five counties: “While trial delays are a problem in many New York courts, none are remotely as bad as in the Bronx – the poorest of New York City’s five boroughs and the poorest county in the state. Misdemeanor defendants... wait twice as long [to go to trial] as those in Manhattan, two-thirds longer than in Brooklyn, and half again as long as in Queens, according to the complaint” (The New York Times 2016). Further, in the Bronx in 2016, the problem was reportedly greater in Criminal Court, where misdemeanor cases were heard, than in the (oddly named) Supreme Court, the trial-level court hearing felony cases, because court administrators had recently shifted resources to the Supreme Court in response to complaints about delays in processing felony cases (The New York Times 2016).

Needless to say, if the problems of the Bronx Criminal Court are not representative of either Bronx courts in general or of the state’s Criminal Courts in general, one certainly cannot take them to be representative of courts in New York State or of courts throughout the United States. To answer whether there are too few US judges, one might have to study one court at a time. That would be a daunting task. There are many courts of many kinds throughout the US, where the national population exceeds 320 million, including both federal and state courts, which operate independently and are funded by different sources.

Federal trial courts generally adjudicate criminal cases involving federal crimes and civil cases where federal rights are at issue or where a dispute is between citizens of different states or countries. There are both federal trial-level courts of general jurisdiction – federal District Courts – in each of the 94 different federal districts, and specialized federal trial courts – e.g., federal bankruptcy courts and tax courts, the Court of Federal Claims, and the U.S. Court of International Trade. Then there are federal appeals courts of general jurisdiction in each of twelve federal circuits and a specialized federal appeals court, the Court of Appeals for the Federal Circuit, which hears patent and trademark appeals and appeals involving various other federal laws. There may be enough judges in one federal courthouse and too few in another, and the adequacy of resources in federal courts says little about conditions in state courts, where most adjudication takes place.

Every state has its own court system, so one cannot generalize based on conditions in any one state judiciary. Like the federal courts, state courts can be divided both geographically and based on subject matter. In New York City, the Criminal Courts hear misdemeanor cases, the Civil Courts hear civil cases involving amounts up to $25,000, and the Supreme Court (which, as noted, is a trial-level court in New York State) hears felony criminal cases and higher-value civil cases. At the trial level, there are also both specialized courts and specialized parts of courts – for example, Housing Court, Family Court, Surrogates Court, and Small Claims Court. Outside New York City, towns and villages in the state fund more than 1,200 different Justice Courts with both civil and criminal jurisdiction (Unified Court System n.d.). The Justice Courts are staffed by “justices” who need not be lawyers (Provine 1986, Glaberson 2006). The state has two levels of appellate courts as well.

In New York and any other given state, the adequacy of judicial resources may vary from court to court and from locality to locality. For example, it was observed eight years ago that in Indiana, “some courts suffer from overcrowded dockets, while other courts function only part-time due to their lighter caseload” (Baker 2008).

Additionally, in assessing whether there are “too few judges,” one must consider who counts as a “judge” in the US. In the US, just as there are ambiguities regarding what counts as a “court” or a “trial” (Kritzer 2013), there are ambiguities about who counts as a “judge”.

For example, in some jurisdiction, senior or retired judges are pressed into service to assist with otherwise excessive caseloads (Cherry 2017). If one takes senior
judges into account, one might conclude that there are enough judges in the jurisdiction. However, if one believes that cases should be decided by active judges, one would conclude that there are too few judges and regard the employment of senior judges as a procedural measure to cope with the shortfall.

Administrative Law Judges (ALJ), some employed part-time, adjudicate hundreds of thousands of disputes each year where a federal, state or local agency is a party. For example, ALJs in New York City’s Office of Administrative Trials and Hearings (n.d.) annually conduct more than 300,000 hearings involving environmental infractions and other violations of city law. New York State ALJs hear cases on matters such as tax appeals, alleged overpayments by the state to health care providers, and denials and revocations of professional licenses (Office of Administrative Hearings n.d., Tax Appeals and Tax Appeals Tribunal n.d., Department of Health 2017). Federal ALJs conduct hearings involving federal enforcement and benefits in agencies such as the Federal Communications Commission (n.d.), the U.S. Federal Labor Relations Authority (n.d.), and the Social Security Administration (Nolo n.d.). At times, complaints have been raised about backlogs in certain administrative settings because there were not enough ALJs in a particular agency (Costa 2014, Aaron and Chen 2016). If one counts ALJs as judges, it becomes even harder to say whether there are enough judges in the US to afford fair and timely resolutions of disputes.

Further, there are various personnel who are not called “judges” but who perform judicial or quasi-judicial functions. Referees and special masters are assigned to perform some of judges’ work. Many legal disputes are not heard in court but in arbitrations overseen by arbitrators who may or may not have legal training. One may view the employment of non-judges as symptomatic of the inadequacy of judicial resources, or one may conclude that referees, special masters and/or arbitrators are sufficiently judicial to be counted as judges when evaluating the adequacy of judicial resources.

A review of judges’ annual reports, writings and speeches, not-for-profit organizations’ studies, and news accounts, underscores the impossibility of generalizing from court to court in the US. The workload of one court does not necessarily provide insight into that of other courts.

As one might expect, many courts are said to have too few judges. For example, a 2013 study of Virginia state courts concluded that: “[T]he current judicial workload for circuit court, general district court, and juvenile and domestic relations district court exceeds the capacity of the existing complement of judges. Additional judges are needed to enable Virginia’s trial court judiciary to manage and resolve court business effectively and without delay” (Ostrom et al. 2013). Similarly, a not-for-profit organization promoting civic reform has asserted that there are not enough US federal judges to handle federal civil litigation because federal courts must give priority to criminal cases, and that “[w]hen there aren’t enough judges on the bench, many plaintiffs are forced into inadequate settlements, and small businesses are pressured to make unnecessary settlements to end the expenses and uncertainty of litigation” (People for the American Way 2015).

In other jurisdictions, however, the inadequacy of judicial resources is perceived to be a problem only for particular courts of the jurisdiction. For example, the 2014 annual report of New York’s Chief Judge focused particularly on the need for more Family Court judgeships, observing that this court’s caseload had doubled over the previous three decades and that “changes in the law and the effects of the economic recession” had further burdened its workload, but that “only a very few new Family Court judgeships have been established to meet this growing demand” (Lippman 2014). Elsewhere, population shifts within the state may lead to increased demand on courts of some localities and a corresponding reduction in the workloads of judges in other localities (Burdick 2017, Symons 2017). Consequently, in some jurisdictions, judges of certain courts may be perceived to have time on their hands. For example, several years, ago, a Connecticut newspaper described a small-town probate court
that opened only six hours a week because of lack of work, and recommended that some small-town probate courts be closed (Hartford Courant 2008).

In sum, the question of “too few judges” in the US raises contestable questions regarding who qualifies as a judge and what qualifies as a court. Even if one were to leave aside administrative proceedings and arbitrations and focus on state and federal tribunals, one cannot answer the question, “too few judges,” with a simple yes or no. The answer – insofar as one can ascertain it – will differ from court to court.

4. Problems of Values

In the end, an analysis of whether there are enough judges is likely to be value-laden for at least two reasons. First, as discussed in Section 4.1., there is room for disagreement whether the various processes developed to conserve and leverage judges’ time promote welcome efficiencies or are symptoms of judicial scarcity that constitute sacrifices in fairness and legitimacy. Second, as discussed in Section 4.2., insofar as one concludes that judges’ caseloads are too heavy for them to handle, one might question whether the solution should be more judges (or alternative adjudicative resources such as administrative or magistrate judges) or simply less litigation and fewer rights. In other words, the pressure on judges to handle high caseloads may be regarded as beneficial in encouraging courts and legislatures to develop innovative procedures that conserve resources in ways that promote fairness and reliable outcomes, or to cut back on substantive claims that are litigated unnecessarily. Alternatively, one may regard judicial labor-saving measures and substantive restrictions as unfair to parties. One’s perspective depends on one’s values and, consequently, the “too few judges” question may defy objective resolution.

4.1. Judicial efficiencies, or denials of fair process?

Modern US judiciaries and judicial processes are designed to reduce the number of judges needed by enabling them to use their time efficiently. This involves both the employment of judicial personnel who are not judges and procedural streamlining. It is sometimes unclear whether these are welcome efficiencies or regrettable symptoms of, or capitulations to, judicial scarcity. The answer turns on one’s assumptions about how judges and adjudication should function.

In many courthouses in the US, a substantial judicial apparatus supports the judges. Much of the work of judicial personnel is ministerial and therefore uncontroversial. It would be inefficient for judges to process filings and distribute orders and opinions rather than relying on support staff. Questions might be raised, however, when judges delegate work that relates to their core functions. For example, in some trial courts, judicial personnel interact with lawyers or parties on questions of procedure and substance, including by expressing views on legal questions and attempting to settle cases. Judges’ law clerks typically take a substantial hand in writing judges’ opinions. It is well known that many federal appellate judges, who generally have three law clerks, and the US Supreme Court Justices, who generally have four law clerks, often delegate substantial opinion-writing responsibility to these recent law graduates. Much less well known is that, as a recent study revealed, even ALJs (in the social security system, at least) rely on “decision writers” – typically either attorneys or “paralegal specialists” – to draft their opinions (Gelbach and Marcus 2016).

For the most part, the role of judicial personnel is taken for granted in the US. But one might see over-reliance on judicial support staff as a symptom of judicial scarcity no less troubling than delays in resolving cases. One might be concerned about the quality of the work performed by judicial personnel who have not attained the same level of experience as judges. Regardless of their capabilities, their substantive responsibility for aspects of judicial work may impair the quality of judges’ own work,
by distancing judges from the relevant law and facts. The role of non-judges may also undermine the legitimacy of judges’ decisions, which may appear to be heavily influenced by employees who have not attained their positions through a formal, democratic process.

A second set of judicial efficiencies involve procedural innovations to reduce judges’ work. To take two early examples, Workers’ Compensation and no-fault automobile insurance were designed in part to reduce tort litigation by affording injured individuals more timely and more certain compensation for certain injuries than they would attain through litigation (Hood et al. 2005). But one can argue whether, leaving aside limited judicial resources, these laws would best serve the public, or whether the plaintiffs would prefer to pursue higher damages awards in the civil tort system.

Other procedural innovations inspired by the interest in judicial economy may be laudable even apart from whether judicial caseloads are reduced. For example, class action procedures, which allow similarly situated plaintiffs to aggregate their claims, are justified in part by judicial economy: “The procedural purpose of the class action is judicial economy. The theory is that, if the interests of the active and absent plaintiffs are aligned closely enough, then simply resolving the interests of the active plaintiffs can save time, effort, and expense, thus also resolving the issues of the absent plaintiffs” (Sanders 2013). But the general effect of class actions is to expand access to justice by allowing the aggregation of claims that could not otherwise be pursued individually.

Ultimately, the question of “too few judges” calls for a judgment, in part, about whether procedural steps taken by courts and others in response to excessive caseloads promote or undermine justice. Arguably, in the name of judicial economy, courts and legislatures are moved to innovate and leverage judicial resources in ways that, for the most part, promote, or at least do not compromise, justice. But whether or not one considers procedural innovation to be benign depends on contestable values regarding how disputes should be adjudicated.

4.2. Too few judges, or too much litigation and too many rights?

One might perceive judges to be overburdened but nevertheless resist the conclusion that there are too few judges, perceiving that the fault lies elsewhere. In particular, critics periodically charge that the problem is that the popular culture is overly litigious or that the US affords too many judicially enforceable rights (Drum 2010). One might debate, for example, whether legislatures appropriately add to the judicial workload by allowing private enforcement of civil rights, environmental, consumer protection and other laws, rather than relegating these exclusively to administrative and regulatory enforcement by government agencies (Farhang 2010).

In the late 1980s and early 1990s, to give another example, representatives of the business community asserted that civil litigation – especially tort and products liability litigation – was making US businesses less competitive than their foreign counterparts (McCormack 1987, Huber 1988, Iacocca 1988, Olson 1991). The critique, based largely on anecdotes, was that the fear of litigation had led companies to avoid normal risks, and that the costs of litigating and of paying lawyers to avoid personal-injury lawsuits had to be passed on to consumers, creating disincentives to development and expenses that foreign companies do not have. For example, Chrysler’s president, Lee Iacocca, asserted: “The United States is by far the most litigious country in the world. We sue one another at the drop of a hat. About 90% of all the civil actions tried before juries in the whole world are tried in the United States, and we’re spending about $30 billion a year just suing each other” (Iacocca 1988). He blamed “greedy lawyers” and “greedy clients” equally (England 1990). The critique became the basis for efforts not to speed resolutions of civil disputes by increasing the number of judges, but rather to discourage civil lawsuits by, for example, eliminating or reducing punitive damages for corporate wrongdoing (McCormack 1987, Iacocca 1988).
The proposition that the US is the most litigious country has been challenged (Wollschlager 1998). Nonetheless, accepting the premise that the US civil legal system was overused and abused, the first President Bush appointed a Commission on Competitiveness to study the perceived problem and make recommendations. In 1991, the Commission reported back with an “Agenda for Civil Justice Reform in America.” Echoing the prevailing corporate critique, its conclusions included that there were too many lawsuits in the US, that litigation had undermined the competitiveness of US businesses, and that products liability lawsuits were particularly blameworthy. Among other things, the report recommended capping punitive damages and requiring more proof before awarding them, requiring losing parties to pay the prevailing parties’ costs, reforming civil discovery, and encouraging voluntary dispute resolution.

On the criminal side, one might make analogous arguments about the overtaxed judiciary. If criminal caseloads are too high, one might attribute the problem to over-criminalization: that legislatures establish too many criminal laws targeting petty wrongdoing that should be addressed civilly, and that prosecutors bring too many prosecutions under those and other laws that do not deserve to be the subject of criminal proceedings (Schwartz 1963, Alschuler 1979). Others might argue that criminal defendants are accorded too many rights and too much process, and that excessive judicial caseloads are among the byproducts (Amar and Cochran Jr. 1996). In either case, the solution may be, not to put more judges on the bench, but to change the criminal law or prosecutorial practices to reduce the number of prosecuted defendants, or to reduce the amount of work that judges must perform in criminal adjudication.3

Again, the “too few judges” question turns in part on contestable values. If one favors tort litigation and broad criminalization, one might conclude that there are too few judges to handle the resulting caseload. But one might argue in the alternative that the real problem is that the tort law is too permissive and that the criminal law is too sweeping.

5. Judicial Economy and its Costs

It would be impossible to prove that judicial scarcity is a near-universal problem in US courts, and it is unlikely that anyone will undertake the necessary empirical study to reach this conclusion. Nevertheless, as Section 5.1. discusses, US judges appear to assume that judicial scarcity is a universal condition, and, as Section 5.2. discusses, they justify their behavior based on this assumption. This leads to the question of whether the measures courts take in the name of judicial economy have too high a cost. Section 5.3. explores this question in the context of the resolution of criminal cases via plea bargaining, including in the context of Bronx Criminal Courts with which this Article began.

5.1. The judicial assumption of judicial scarcity

Although US courts have not put forth a well elaborated theory of judicial economy, contemporary courts often refer to the concept. This in itself does not necessarily mean that courts are concerned that judges have too little time to handle too many cases. One could view judicial economy as an important value even if judicial resources were abundant, on the theory that public resources should not be squandered, and therefore courts should always be efficient and mindful of their time. But besides invoking the interest in judicial economy, courts often refer to “scarce” or “limited” judicial resources (Robbins v. Milner Enterprises, Inc. 1960, Tidewater Oil Co. v. United States 1972, People v. McCray 1982, Painter v. Harvey 1988, Leung

3 Those arguing for decriminalization do not principally point to the benefits to the judiciary, however, but rather invoke fairness to defendants and a reduction of prosecutorial and defense burdens (Natapoff 2015).

Of course, even explicit references to limited resources do not necessarily denote a concern that there are too few judges. In some cases, the perceived problem may be too few courtroom deputies, law clerks or other personnel, or their unwillingness to work at hours when judges would otherwise hear cases; alternatively, the perceived problem may be inadequate funding to keep the courtrooms lit and heated. But judges are also explicit in acknowledging “heavy” or “excessive” judicial caseloads (Patsy v. Bd. of Regents 1982, Long v. State 2008, Pearson v. Callahan 2009, People v. Bankers Ins. Co. 2009). One can infer that judges are at least partly concerned about the burden on judges, not just the strain on other judicial resources.

5.2. The influence of perceived judicial scarcity on judges’ conduct

The institutional interest in judicial economy can influence US judges' work, or be used to justify their behavior, in various ways. First, judges develop and implement procedures, big and small, in order to conduct their work more quickly or to cut down on their work. On the big side, the concept of stare decisis, which lies at the heart of the US judicial process, has been termed “a maxim of judicial economy” (Cribben v. Ellis 1887). On the small side, judicial economy may be invoked to explain courtroom and deliberative processes. Judges themselves acknowledge processes adopted by overburdened courts to expedite cases – “procedures that ensure that individual cases will get less attention” (Reinhardt 1993). In situations where judges might once have allowed lawyers to argue legal questions at length, judges might shorten the amount of time for legal argument or decide a matter based solely on the written legal memoranda, without oral argument. Likewise, judges might rely on written submissions without conducting hearings on disputed factual questions. Or they may employ other procedural shortcuts to expedite proceedings. Not all economies may be visible; for example, judges may give less time and thought to a question than may be warranted in order to reach a fully-considered resolution. One academic critic of state criminal proceedings observed almost two decades ago that: “The unfortunate reality in the criminal courts today is that the judge's primary concern is not with the adequacy of the representation provided by overburdened defenders, but rather to convey the message with which the criminal justice system is most concerned: move the calendar and process cases as rapidly as possible” (Klein 1988).

The question of whether judges are giving matters enough time, either in general or in particular cases, calls for a value judgment resting on empirical assumptions that may be unprovable – e.g., how much time does a judge spend studying a legal

---

4 See, e.g., Tidewater Oil Co. v. United States (1972): “Given the potential waste of limited judicial resources -- those either of this Court or of the district court -- associated with each choice, neither can be considered attractive.”; Painter v. Harvey (1988): “Holding counterclaims compulsory avoids the burden of multiple trials with their corresponding duplication of evidence and their drain on limited judicial resources.”; Robbins v. Milner Enterprises, Inc. (1960): “To require a trial which will (or may) end in the very same judgment of dismissal which we here reverse may appear to be an unnecessary and wasteful expenditure of precious limited judicial resource.”; Leung v. Verdugo Hills Hospital (2012): The public policy favoring pretrial settlements “lessen[s] the burden that litigation imposes on scarce judicial resources”; People v. McCray (1982): “In this era of over-crowded court calendars and scarce judicial resources, we should not alter the trial stage in such a way as to necessitate or encourage unwarranted additional lengthy delays.”
question, and would the case be better resolved if more process were afforded? When cases are moving expeditiously, it may be hard to say whether judges are judging badly – shortchanging litigants in response to caseload pressures – or judging well, fairly and efficiently.

Some procedural efficiencies may have no significant implications for how cases are ultimately resolved. For example, where a case may be decided on multiple grounds, no one can complain if the judge decides the case on only one ground, and addresses the question that requires the least effort (Lambrix v. Singletary 1997, Pearson v. Callahan 2009). In other situations, however, judicial efficiency arguably comes at the expense of fairness to a party. For example, it is efficient to try a defendant charged with multiple related crimes in a single trial rather than separate trials, or to try multiple defendants charged with related offenses together rather than separately (Mutual Life Ins. Co. v. Hillmon 1892). But if parties claim that they will be prejudiced when charges or parties are combined, a judge must decide how much significance to ascribe to the institutional interest in judicial economy: does the balance tip in favor of efficiency or fairness (United States v. Hawkins 2009)?

Besides structuring their own work to promote judicial economy, courts sometimes interpret procedural statutes and rules in light of the interest in preserving judicial resources. This seems fair insofar as judicial procedures are designed with judicial economy in mind. All things being equal, courts might read the law in a manner that saves them time. But the principle is sometimes invoked regarding questions with considerable significance for the outcome of litigation. For example, judicial economy figures into decisions about whether to close the courthouse door to certain claims or litigants – e.g., decisions whether a court will take jurisdiction over a case or whether a particular litigant has legal "standing" to bring a case (United Mine Workers v. Gibbs 1966, United Food & Commercial Workers Union Local 751 v. Brown Group, Inc. 1996).

Further, courts invoke judicial economy in cases that are not centrally about judicial process. For example, in deciding on the scope of substantive rights, courts sometimes anticipate the burden on courts of enforcing the rights. Virtually any case about whether a party has a right that would be vindicated in court implicates the judicial workload. Judges may define rights in a way that minimizes the judicial burden – e.g., by adopting a standard that is easy to implement because it necessitates limited fact finding (United States v Drummond 1990, Giordano v. Mkt. Am., Inc. 2010). Or judges may decline to recognize rights altogether out of concern for the judicial burden. For example, four dissenting Justices in a 2009 case predicted that trial courts would be flooded with recusal motions because of the Court’s ruling, and therefore opposed recognizing that due process can be violated when a judge may be biased by a party’s large campaign contributions (Caperton v. A.T. Massey Coal Co. 2009).

One might question whether judicial economy is given as much weight as opinions suggest. It seems likely that some judges invoke judicial economy primarily for rhetorical effect to justify decisions they favor for other reasons and that may not in fact reduce courts’ workload (Green 2010). Even judges who seriously consider judicial economy probably do not regard it as the paramount interest in a given case, though it may tip the balance. One might also question how much weight this interest deserves. As suggested above in Parts 2, 3 and 4, the assumption that courts are overtaxed may be unfounded. Further, courts may value judicial economy even when they have a manageable workload simply because they would prefer to work less (Feeley 1978).

---

The question of whether a case is well resolved in itself implicates various contestable premises. For example, one might contest whether a case is well resolved if a court reaches an appropriate decision but does not adequately explain the decision, or if a participant or the public perceives that the proceeding was unfair because a party did not have a sufficient opportunity to be heard.
5.3. The cost of judicial economy

Even if courts are right to assume that there are too few judges, and that there is a need to use judicial resources economically, one might consider whether some efficiencies involve too great a sacrifice. Some imaginable judicial economies would undoubtedly involve an intolerable trade-off between efficiency and fairness. Courts could race through their dockets by deciding cases on a coin flip, and they could cut down on litigation by denying rights arbitrarily, but either approach would be an abdication of their responsibility to decide questions of law and fact fairly. The price of judicial economy, even assuming there are too few judges, would be just too high in anyone’s estimation. One question to consider is whether, in some respects, that point has already been reached under the adjudicative process now employed in the US.

The US plea bargaining process is a dramatic context in which to explore this question. Almost a half century ago, the US Supreme Court explained: “For a defendant who sees slight possibility of acquittal, the advantages of pleading guilty and limiting the probable penalty are obvious – his exposure is reduced, the correctional processes can begin immediately, and the practical burdens of a trial are eliminated. For the State there are also advantages – the more promptly imposed punishment after an admission of guilt may more effectively attain the objectives of punishment; and with the avoidance of trial, scarce judicial and prosecutorial resources are conserved for those cases in which there is a substantial issue of the defendant’s guilt or in which there is substantial doubt that the State can sustain its burden of proof” (Brady v. United States 1970). A year later, in a leading case on plea bargaining, the Court acknowledged that: “If every criminal charge were subjected to a full-scale trial, the States and the Federal Government would need to multiply by many times the number of judges and court facilities” (Santobello v. New York 1971). A lower federal court similarly observed a few years later: “[P]lea bargaining is sanctioned because without it the system of criminal justice could not function effectively (…). Even given the institution of plea bargaining, this country confronts severe backlogs, scarce judicial resources, and overworked attorneys” (United States v. DeMarco 1975).

More recently, the US Supreme Court has recognized that plea bargaining is “not some adjunct to the criminal justice system; it is the criminal justice system” (Scott and Stuntz 1992, quoted in Missouri v. Frye 2012). Criminal defendants are offered leniency in exchange for a guilty plea that avoids the state’s obligation to prove the defendant’s guilt at trial. The process serves a number of public interests, including conserving prosecutorial resources, achieving finality, and sparing witnesses the anxiety of having to testify, but judicial economy is certainly among them. As one commentator noted: “Although plea bargaining has its critics, it is nevertheless viewed as an effective means of achieving judicial economy through the avoidance of full-scale trials” (McClanahan 2013). And as another observed: “Plea bargains are essentially built-in leniency, the reward for judicial economy” (Strutin 2013). Cognizant of courts’ institutional interest in judicial economy, judges contribute to the process both by developing law that legitimizes or facilitates plea bargaining and by exercising sentencing discretion in a manner that encourages guilty pleas – that is, by sentencing defendants more leniently following a guilty plea (McCoy and Mirra 1980, Bowers 2008).⁶

Debates over the plea bargaining process illustrate how, as discussed above, pronouncements about the adequacy of judicial resources may turn on contestable value judgments. Some observers might charge that plea bargaining denies defendants a fair process – both by inducing most defendants to give up their procedural rights in exchange for leniency and by subjecting those few who elect to

⁶ Some maintain that the disparities encourage guilty pleas by innocent no less than guilty defendants (Bowers 2008).
go to trial to harsh sentences (“trial penalties”) if convicted, to induce future defendants to plead guilty. For example, a federal trial judge, disturbed by the onerous sentence required after a defendant turned down a plea bargain and was convicted following a trial, recently observed: “[A] serious question posed by the facts of this case – and perhaps the elephant in the room – is the fairness of the fifty-two-year mandatory minimum sentence that Defendant is facing because he opted to forego a ten-year plea offer and go to trial... [I]t is difficult to see how a forty-two-year trial penalty for a twenty-seven-year-old defendant who is not alleged to have fired a weapon, much less killed anyone, could be ‘the right thing’ – no matter how that vague term is defined” (United States v. Washington 2014).

Others might see plea bargaining as a legitimate way to enable courts to conserve limited judicial resources by adopting or facilitating worthwhile procedural economies. The differing views were illustrated by the reactions to the work of the Brooklyn judge who helped clear up the backlog of Bronx felony cases. On one hand, a Bronx supervising prosecutor perceived that the judge was simply more adept than her contemporaries at encouraging both sides to reach a fair resolution expeditiously: The prosecutor expressed the view that the judge’s “straightforward discussions with defendants, detailing their options and what a jury might do, have brought many to face reality.” On the other hand, a Bronx judge, privately complaining that dangerous defendants had been treated too leniently, reportedly observed that the Brooklyn judge “was presiding over a ‘fire sale’” (Rivera 2013).

To be sure, plea bargaining can be explained in various ways. But particularly insofar as it is justified as a measure to promote judicial economy, one might question whether it is worth the costs, such as the sacrifice of procedural due process. A similar question might be asked about other measures adopted in the name of judicial economy. Courts take for granted that judicial economy is important because judicial resources are inadequate and judges in particular are in too short supply, but their assumption is unproven and would be hard to prove.

6. Conclusion

The question of “too few judges” raises various complexities in the US setting, ultimately defying easy resolution. Judges’ reports suggest that judicial scarcity is not a universal problem but that the adequacy of judicial resources, and of the supply of judges in particular, varies from court to court and jurisdiction to jurisdiction. In courts where judges or others assume that excessive caseloads and delays or other perceived problems result because there are too few judges, they may be wrong about the causes. The addition of judges may not solve the problems. Further, the premise that all the cases filed deserve to be in court and deserve to be processed in particular ways reflects value judgments that often may be contested. But despite cause for skepticism, the concept of judicial scarcity, and the related concept of judicial economy, are recurring themes in judicial opinions.

References


**Case Law**

*Aero Spark Plug Co. v. B.G. Corp.*, 1942. 130 F.2d 290.


Cribben v. Ellis, 1887. 34 N.W. 154.
Samara v. United States, 1942. 129 F.2d 594.
Union P.R. Co. v. Bean, 1941. 119 P.2d 575.