Delay by the Parties and Delay by the Courts

Hans Zeisel
DELAY BY THE PARTIES AND DELAY
BY THE COURTS

HANS ZEISEL *

I

The one good thing to be said for court congestion is that it has spawned a number of careful studies on the operation of our courts and has, indeed, marked the beginning of scientific inquiry in the area of judicial administration. We have had a series of stimulating investigations by the Columbia University Project for Effective Justice, dealing with, among others, such proposed delay remedies as the master system in Massachusetts and compulsory arbitration in Pennsylvania. Then came a modest but elegant study of the Pittsburgh courts by Milton D. Green, under the auspices of the Institute for Judicial Administration of New York University. And now we have before us a major contribution to the field from the University of Pennsylvania Law School, setting out to investigate with care and circumspection the extent and causes of delay in the major civil courts of Pennsylvania.¹

The study is built around a framework of nine stages, which together form the full course through which a litigated case may pass, beginning with the period that elapses between the cause of action and plaintiff's first visit to counsel and ending with the final disposition after appeal. From a random sample of cases on the trial list, a count was made of how much time elapses, on the average, in each of these stages. A similar count was made for a sample of "long cases" that had stayed in the court far beyond the average time. Afterwards, counsel on both sides were interviewed to establish the reasons for the various delays. These

* Professor of Law and Sociology, University of Chicago Law School.

two approaches are supplemented by a modest effort to ascertain how much time judges actually spend on the bench trying cases.

Thus, this study concentrates on ascertaining by how much civil cases in the Pennsylvania courts are delayed; at what stages the delay occurs; and, finally, what caused these delays. This was done for seven Pennsylvania counties, with special emphasis on Philadelphia and Pittsburgh, the two major urban centers.

The following table, which records the delay for jury trials in the Philadelphia courts, shows the type of data that were obtained from the first research approach, the counting operation. Each stage is defined by the two listed boundary events, so that stage one covers the time between the cause of action and plaintiff's first visit with counsel. These time intervals are measured for two types of cases: a random sample of all disposed cases, and a subsample of the cases that stayed in the court for an abnormally long time. The table is to be read as follows: on the average, it took about two months (the figures are rounded off) for a random case to cover this first period, and so forth.

### Delay of Jury Trials in the Philadelphia Courts*

<table>
<thead>
<tr>
<th>Stage and Boundary Event</th>
<th>Random Sample of Cases</th>
<th>&quot;Long Cases&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cause of action</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Plaintiff's first visit to counsel</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Filing of complaint or summons</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Case is at issue</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Case first ordered for trial</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Case first on trial list</td>
<td>8</td>
<td>45</td>
</tr>
<tr>
<td>Trial begins</td>
<td>0**</td>
<td>0**</td>
</tr>
<tr>
<td>Trial ends</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Posttrial motions are decided</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Final disposition after appeal</td>
<td>38 months</td>
<td>89 months</td>
</tr>
</tbody>
</table>

*This table is a simplified adaption of chart I, on p. 311 of the study.

**Zero months, because this stage lasts only 2 to 3 days, on the average.
This table permits us to see some of the difficulties that confronted this inquiry. Delay in stages one through three came, of course, almost entirely from the client and his counsel. Delay in stages four through six is caused to some extent by the court (in many cases exclusively by the court); but in many cases, counsel contributes to the delay in these stages. Stage seven is of little consequence, because cases tried before a jury are almost never interrupted once the trial has begun. And the posttrial stages eight and nine apply of course only to a minority of cases. As to the size of these delays, the concentration is clearly in stages five and six—an average of eighteen months for the random sample of cases, and of fifty-five months for the long cases. These are the stages for which primary responsibility rests with the courts.2

Such measurements, together with the interviews with counsel in these cases, provide invaluable data on the client’s own tardiness to visit a lawyer; on the plaintiff lawyer’s delaying the trial date offered to him by the court; on courtesy consents to the defense lawyer’s motion to continue a case; and, of course, on the court’s inability to offer an early trial. All this is of great interest, and at some points, it would seem that serious questions of professional ethics might be involved—for instance, when it is reported that:3

. . . in close to 50% of the long [lasting] motor vehicle . . . cases it was the negligence of the plaintiff’s attorney in failing to answer a call of the trial list or in failing to re-order for trial promptly . . . which was the cause of substantial delay.

Unfortunately, the study does not always make it clear what moved counsel to delay. To be sure, counsel may not answer a call out of sheer negligence, but it may also be an intentional move in the interest of his client because, for instance, the evidence is not fully ready for trial, or because counsel may prefer that settlement negotiations be continued. But even if counsel should have so many cases on hand that he is forced to postpone some, such delay might well be construed to be in the interest of his client if the client prefers a delayed trial with this particular counsel to an earlier trial with a counsel who would be less busy but also less desirable. Thus, the study does not always permit us to distinguish between delay that is in the client’s interest, delay that is in the counsel’s interest, delay that is caused by the counsel’s courtesy to opposing counsel, and delay that is simply due to negligence.

2 It would seem, incidentally, that the delay figures for the average cases might be somewhat on the low side, because the random sample did not include (because they were not yet disposed of) some of the long cases that should have been part of it.

3 P. 317.
The study also has some data that are relevant to the system delay. It is reported, for instance, that the Philadelphia court keeps a separate list for jury cases and for nonjury cases; eighty-five per cent of all cases are on the former, fifteen per cent are on the latter. But since twenty-seven per cent of the jury cases reach trial as against twice as many—fifty-four per cent—of the nonjury cases, the relationship of jury to nonjury trials is like \(0.27 \times 0.85 = 0.23\) to \(0.54 \times 0.15 = 0.08\)—or, roughly, like 3 to 1. But the rules of the court require that only six months of the ten-months court year be devoted to the trying of jury cases and that four months be devoted to the trying of nonjury cases,\(^4\) a ratio of 3 to 2. Accordingly, it is not surprising that while jury cases keep piling up, the “time allotted for non-jury trials was not always needed . . . some . . . went unused altogether. . . . [T]he lists were so current that judges often did not have enough trial work during non-jury terms.”\(^5\) An actual count of the average number of trial hours on nonjury days yielded 2 hours, 17 minutes, including “all the five and ten-minute recesses, of which there were a number.”\(^6\) These averages cover only the days on which a judge actually sat in court; they do not cover the days on which a judge should have tried a case, but did not. And on the jury calendar, in six per cent of the cases, “the judge refuses to hear a case because it is complex and the end of the week is approaching.”\(^7\)

Such data make it somewhat difficult to share the local (in this case, not the authors’) enthusiasm for the Pennsylvania system of compulsory arbitration for small claims, where Ersatz-judges work for substandard fees and supply their overhead free of charge.\(^8\)

The Levin-Woolley study also has some interesting figures on local peculiarities of the various systems. For instance, the percentages of jury cases that are disposed of by adjudication (that is, not by settlement) differ as follows between the counties:\(^9\)

\(^4\) P. 277.
\(^5\) Pp. 377, 381.
\(^6\) P. 85.
\(^7\) P. 298.
\(^8\) The book calls it “The Pennsylvania Contribution” to the cures of court delay.
\(^9\) This is a summary of chart II, p. 67.
Percentage of Jury Cases that Go to Verdict

<table>
<thead>
<tr>
<th></th>
<th>Philadelphia</th>
<th>Allegheny (Pittsburgh)</th>
<th>Five Other Counties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>19%</td>
<td>30%</td>
<td>43%</td>
</tr>
</tbody>
</table>

These are challenging figures, because if one knew the reasons why the settlement ratio in Philadelphia is so high, one might be able to reduce the trial load in the other counties. Unfortunately, we do not learn these reasons; hence, we are frustrated in drawing the proper conclusions.

The study thus provides an illuminating analysis of the delay of the individual case on its way from the cause of action to its final disposition. These are highly relevant data from the point of view of the individual litigant. Yet, the main target of the investigation is the delay in the system. And, indeed, at first glance, it would seem that the system is but the aggregate sum total of the individual delays. Hence, an analysis of these individual delays should be tantamount to an analysis of the system. Curiously enough, this is not the case.

II

It is true that at any one point of time, the sum total of the individual delays in a court constitutes (and is, therefore, equal to) the delay of the system as a whole. But, as we shall presently show, the individual delays are to a surprising degree independent of the over-all delay of the system. Increasing the delay of an individual case does not necessarily increase the delay of the system; nor does a reduction of the delay of the individual case necessarily reduce the delay of the system. The reason why the system remains unaffected by such changes is that the system, as a rule, automatically compensates any increase of delay in one case with the advancement of other cases, and vice versa, thus keeping its over-all delay unchanged.

To appreciate this, we may visualize all the reasons for delay in the individual case falling into three broad categories: when the parties are ready, and the court keeps them waiting; when the court is ready, and the parties are not; and, when neither the parties nor the court are ready, meaning that even if the parties were ready, the court could not oblige them.

Let us now relate these three types of delay to a somewhat simplified model of a congested court system. We will assume that this court system has the following properties:

1. In each pending case, both sides declare to be ready for trial exactly six months after suit was filed.

---

15 Journal of Legal Ed. No.1–3
2. From that point on, each case has to wait exactly three months before the case is reached for trial by the court.

3. The trial of each case lasts exactly two court days, and (to simplify our model still further) suits are also being filed at intervals of exactly two court days apart.

4. The court consists of one judge who tries cases to capacity—that is, ten cases, or twenty trial days a month. In such a court system, each case will reach the trial list exactly six months after suit was filed, and the trial list itself will always consist of thirty cases, a three-months’ supply for the judge. Every two days he will remove one case by trial, and a new one will be added at the bottom of the list. This case, like all others before it, will reach trial after three-months’ waiting time. Adding the six months’ time for preparation, each case will be tried exactly nine months after suit was filed. Such a court system would be characterized as having no delay caused by the parties and three months’ delay caused by the court.

Let us now introduce our first modification of the system. Suppose that, for whatever reasons, some lawyers begin to need more time for the preparation of their cases. Let us assume that beginning on a certain date, every other case will require seven instead of six months’ preparation. And let us now see what will happen to the system. Up to that point, every case, according to our simplified assumptions, will have reached the trial stage exactly \( (6 + 3) = 9 \) months after the day suit was filed, and in the order in which it was filed. After that date, every other suit will reach the ready-for-trial stage only after seven months. As a result, for the duration of one month, only the six-months’ cases will line up on the trial calendar; but not at the rate of ten per month as before, but at the rate of five per month. Hence, after one month, the trial list will contain not thirty cases (a three months’ supply), but only twenty-five cases—that is, a two-and-one-half months’ supply. Ten cases will have been tried during that month, but only five replaced; and \( 30 - 10 + 5 = 25 \). During the following month, again, five six-months cases will be filed, but also five seven-months cases, delayed from the previous month will have matured in the meantime. Hence, the waiting list will from now on remain \( (25 - 10 + 10) = 25 \) twenty-five cases long.

As a result, once a case reaches the trial list, its waiting time is reduced to two-and-one-half months. But let us now see how this affects the total time elapsed between filing of suit and trial. For half of all cases (the six-months cases), the waiting time will be \( (6 + 2 \frac{1}{2}) = 8 \frac{1}{2} \) months; and for the other half (the seven-months cases), the waiting

10 All the numbers are arbitrary and irrelevant for the point we are about to make.
time will be \((7 + 2\frac{1}{2} =) 9\frac{1}{2}\) months. But note that the average over-all waiting time for all cases has remained the same, because \((\frac{1}{2} \times 8\frac{1}{2}) + (\frac{1}{2} \times 9\frac{1}{2}) = 9\) months.

Thus, increasing the waiting time of half of these suits had the automatic effect of correspondingly shortening the waiting time of the other half of the cases. This will hold true, as we were able to show in Delay in the Court,\(^{11}\) for any set of figures, not only the simplified ones in our example. A number of important conclusions derive from this:

1. No increase in delay of individual cases can have any effect on the average delay—that is, on the system as a whole—as long as the court continues trying cases at its normal rate, and the over-all number of cases requiring trial does not change.

2. Individual delay becomes a concern of the system only when a trial continuance is effected so late that no substitute case can be obtained and a gap ensues in the trial schedule. Such continuances should be disallowed with utmost rigor.

3. But except for this situation, the system can only benefit by delaying the trial of an individual case. As long as a case is not tried, there is always the possibility that it will be settled without trial, thereby reducing the system's over-all trial load.

4. Obviously this is only a minor hope for the system; hence, it must not interfere with the legitimate interests of the litigants for continuance or advancement. If a motion for continuance is made jointly by both litigants, the court's only possible concern is with the protection of the litigant against a courtesy consent that is really not in his true interest. If such a motion is opposed by the other side, the court must weigh these interests; also the defendant's right to have his case disposed of at the proper or earliest possible time, whichever the case may be.

To round out these considerations, we shall now reverse our modification of our court model and assume that some lawyers were able to speed up their preparations. As a result, every other case reaches the waiting list now a month earlier—that is, in five instead of six months. It requires little mathematics to see that the result will be that this half of the cases will reach trial half a month earlier, in \((5 + 3\frac{1}{2} =) 8\frac{1}{2}\) months. But as a result of this advancement, the other half of the cases will now be tried half a month later, in \((6 + 3\frac{1}{2} =) 9\frac{1}{2}\) months. Again, the average waiting time remains unaffected, at nine months. From this, still another conclusion follows:

\(^{11}\) ZEISEL, KALYEN & BUCHHOLZ, op. cit. supra note 1, at 53.
\(^{12}\) Id. at 193.
5. Just as delay of individual cases cannot affect the delay of the system, advancement of individual cases cannot affect the system, but benefits only these litigants. The average delay will remain the same, since the cases whose preparation time did not change must now join a longer waiting line.\textsuperscript{13}

To achieve full clarity in the relationship between individual delay and system delay and derive an important sixth conclusion, we shall now modify our simplified court model in a slightly different way. Suppose the preparation time of all cases, through greater solicitousness of the lawyers, could be reduced from six months to three months. How would this affect the over-all waiting time for each case? Again, the answer is: It would not affect it at all. All that would happen is that the trial (waiting) list would be formed at the end of the third month instead of at the end of the sixth month, and would, therefore, become twice as long. The waiting time in the first stage would, indeed, have been reduced to three months, but the waiting time in the second stage would automatically increase to six months; the over-all waiting time would be, as before, nine months, because $3 + 6$ is as much as $6 + 3$.

We may think of a supermarket in which there is a five-minute waiting line at the check-out counter and in which every customer shops for ten minutes before she joins the line. If a speed-up in the shopping method were to curtail shopping time to five minutes for all customers, they would reach the waiting line five minutes earlier but they would then have to wait ten minutes (instead of five) in line. Since $10 + 5$ is as much as $5 + 10$, the over-all waiting time would remain fifteen minutes. All, of course, under the proviso that there is no change in the number of check-out counters and the time it takes to check one customer out.

Applying our supermarket example to the court system, we then derive:

6. If all litigants shorten their preparation time prior to the bottleneck by an equal amount of time (however great this time amount is), the over-all waiting time for each case will thereby not be affected. Or, put differently: No shortening of individual delay time in stages prior to the bottleneck can have any effect on the system, because it would only increase correspondingly the waiting time at some other stage prior to the bottleneck.\textsuperscript{14}

\textsuperscript{13} The important note played by the average (mean) delay in the measurement of court delay makes it clear why such measurements ought always to be given in terms of the mean delay, not the median delay. Many studies give often only the latter, because it is easier to compute; but it is only with respect to the mean delay that the system remains invariant against individual delay.

\textsuperscript{14} Hence, remarks by Levin-Woolley that “an efficient judicial system can eliminate stage 4 entirely” (p. 262), while undoubtedly true, is relevant only for a court system in which this is the only delay; obviously it is not a characteristic of the courts under investigation.
This is, indeed, the final insight into the relevance of individual delays. Aside from being important to the individual litigant, they are important to the system only if the system is up-to-date or becomes up-to-date because the individual delays have removed the waiting line.

In a delayed system, individual delay can affect the system only in one quite specific way. This follows from the basic insight that must govern all inquiries into court congestion—namely, that the over-all delay of the system can be affected (reduced or increased) only in one of two ways: by affecting the over-all trial load, either by changing the number of cases requiring trial or the time it takes to try a case; or by affecting the over-all trial capacity of the court, either by changing the trial hours per judge or the number of trial judges. Therefore, only when an individual delay leaves a gap in the trial calendar that cannot be filled by another waiting case (and the court’s trial capacity is thereby reduced) will it affect the system. But no assurance comes from this study that the courts appreciate how crucially different such postponements are from all others.

The Levin-Woolley book raises still another question. What, if any, research on the problem of delay should be done in one particular court system, for which remedial action is considered? To ask the question differently: which problems are bound to vary significantly from court to court, and which are bound to show only small differences? Data from the Levin-Woolley study indicate the general direction of the answer. They show, for instance, for Allegheny County an inordinately high percentage of cases settled immediately after the trial jury is sworn in, giving rise to the often-heard suspicion that something is wrong with the fee arrangements for defense counsel. But in the main, if one wants to study delay in a particular court system, the task is threefold: First, determine the demand end of our equation—how much court-time would be required to remove the existing backlog? The second job is to determine the supply end—how much court-time is available? The third necessary step, how either to reduce the workload or increase the available court-time, will seldom require basic local studies. The curative powers of certain specific remedies do not vary much from court to court. If, for instance, pretrial, centralized calendar, separate liability trials, or court appointed medical experts should prove useful remedies, the magnitude of their remedial power should be more or less the same everywhere. And the remedial power of the major cures, more judges or more trial time per judge, is clearly assessable without special investigation.

15 Conceivably, there are indirect, secondary effects of minor magnitude—i.e., long delayed cases might become easier (or more difficult) to settle, but such effects are difficult to ascertain. See ZEISEL, KALVEN & Buchholz, op. cit. supra note 1, at 111.

16 For instance, nowhere in the Levin-Woolley study are we told how many postponements fall into the one and how many in the other category.
Seen in this perspective, searching for the many causes of individual delay performs an important and useful function. But valuable as these insights are for a variety of purposes, they have no direct relevance for the problem of court congestion. *Dispatch and Delay* in a way complements *Delay in the Court* by laying bare the anatomy of individual delay where that other study aimed at clarifying the nature of a delayed system; it also strengthens the important trend toward more fact-finding as an aid to the policy-makers in the area of judicial administration.

Yet, the usefulness of this trend is forever threatened by two dangers: one is, simply, that the facts may by-pass the issue; and the other is that the policy-makers may ask for ever more facts—not because they need them, but because such a request is a convenient way to procrastinate. And court delay is replete with procrastination—so much so that time may be running out not only for individual litigants, but also for the system as a whole.