THE SUPPLY OF JUSTICE: THE EFFECT OF JUDGESHIPS ON DISPOSITIONS

by

Michael Beenstock
&
Yoel Haitovsky

The Hebrew University of Jerusalem
Department of Economics

April 12, 2000

We use panel data on Israeli courts to estimate the "production function" for case dispositions. Our results show that the number of case dispositions is independent of the number of serving judges, and that "productivity", as measured by completed cases per judge, varies directly with the caseload per judge. These results suggest that the productivity of judges is endogenous; for the same caseload judges complete more cases under pressure, and complete less when new judges are appointed. They also suggest that the practice of determining the number of judges by fixed "Leontieff" input-output coefficients is not appropriate.

We are grateful to Natasha Lintsky for her excellent research assistance, and to the Eshkol Foundation for financial support in completing the research, and to three referees for their helpful remarks.

"..as the court grows larger, the productivity of individual judges on the court declines."

Judge Gerald Bard Tjoflat (1993)

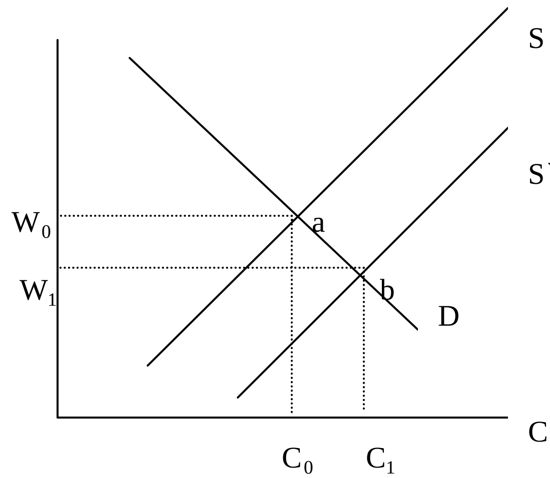
1. Introduction

A key factor in planning the judicial system, and especially in planning the number of judges and their allocation among the different courts, is the effect of the number of judges on the output of the judicial system. Our central concern in this paper is with the effect of the number of judges on the output of the judiciary. In courts where bottlenecks are caused by the number of judges, one might expect, *ceteris paribus*, that the greater the number of judges, the greater should be the output of the judiciary, so that more cases are decided, resulting in reduced waiting times. In short, more judges should enable the system to handle more cases. Indeed, the policy decision to appoint judges is most probably motivated by the desire to ensure that the supply of judicial services matches the demand as measured by the flow of new cases and the number of pending cases.

The Administrative Office of the United States Courts determines the number of judges, *inter alia*, by using "input-output" coefficients for different cases. These coefficients express the amount of judge-time required to judge a case of a given type (currently classified into 153 categories). Using projections of lodged cases by type, these coefficients may then be used to calculate the number of judges required to handle the projected caseload. Our analysis suggests that the application of fixed "input-output" coefficients in the case of the judiciary may be inappropriate (as it generally is in economics) on both theoretical and empirical grounds, because the coefficients are, in fact, variable.

This paper is concerned with the supply-side of the judiciary. To clarify the matter we refer to the diagram where W denotes waiting-time per case and C denotes the output of the judiciary, measured below by case dispositions. The demand for litigation is assumed to vary inversely with waiting-time since civil litigants prefer justice sooner rather than later. Hence, schedule D slopes downward. The supply schedule of court services (S) slopes upwards because, as discussed in section 2, the productivity of judges rises with their caseload. Since, given the number of judgeships and ancillary judicial inputs, waiting-time varies directly with caseload backlog, the output of the court system varies directly with W.

Because court fees are not used to equilibrate supply and demand for court services, waiting-time and the caseload backlog fulfill this function instead. Schedule S is drawn for a given number of judgeships. The initial equilibrium is determined at **a**. This paper is concerned with the slope of schedule S. It is also concerned with the shift in the schedule to S` that is brought about when the number of judgeships is increased. These two concerns are related because the flatter is schedule S the smaller must be the equilibrium change in C. In the limit, if schedule S is horizontal over the relevant range, the output of the judiciary will remain unchanged at **b**, i.e. $C_1 - C_0 = 0$.



Equilibrium in the Market for Justice

To the best of our knowledge there has been no previous empirical attempt to investigate the empirical relationship between judicial inputs and outputs. This explains our unusually thin list of relevant references on this issue. Yet the issue is of central importance in determining the number of judges that should be appointed. If too many judges are appointed and the number of cases pending tends to zero, judges will be under-employed and unnecessary expense will have been incurred. If too few judges are appointed, the number of cases pending will grow without limit since demand outstrips supply. In this case delays will lengthen thereby incurring miscarriage of justice as well as financial costs to litigants.

Although our hypothesis of interest has not been previously investigated it is usually taken for granted that the appointment of more judges necessarily increases the output of the judiciary. Although he does not recommend expansion of the judiciary as a solution to the caseload crisis in the US, Posner (1985) assumes that the output of the judiciary depends, *inter alia*, upon the number of judges. Others, such as McCree (1981) and Gabrys (1998), see expansion of the judiciary as a crucial component of the solution for solving the caseload problem in the US. Starr (1992), Figueiredo and Tiller (1996) and Carlton (1997) show that judicial expansion in the US varies directly with caseload pressure, which is consistent with the hypothesis held by the authorities that more judges necessarily means more justice (but see Tjoflat, 1993 for a dissenting view). We hope that the line of research that we open in this paper will encourage others to examine this issue critically instead of taking the matter for granted.

We use panel data for courts in Israel to investigate the relationship between judicial inputs, as measured by the number of serving judges, and outputs, as measured by the number of cases completed. We find that the relationship is far from straightforward. Our main finding is that judges' productivity, as measured by case completions per judge, is highly adaptive and endogenous; it varies directly with the caseload pressure under which they work. Judges complete more cases the greater is their workload. This happens not because judges work longer hours in court when they are under pressure. Their court room hours are set exogenously. It happens because they use their courtroom time more intensively, so that cases on average consume less courtroom time, whatever its consequences for the quality of judgements. In civil cases litigants may be pressurized to settle out of court, or they may be pressurized to compromise after the commencement of proceedings. Or judges may simply manage courtroom proceedings with greater efficiency, by shortening cross-examinations and reducing the number of witnesses. In criminal cases judges may press for plea-bargains to shorten the courtroom time allotted to the case. Increased efficiency requires effort on the judge's part. They will come to court better prepared by doing more "homework" or they will increasingly defer to "homework" matters that do not strictly require court-time. Robel (1990) shows that all of these responses to caseload pressure have been reflected in the behavior of US judges.

In short, judges' productivity as measured by case completions is potentially elastic and, as described below, can be expected to respond to economic incentives. Just as input-output coefficients in production are unlikely to be fixed, so are they unlikely to be fixed in the judiciary. The appointment of new judges can be expected to reduce the productivity of existing judges since the workload of each incumbent judge will become lighter on average. In the limit, their productivity may fall to such an extent that judges' services as measured by case dispositions may remain unchanged, in which event new appointments would be ostensibly self-defeating. Indeed, we reach this conclusion for some of Israel's courts.

It may, of course, be the case that even if the quantity of judicial output (as measured by case completions) does not change its quality will be improved. Judges may pass better judgments when they are less pressurized. By reducing plea-bargains and out-of-court compromises, and by allowing more witnesses and courtroom deliberation, more justice may be done. However, our data do not enable us to measure judicial output hedonically. Moreover, our data force us to measure output by cases completed. This may be deficient for several reasons. First, judges may deal with specific cases intermittently. Hence the fact that the case is not completed does not mean that there has been no output. Secondly, even judges do not judge intermittently, cases will be as yet incomplete at the end of the data period; output has occurred but it will not be registered until the next time period when completion takes place. Thirdly, cases may be completed without incurring any judicial output; e.g. if litigants reach a compromise prior to courtroom proceedings. Unfortunately, an ideal measure of output which takes account of quality and quantity is not available.

Nor are ideal data available on judicial inputs. Due to data limitations we are forced to measure judicial inputs by the number of serving judges. While the number of judges is no doubt the key judicial input, other inputs include clerical support staff, the number of clerks, computer support and physical conditions in the courtroom. Therefore we cannot explore how judges' productivity depends upon these other inputs. In section 6 we argue that this deficiency most probably does not bias our main findings.

The remainder of the paper is organized as follows. In section 2 we present a theory of endogenous productivity for the representative judge. In section 3 we discuss how this theory may be used to derive a "production function" for case completions for the representative court. The data are presented in section 4 for three separate court systems in Israel: the High Court (including the Supreme Court), the 5 district courts and 19 magistrate courts. The data are annual. In the case of the district and magistrate courts the observation periods are too short for separate econometric analysis of each court. We therefore pool the cross-section and time series data. Since these data are nonstationary we present in section 5 recently developed estimation methods for nonstationary panel data. Results are presented in section 6. Section 7 concludes.

2. Microfoundations

In this section we develop a simple representative agent model which motivates the empirical analysis in section 6. We denote by K the stock of pending cases for a given judge. The instantaneous change in K is defined as the difference between the new cases that are allocated to him (S) and the cases that he completes (C). A proportion δ of cases pending is assumed to terminate before reaching the courtroom due to out-of-court settlements etc. In practice, δ might vary directly with K, since the pressure to settle out of court is likely to vary directly with the caseload. Here, in the interest of simplicity of exposition, we assume that δ is constant, hence

$$\dot{K} = S - C - \delta K \qquad \textbf{(1)}$$

The instantaneous utility of the judge is assumed to be:-

$$U = U(E, K) \qquad \textbf{(2)}$$

where E denotes the judge's effort. We assume $U_E < 0$ and $U_K < 0$, i.e. judges prefer to take it easy, but they also prefer to have a smaller backlog of cases. Marginal utility is assumed to decrease in both E and K, i.e. $U_{EE} < 0$ and $U_{KK} < 0$. We further assume that the "production function" for completions is C(E) where $C_E > 0$ and $C_{EE} < 0$, i.e. the marginal productivity of effort is positive but diminishes. Finally, judges are assumed to maximize discounted utility over their career (T) where:-

$$U = \int_0^T U(t)\, e^{-rt}\, dt + f(K(T)) \qquad \textbf{(3)}$$

and r denotes the relevant discount rate. The function f( ) is specified to ensure that judges do not "bequeath" an unreasonably large caseload upon retirement.

The Hamiltonian of the system is:-

$$H = U(E, K)\, e^{-rt} + \lambda\,(S - C(E) - \delta K) \qquad \textbf{(4)}$$

in which E serves as the judge's control variable and K his state variable, and λ is a costate variable equal to the shadow cost of pending cases. S is assumed to be exogenously decided by the president of the court. The first order conditions for a maximum are:-

$$U_E e^{-rt} - \lambda C_E = 0 \qquad (5)$$

$$-\dot{\lambda} = U_K e^{-rt} - \lambda\delta \qquad (6)$$

with transversality condition:-

$$l_T = f_K(T) \qquad (7)$$

and Legendre condition:-

$$U_{EE}e^{-rt} - lC_{EE} < 0 \qquad (8)$$

which from equation (5) may be rewritten as

$$U_{EE} - \frac{U_E C_{EE}}{C_E} < 0$$

Equation (5) states that judges should equate the discounted marginal disutility of effort with its marginal product weighted by the shadow cost of pending cases. If the latter increases they should work harder to reduce it.

It is intuitively obvious that if a judge's caseload is eased as a result of a reduction in S he will reduce his productivity by lowering E. This happens because he can sustain a given stock of pending cases with less effort, and both E and K will tend to fall. This may be proven formally by solving equations (1), (5) and (6) for K and E in terms of S and other parameters[1]. No doubt this simple model does not capture the strategic behavior that may take place both between judges and with the president of the court. The president may allocate new cases to his most efficient judges. Moreover, promotion of judges most probably depends on their productivity. Therefore competition between judges can be expected to increase the disposition rate, and in a more general model S would be endogenous (Beenstock and Haitovsky, 1999).

3. From Micro to Macro: The Model of the Court

Suppose that J_0 homogeneous judges serve in the court and each judge is allocated $TS_0/J_0 = S_0$ new cases, where the prefix T denotes court aggregates, i.e.TS denotes the number of new cases lodged at the court in the relevant time period. Following the analysis in section 2, each judge will decide on an optimal amount of effort (E_o) given S_0. The total number of cases completed per time period will be $TC_0 = J_0C(E_0)$, and the number of cases pending will be $(TS_0 - TC_0)/\delta = TK_0$ in the stationary state. This shows that even if TS – TC is small TK may tend to infinity or to a lower bound of zero if δ is small.

Suppose that the number of judges is raised to J_1 and the number of cases lodged is unchanged, so that each judge is allocated $S_1 < S_0$ new cases. The number of cases completed per time period will be $TC_1 = J_1C(E_1)$. Judges respond to the new appointments by "taking it easier"; their productivity is endogenous. Hence, if the number of judges is raised by one percent the number of completions by the court can be expected to increase by less than one percent. In the limit the elasticity may be zero. This happens when E_1 is sufficiently below E_0. The size of this elasticity is an empirical matter. It depends on the nature of the utility function. However, even competitive judges will tend to take it easier when the pressure is off.

Suppose next that J is unchanged but TS rises to TS_2, in which case each judge is allocated $S_2 = TS_2/J_0$ new cases instead of S_0 cases. In this case (Robel, 1990) judges will raise their efforts to E_2, and the output of the court will rise to $TC_2 = J_0C(E_2)$. Since there is a natural upper limit to the ability of judges to boost their effort and $C'' < 0$, we may expect that successive increases in TS will induce smaller increases in TC.

We do not observe the efforts of individual judges or even their caseloads. Instead, as described more fully in section 4, we assembled a panel data base for individual courts. These data include case completions, cases lodged and cases pending (disaggregated by type), as well as the number of judges serving in the court. The basic model that we estimate in section 6 is:-

$$\ln TC_{it} = \alpha + \beta \ln TS_{it} + \gamma \ln TK_{it} + \delta \ln J_{it} + \varepsilon_{it} \qquad (9)$$

where i = 1, 2, ..., N identifies the court, t = 1, 2, ..., T identifies the time period, and ε is an error term. Although in the simplified model in section 2 $\gamma = 0$, in practice the backlog may have an independent effect on judges' efforts, since at any point in time the backlog may be out of equilibrium. If K is larger, *ceteris paribus*, judges may intensify their efforts to reduce it. According to the discussion above we expect $0 < \delta < 1$.

As noted in section 1, equation (9) should also specify other judicial inputs, such as clerical support and clerks. Absence of the necessary data means that these variables are incorporated into ε. We discuss the implications of these omitted variables below. However, in contrast to the US there has been no widespread use in Israel of law clerks, staff attorneys and ADRs. Intercourt differences in the allocation of these inputs will be picked up, to some extent, by the specification of specific effects, as discussed in section 5. Intertemporal differences may be captured, to some extent, by specifying a time trend in equation (9), as discussed in section 6.

An alternative formulation is to define "productivity" in terms of average completions per

---

[1] For a formal exposition see Beenstock and Haitovsky (1999).

judge (P = TC/J) and "caseload" as average caseload per judge (L = (TS+TK)/J). We then regress P on L and J:-

$$P_{it} = q + l\ L_{it} + f\ J_{it} + e_{it} \qquad (10)$$

Equations (9) and (10) differ only in their functional form, the choice being determined on empirical grounds. In practice we disaggregate L into TS/J and TK/J with coefficients λ_1 and λ_2 respectively.

4. The Data

The data consist of annual observations on the three court systems in Israel. With the exception of the data for judges (which were abstracted from various issues of the Government Yearbook) all the data were collected from various issues of Judicial Statistics, published annually by the Central Bureau of Statistics.[2] The number of courts and period coverage are summarized in Table 1. The order reflects their position in the judicial hierarchy. District Courts hear more important cases and serve as the court of appeal for magistrate courts. High Court judges serves two functions; they hear appeals from the District Courts in the High Court and they hear constitutional cases in the Supreme Court.

In Figure 1 we plot the time series for the High Court (including Supreme Court cases). All the series trend upwards. Figure 2 plots the relationship between "productivity" and "caseload" that is implied by these data. The pattern is clear; λ (in equation 10) is apparently positive. Figures 3 parallel Figure 1 for the district courts. Here too the data trend upwards, including cases pending. Figure 4 parallels Figure 2, suggesting once more that productivity varies directly with caseload. Since there are too many magistrate courts to present individually we limit ourselves in Figure 5 to plotting the relationship implied by the data between productivity and caseload. Here too the data trend upwards. Hence, Figure 5 parallels Figures 4 and 2, further confirming the positive association between productivity and caseload.

Note that while the number of judges and lodged cases tend to increase over time, i.e. the system was generally expanding, there are periods in the data during which the system temporarily contracts (see e.g. Figures 3a, 3b and 3e). The positive association between productivity and caseload holds both when the system is expanding and when it is contracting.

[2] We wish to thank Tamar Turjman for her painstaking efforts in assembling the data. The data are available at the Social Sciences Data Archive at the Hebrew University of Jerusalem.

5. Econometric Methodology

The underlying model to be estimated is either equation (9) or equation (10). Since, as discussed in section 4, these data are nonstationary conventional estimation of these equations would run the risk of spurious correlation. If equations (9) and (10) held instantaneously, a solution to the spurious correlation problem would be to estimate using first differences of the data. Since, however, these equations refer to the steady-state and are hypothesized to hold in the long-run it cannot be assumed that they hold instantaneously. Hence, first differencing would generally induce misspecification (see e.g. Hendry, 1995). We therefore use cointegration analysis to estimate the long-run parameters of the model. Equation (9) will be cointegrated when $\varepsilon \sim I(0)$, i.e. when its residuals are stationary. Equation (10) will be cointegrated when $e \sim I(0)$. Taken together, equation (9) and the discrete counterpart of equation (1) form a multicointegrated system.

When the number of time series observations is insufficient we pool the time series and cross section observations to generate sufficient degrees of freedom. However, we do not pool courts from different hierarchies (e.g. District Courts and Magistrate Courts) because they are essentially different. Im, Pesaran and Shin (1997), hereafter IPS, have suggested the following unit root test for nonstationarity in panels, based on the assumption that both N and T go to infinity, with the condition that T goes to infinity faster than N. First, estimate the augmented Dickey-Fuller statistics (ADF_i) for each panel using p augmentations. Then calculate

$$z = \frac{\sqrt{N}\left(|\overline{ADF}| - a(T,\ p)\right)}{\sqrt{b(T,\ p)}} \sim N(0,\ 1) \tag{11}$$

where $|\overline{ADF}|$ is the absolute mean of the N ADF statistics and a and b are derived by IPS from Monte Carlo simulations. Harris and Tzavalis (1999), hereafter HT, have suggested an alternative test for stationarity in panels in which T is fixed, but N tends to infinity. This procedure, unlike equation (11), assumes that the panel unit root is homogeneous. Harris and Tzavalis (1988) have extended their approach by allowing for cross-sectional serial correlation of the MA(1) variety and heteroscedasticity in the disturbances. Like the IPS test for unit roots, the HT test is normally distributed.

Equation (11) tests for stationarity in panels with heterogeneous dynamics. Strictly speaking it is not a cointegration test. However, Pedroni (1999) suggests that if the slope coefficients in equations (9) and (10) are homogeneous (matters are quite different otherwise) the application of the IPS or HT panel unit root tests to panel model residuals is appropriate. The null hypothesis of no cointegration is rejected when z exceeds its critical value. This application of unit root econometrics to nonstationary panel data joins the rapidly expanding number of empirical applications (see, e.g., Breitung and Meyer, 1994, Coakley and Fuertes, 1997, Coe and Helpman, 1995, MacDonald, 1996, and Maddala and Kim, 1998).

Because the data are nonstationary the parameter estimates in equations (9) and (10) will be "super-consistent", especially in our case where, due to lack of data, the structural parameters are assumed to be homogeneous. This has the effect of asymptotically removing possible simultaneous equations bias from the parameter estimates (Stock, 1987 and Hsiao, 1997), induced by the identity in equation (1) and the feedback effect, noted by Figueiredo and Tiller (1996) and others, of caseload pressure on the appointment of judgeships. Therefore, as long as our concern is limited to the long-term production function for justice, we need not concern ourselves (asymptotically at least) with the possible endogeneity of judges, i.e. that the number of judges serving in court i depends *inter alia* upon its caseload.

Finally, we estimate equations (9) and (10) under different assumptions about their disturbance structures. These include fixed and random effects models by court, weighted least squares (WLS, allowing for heteroscedasticity by court), and seemingly unrelated regression (SURE, allowing the disturbances to be correlated between courts), in addition to ordinary least squares (OLS). Since the sample of courts comprises the whole population of courts in Israel a fixed effects specification is preferable to a random effects specification (Hsiao, 1986). We do not specify fixed effects for time periods since they were not statistically significant. However, we do experiment with homogeneous time trends in section 6.3.

We have chosen not to report “best” models. Instead we have taken a pluralistic approach in reporting our econometric findings so that the reader may ascertain their robustness. Hence we experiment with different specifications of heterogeneity by court, different specifications of disturbance structures, different functional forms, and we report different tests for cointegration. The latter are particularly important given that unit root tests are known to have low power.

6. Results

6.1 High Court Judges

We begin by presenting results for the High Court. Since the data span 30 years there is no need for pooling. We conduct two types of cointegration test: the two stage OLS method proposed by Engle and Granger (1987), and the maximum likelihood (ML) method proposed by Johansen (1995). These tests are applied to two specifications of the model as represented by equations (9) and (10) in parts A and B respectively of Table 2.

The main conclusions arising from part A are that the specification in equation (9) is cointegrated[3], however, the number of judges (J) is not statistically significant; it does not form part of the cointegrating vector. This means that in the long run the number of judgeships does not affect the number of dispositions. In models 1 and 3 $\hat{d}$ is negative, implying that dispositions vary inversely with the number of high court judges. However, this effect is not statistically significant. Indeed, all models indicate that cointegration is entirely due to lodged cases (TS); cases pending have no significant effect on dispositions.

The results in part B of Table 2 indicate that "productivity" in the High Court varies directly with the caseload. However, new cases lodged have a significantly greater effect than cases pending (i.e. λ_1 is greater than λ_2). Both the ML and OLS models are cointegrated[4]. Models 1 and 2 indicate that cointegration is not sensitive to the specification of the VAR, while models 4 and 5 suggest that productivity varies inversely with the number of judges. The implied partial derivative[5] of case completions with respect to judges in model 4 is 208 - 30.6J. Since typically J=14 (see Figure 1) model 4 implies that completions tend to vary inversely with the number of judges.

We have already observed that we do not have data for judicial inputs other than the number of serving judges. If these omitted variables (Q) happened to be nonstationary and proportionate to J, i.e.

$$\ln Q_{it} = \phi_i + \omega \ln J_{it} + u_{it}$$

with $u \sim N(0, \sigma^2_u)$, then δ in equation (9) would capture the combined effect of both observed and unobserved judicial inputs. If Q happens to be stationary its omission from the model will not asymptotically affect the parameter estimates and cointegration. If Q is nonstationary and does not cointegrate with J then its omission will induce non-cointegration. Unlike the US where the growth in law clerks and staff attorneys has most probably induced nonstationarity in Q, in Israel there has

[3] Although the ADF statistic falls short of its critical value the PP statistic exceeds it.
[4] Here too ADF falls short of its critical value while PP exceeds it.
[5] To calculate $\partial C/\partial J$ we multiply both sides of equation (10) by J, differentiate and rearrange terms to obtain $\partial C/\partial J=\theta+2\phi J$.

been no such growth. Hence we do not expect (asymptotically) that in our case the omission of Q will induce non-cointegration.

In summary, we could not find any evidence to support the hypothesis that in the long run dispositions of the High Court depend upon the number of judges. The productivity of judges varies directly with new cases lodged and to a lesser degree with pending cases.

6.2 District Courts

Here we pool the cross section and times series data for the five district courts because there are only 19 annual observations for each court[6]. Since the Johansen approach to cointegration has not been developed for pooled data we only report least squares estimators, which vary according to their specification of the specific court effect (fixed vs random) and their disturbance specification (OLS, GLS and SURE). The results are presented in Table 3.

Parts A and B of Table 3 parallel their counterparts in Table 2. All models easily pass the IPS[7] and HT cointegration tests for panel data (the critical one-tail value for z is 1.68 at 0.05 significance level). This is also reflected in the panel Durbin-Watson statistics which suggest that the residuals do not contain a unit root. Part A indicates that case dispositions do not depend on the number of judges. As in Table 2, $\hat{d}$ is mostly negative but not statistically significant. This result is independent of the estimation method used. Models 1 to 3 imply that dispositions are close to being linear, homogeneous in pending and lodged cases with the dominant effect taken by lodged cases. However, a comparison between models 1 and 4 indicates that this depends upon the specification of the specific effects. Since a random effects model fits the data less well than a fixed effects model, and since it is conceptually less appropriate, we prefer model 1 to model 4. Model 1 in part B indicates that productivity varies directly with the caseload. Here too the residuals are stationary, except perhaps when p=2, suggesting that the model is cointegrated. A random effects version of model 1 generates similar results. Model 2 indicates that judge productivity increases with the number of judges (see also model 5).Model 3 suggests that the effect of caseload pressure may be nonlinear and increasing. Models 4 through 6 suggest that lodged cases have more than twice the effect on productivity than pending cases. Finally, model 7 indicates that, although these results are not sensitive to the method of estimation, they are sensitive to the specification of effects. The

[6] We take the view that pooling is a necessary evil to be applied when the number of individual time-series observations is insufficient.

[7] The p-value varies inversely with the number of augmentations (p) in the case of IPS.

estimates of the fixed effects imply that court productivity is greatest in Nazareth and smallest in Haifa and Jerusalem. We checked to see whether this results from the case material heard in the different courts. Perhaps the cases heard in Nazareth happen to be easier. However, we found no effect of material composition (criminal vs civil, pecuniary vs injury, etc) on the parameter estimates.

In summary, there is apparently no evidence that dispositions in the District Courts depend upon the number of judges. Instead, the caseload (especially lodged cases) induces judges to increase their productivity.

6.3 Magistrate Courts

Tables 4 and 5 report estimates of equation (9) for the 3 large magistrate courts and the 16 smaller ones respectively. We do not pool them because the number of observations differs, and because the smaller courts tend to hear lighter cases. We do not report estimates of equation (10) in view of the fact that $\hat{d}$ turns out to be both statistically significant and positive, implying for the first time that dispositions vary directly with the number of judges. However, this effect is neither large nor robust. For example, a comparison of models 1 and 5 in Table 5 indicates that it depends upon the specification of the effects, and a comparison of models 3 and 4 indicates that it depends upon the specification of the residuals. In the three large magistrate courts (Table 4) this coefficient is robust; the estimated elasticity of completions with respect to judges is of the order of 0.14. Table 4 further suggests that cases pending have no effect on dispositions and that there may be a negative time trend in "total factor productivity" of the order of about 0.4% per year. According to the estimated fixed effects in Table 4 dispositions in the Tel-Aviv and Haifa courts are about 4% higher than in Jerusalem.

Model 3 in Table 5 implies that lodged cases have a greater effect on dispositions than pending cases, suggesting that judges dispose of new cases in order to keep pending cases from growing. Model 3 also implies that while the marginal product of judges is positive, it varies inversely with the number of judges. Model 3 further indicates that there is a weak but negative time trend in "total factor productivity" for the smaller magistrate courts. The estimated fixed effects imply a 33% difference between the courts with the highest conditional dispositions (Tiberias, Afula and Safed) and the court with the lowest (Herzlia).

6.4 Discussion

Only in the case of magistrate courts do we find any evidence that the production of justice as measured by case dispositions depends upon the number of judges. And even in this case the effect is small and not robust. By contrast all the models indicate that productivity varies directly with the caseload. These findings suggest that the productivity of judges is endogenous. Indeed, it is endogenous to such an extent that changes in the number of judges are offset by countervailing changes in productivity, such that the total dispositions by court are not sensitive to the number of judges.

This result may be understood in terms of the theoretical model presented in section 2. When their caseload increases their productivity as measured by dispositions rises. When more judges are appointed their caseload falls and judges on average reduce their productivity. The data indeed indicate that productivity has tended to increase over time due to the growing caseload. However, it has not grown sufficiently fast to stabilize the backlog per judge; the backlog has grown too. This process has persisted for more than 20 years in the case of some of the courts studied. One wonders how much longer this can continue before judges reach some upper limit of coping.

On the other hand, appointing more judges will not stabilize the backlog, except possibly in the magistrate courts. If more judges are appointed, the existing judges will simply adapt their behavior by reducing their productivity as measured by dispositions. This implies that the backlog apparently has a life of its own and that it does not depend on the number of judges. However, if there is autonomous growth in cases lodged, judges will tend to step up their disposition rate in an attempt to keep the backlog from growing too rapidly.

Our results show that after controlling for caseload courts have different rates of disposition which seem to be independent of caseload material. On the face of it this suggests that some courts are more efficient, for example the district court in Nazereth and the magistrate courts in Tiberias, Afula and Safed. These courts happen to be in close geographical proximity.

7. Conclusion

To our best knowledge this is the first attempt to estimate empirically the relationship between the change in the number of judges and the change in the number of case dispositions. In doing so, we have used recently developed panel unit root tests to investigate the long-run determinants of case dispositions in Israeli courts. The main findings are twofold. First, pressure produces productivity; the number of case dispositions per judge varies directly with his caseload.

Secondly, the number of case dispositions does not apparently depend on the number of judges. These findings imply that the appointment of judges does not help clear the backlog. Instead, the backlog pressurizes judges into completing more cases. Nevertheless, as the number of lodged cases grows the backlog increases despite the increased "productivity" of judges. These results apply across the hierarchy of courts in Israel.

If our findings for Israel are relevant to other countries, they suggest that the fixed coefficient “input-output” approach as practiced, for example, by the Administrative Office of the United States Courts, will generally be inappropriate. Our results imply that these coefficients are adaptive; they depend upon the workload. The amount of judge-time required to complete a case varies inversely with his caseload. If lodged cases tend to grow over time, this implies that projections of judge shortages will generally be over-estimated. Perhaps this is the reason why the US federal court system has not collapsed 15 years after Posner’s first discussed the caseload explosion.

In terms of the theoretical model presented in section 2 this pattern of behavior is consistent with the hypothesis that incumbent judges work less hard when new judges are appointed; the smaller caseload reduces the pressure under which they operate. A more normative interpretation is that the quality of justice increases even if the quantity, as measured by case dispositions, does not. When under pressure, judges may press for compromises and plea bargains, streamline courtroom hearings, and undertake other initiatives to save courtroom time. Hence, it is conceivable, as argued by Posner (1985), Robel (1990) and others that cases are completed with less justice. This would imply, that the appointment of new judges enables incumbent judges to bring cases into the courtroom which otherwise would be closed in a technical manner, allow more witnesses and cross examination, and spend more time on writing judgments, in addition to working less hard. Unfortunately, we cannot corroborate this hypothesis due to lack of sufficiently detailed data on variables such as the quantity and quality of written judgments, courtroom time per case and the leisure time of judges. There is a need for data on the quality of justice and not just its quantity.

Absence of data on other judicial inputs prevented us from investigating whether they are responsible for the increase in judges’ productivity, rather than pressure of work. It is arguable that although the number of judges grew slower than their caseload, other judicial inputs grew sufficiently faster, such that the productivity of judicial inputs as a whole did not change. We reject this argument for several reasons. First, as mentioned, informal data for Israel on other judicial inputs suggest that they did not in fact grow faster than the number of judges. Secondly, experimentation with time trends in section 6, which might have picked up trend-like improvements in productivity, did not change our basic conclusion. Third, the cross section comparison clearly shows that at a given point in time it is the workload that explains inter-court differences in productivity. Finally, when the caseload contracts productivity falls, and when the number of judges contracts productivity increases. Hence our results apply both when the system expands and contracts.

References

Beenstock, M. and Y. Haitovsky (1999), "The Supply of Justice", Working Paper, Department of Economics, Hebrew University of Jerusalem.

Breitung, J. and W. Meyer (1994), "Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?", **Applied Economics**, 26: 353 - 61.

Carlton, C. F. (1997), "The Grinding Wheel of Justice Needs Some Grease: Designing the Federal Courts of the 21st Century", **Kansas Journal of Law and Public Policy**, Summer/Fall.

Coakley, J. and A.M. Fuertes (1997), "New Panel Unit Root Tests of PPP", **Economics Letters**, 57: 17 - 22.

Coe, D.T. and E. Helpman (1995), "International R&D Spillovers", **European Economic Review**, 39:859–888.

De Figueiredo, J.M. and E.H. Tiller (1996), "Congressional Control of the Courts: a Theoretical and Empirical Analysis of Expansion of the Judiciary", **Journal of Law and Economics**, 34: 435-462.

Enders, W. (1995), **Applied Econometric Time Series**, John Wiley, New York.

Engle R.F. and C.W.J.Granger (1987), "Cointegration and Error Correction: Representation, Specification and Testing", **Econometrica**, 55: 251-76.

Gabrys, M.E. (1998), "A Shift in the Bottleneck: the Appellate Caseload Problem Twenty Years after the Creation of the Wisconsin Court of Appeal", **Wisconsin Law Review**, 6: 1547-1594.

Government Yearbook, various years, Jerusalem, Israel.

Im, K.S., M.H. Pesaran and Y. Shin (1997), "Testing for Unit Roots in Heterogeneous Panels", Department of Applied Economics, University of Cambridge, March.

Harris, R.D.F. and E. Tzavalis (1999), "Inference for Unit Roots in Dynamic Panels where the Time Dimension is Fixed", **Journal of Econometrics**, forthcoming.

Harris, R.D.F. and E. Tzavalis (1998), "Inference for Unit Roots in Dynamic Panels with Heteroscedastic and Serially Correlated Errors", mimeo, Department of Economics, University of Exeter (June).

Hendry, D. F. (1995), **Dynamic Econometrics**, Oxford University Press.

Hsiao, C. (1986), **Analysis of Panel Data**, Cambridge University Press.

Hsiao, C. (1997), "Cointegration and Dynamic Simultaneous Equations Models", **Econometrica**, 65: 647-670.

Johansen, S. (1995), **Likelihood-based Inference in Cointegrated Vector Autoregressive Models**, Oxford University Press.

Judicial Statistics, various years, Central Bureau of Statistics, Jerusalem, Israel.

MacDonald, R. (1996), "Panel Unit Root Tests and Real Exchange Rates", **Economics Letters**, 50: 7 - 11.

Maddala, G.S. and I-M. Kim (1998), **Unit Roots, Cointegration, and Structural Change**, Cambridge University Press.

McCree, W.H. (1981), "Bureaucratic Justice: an Early Warning", **University of Pennsylvania Law Review**, 129: 777-797.

Pedroni, P. (1999), "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis", **Oxford Bulletin of Economics and Statistics**.

Posner, R.A. (1985), **The Federal Courts: Crisis and Reform**, Harvard University Press.

Robel, L.K. (1990), "Caeload and Judging: Judicial Adaptations to Caseload", **Brigham Young University Law Review**, 1: 3-65.

Starr, K.W. (1992), "The Supreme Court and the Federal Judicial System", **Case Western Reserve Law Review**, 42: 1209-1222.

Stock, J. (1987), "Asymptotic Properties of Least-Squares Estimators of Cointegrating Vectors", **Econometrica**, 55: 1035-1056.

Tjoflat, G.B. (1993), "More Judges , Less Justice", **American Bar Association Journal**, 79: 70-73.

Table 1. The Data

Court	**Number of Courts**	**Observation Period**
High/ Supreme	1	1964-1995
District	5	1975-1994
Magistrate:		
Large	3	1975-1994
Small	16	1980-1992

Table 2: Results for High Court (1965-1995) : Cointegration Test

Part A

$$\ln T\text{C} = a + b \ln T\text{S} + g \ln T\text{K} + d \ln \text{J}$$

Model	α	β	γ	δ	Cointegration Test
1	1.09 (0.48)	0.859 (0.117)	0.045 (0.035)	-0.127 (0.296)	ADF_2=-2.42 PP_2=–4.66
2	1.171 (0.454)	0.806 (0.079)	0.054 (0.032)		ADF_2 = – 2.29 PP_2 = –4.69
3	1.015 (0.234)	0.932 (0.072)	0.0056 (0.018)	-0.210 (0.137)	J-DSS Rank=1

Notes: In models 1 and 2 ADF_p denotes the augmented Dickey-Fuller "t" statistic with p augmentations, and PP_p denotes the Phillips-Perron statistic with truncation at lag p. See, e.g. Enders (1995). Standard errors are reported in parentheses in all tables. Note that OLS parameter estimates generally have non-standard distributions. The critical values of ADF and PP for cointegration are in MacKinnon (1991). In model 3 J-DSS_p denotes that Johansen's cointegration test has been applied using a VAR of order p in which the data are assumed to be difference stationary. The reported rank is equal to the number of cointegrating vectors at $p < 0.05$. Elswhere we use TSS to denote the data are assumed to be trend staionary in the VAR.

Part B

$$P = q + l_1 L_S + l_2 L_K + f J$$

Model	θ	λ_1	λ_2	φ	Method/coint-egration test
1	128.6 (22.35)	0.511 (0.087)	0.283 (0.061)		J-DSS Rank=1
2	130.2	0.458 (0.096)	0.319 (0.068)		J-TSS Rank=1
3	85.04 (29.13)	0.637 (0.109)	0.192 (0.082)		ADF_3=-1.86 PP_3=-4.76
4	207.62 (39.70)	0.757 (0.08)	0.198 (0.05)	-15.29 (4.21)	J-DDS Rank=1
5	149.96 (46.15)	0.734 (0.12)	0.18 (0.07)	-9.40 (5.31)	ADF_3=-2.57 PP_3=-4.78

See notes to Table 2A.

Table 3: Results for District Courts (1976-1995): Panel Cointegration Tests

Part A

$$\ln TC = a + b \ln TS + g \ln TK + d \ln J$$

Model	α	β	γ	δ	Method	DW	HT	IPS
1	Fixed	0.843	0.168	-0.018	OLS	1.49	5.06	4.65
		(0.049)	(0.049)	(0.081)				2.64
2	Fixed	0.839	0.177	-0.017	GLS	1.82	5.00	4.60
		(0.041)	(0.036)	(0.067)				2.51
3	Fixed	0.805	0.155	0.033	SUR	1.47	4.94	4.31
		(0.035)	(0.026)	(0.057)				2.4
4	Random	1.084	-0.050	-0.021	OLS	0.93	5.69	4.71
		(0.008)	(0.012)	(0.018)				2.56

Notes: IPS denotes the z-statistic using values for a and b in equation (11) supplied by Im, Pesaran and Shin (1997), table 3. The upper statistic assumes p=0, the lower assumes p=2 (two augmentations). HT denotes the z-statistic using the method of Harris and Tzavalis (1998b).

Part B

$$P = q + l_1 \frac{S+K}{J} + l_2 \left(\frac{S+K}{J}\right)^2 + f\,J + l_3 \frac{S}{J} + 1000\, l_4 \left(\frac{S}{J}\right)^2 + l_5 \frac{K}{J} + 10000\, l_6 \left(\frac{K}{J}\right)^2$$

Model	φ	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	ϕ	Method	DW	HT	IPS
1	Fixed	0.703							GLS	1.32	4.13	2.73
		(0.017)										1.31
2	Fixed			0.6	0.116	0.405	-0.476	3.024	OLS	1.647	5.63	3.88
				(0.176)	(0.075)	(0.233)	(1.95)	(4.31)				2.89
3	Fixed	0.491	7.42						GLS	1.36	4.36	3.67
		(0.086)	(2.99)									2.58
4	Fixed	0.703						8.021	GLS	1.32	4.11	2.95
		(0.017)						(4.295)				2.02
5	Fixed			0.575	0.106	0.494	0.074	3.024	SUR	1.46	5.09	3.54
				(0.098)	(0.041)	(0.122)	(1.1)	(4.31)				2.46
6	Fixed			0.857		0.348			GLS	1.60	5.57	3.91
				(0.034)		(0.071)						2.72
7	Random			1.016		0.013			GLS	1.54	6.04	3.87
				(0.023)		(0.032)						2.81

See notes to Table 3.

Table 4: Results for Magistrates Courts: Panel Cointegration Tests

(Tel-Aviv, Jerusalem and Haifa), 1971-1995

$$\ln TC = a + b \ln TS + g \ln TK + d \ln J + t\, t$$

Model	α	β	γ	δ	ι	Method	DW	HT	IPS
1	Fixed	0.898	0.064	0.085	-0.0058	SUR	1.64	5.75	4.95
		(0.04)	(0.028)	(0.045)	(0.0027)				3.57
2	Fixed	0.921	0.0067	0.167	-0.0034	OLS	1.83	6.5	5.873
		(0.05)	(0.038)	(0.72)	(0.004)				3.95
3	Fixed	0.922	0.009	0.159	-0.0035	GLS	1.88	6.48	5.58
		(0.051)	(0.039)	(0.074)	(0.004)				3.24
4	Fixed	0.924		0.004		GLS	1.89	6.39	5.58
		(0.03)		(0.0016)					3.17
5	Fixed	0.912	-0.014	0.142		GLS	1.92	6.67	5.55
		(0.048)	(0.03)	(0.071)					3.99

See notes to Table 3A.

Table 5: Results for 16 Small Magistrates Courts, 1981-1993: Panel Cointegration Tests

$$\ln TC = a + b \ln TS + g \ln TK + d \ln J + t\, t$$

Model	α	β	γ	δ	τ	Method	DW	HT	IPS
1	Fixed	0.741	0.125	0.129		OLS	1.82	2.12	2.81
		(0.042)	(0.026)	(0.044)					0.03
2	Fixed	0.837	0.148	0.06	-0.0085	GLS	1.85	2.26	7.25
		(0.029)	(0.024)	(0.028)	(0.0032)				1.92
3	Fixed	0.744	0.17	0.151	-0.0089	OLS	1.86	2.05	2.87
		(0.042)	(0.036)	(0.045)	(0.0047)				0.65
4	Random	0.962	0.0612	0.003	-0.001	OLS	1.66	2.18	3.87
		(0.026)	(0.03)	(0.03)	(0.003)				0.09
5	Random	0.95	0.059	0.016		OLS	1.717	2.74	3.87
		(0.27)	(0.027)	(0.031)					0.01

See notes to Table 3A.