Crime and Justice Statistics

NSW Bureau of Crime Statistics and Research

Bureau Brief

November 2001

FORECASTING TRIAL DELAY IN THE NSW DISTRICT CRIMINAL COURT

Marilyn Chilvers

This paper describes a simple forecasting mechanism for predicting trial court delay. By the method of time series regression modelling described here, changes in delay are forecast four quarters in advance using information about the number of current trial matters pending. The model shows that a 10 per cent increase in backlog leads to a 6.2 per cent increase in delay the following year.

INTRODUCTION

During the 1990s, the NSW District Criminal Court experienced difficulties in bringing matters to trial expeditiously. For most of the decade, median trial court delay, from case committal to finalisation, exceeded 12 months. A study published recently by the Bureau of Crime Statistics and Research (BOCSAR) identified a number of inefficiencies in trial case processing.[1] A wide range of reforms in the District Court have been implemented to reduce or eliminate these inefficiencies, including centralised committal proceedings, reforms to the trial listing process, more flexible judicial holiday arrangements, and the appointment of additional judges. Since 1999 there have been marked improvements in delay.[2]

Notwithstanding these changes, it is clear that the capacity of the Court to achieve its time standards is inevitably affected by the volume of cases awaiting trial. Many of the factors identified by BOCSAR as affecting delay are difficult to enumerate or to incorporate into the Court's planning process. The development of a simple predictive model which identifies the impact of current changes in pending caseload on future delay may prove useful in the Court capacity planning process. This paper develops such a model using time series regression analysis.

RELATIONSHIP BETWEEN MEDIAN DELAY AND BACKLOG

Figure 1 plots the relationship between median trial delay and trials pending (backlog) on a quarterly basis between January 1997 and December 2000. Median delay is plotted on the left-hand vertical axis and pending cases on the right-hand axis.[3] The backlog series has been lagged by four time periods.

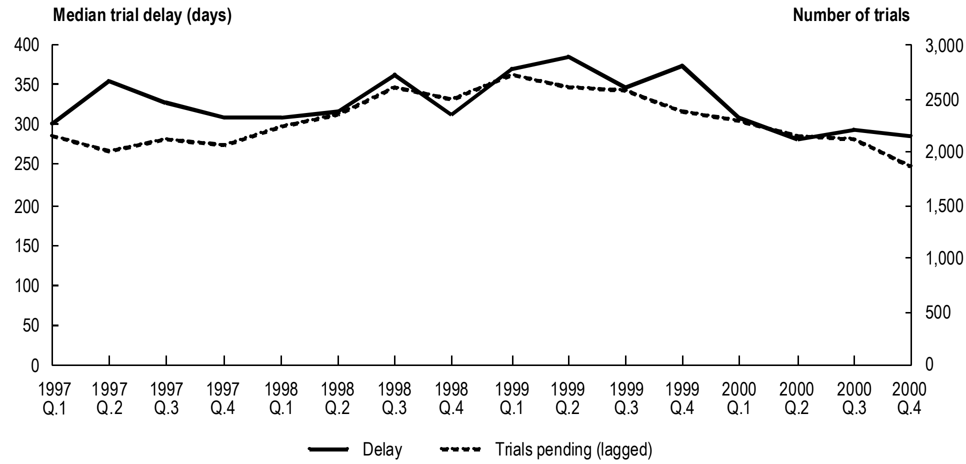
There are two features evident in the series shown in Figure 1 which indicate suitability for time series regression modelling. Firstly, the two series are positively correlated. That is, a movement in one time series is associated with a movement in the other. In particular, the change in the level of median trial delay is preceded by a (same direction) change in the level of the pending caseload. The size of the lag for optimal prediction purposes (i.e. four time periods or quarters) has been illustrated in Figure 1 by graphing backlog, lagged by four periods, against the current level of delay. The method used for determining the optimal lag structure will be described in the next section.

The second time series feature which is evident in Figure 1 is the quarterly seasonal movement in pending caseload. This movement is significantly affected by the functioning of the court. For example, after a period of decreased court activity, such as the January break, the backlog of cases awaiting trial increases, resulting in a systematically high value in the first quarter of each year. Each of these two features, in turn, is incorporated into the predictive models described below.

MODEL A

A simple linear regression model is now developed between the two series in Figure 1 to incorporate the first feature noted in the graph. The lag structure of the modelled relationship was determined by examining the Pearson correlation between the two series at various lags. The highest significant correlation between the two series is for pending

Figure 1: Median trial delay (days) and number of trials pending one year previously, NSW District Court, 1997 to 2000, by quarter



caseload leading median delay by four quarters (r = 0.68). Median trial delay was therefore regressed on pending trial caseload lagged by four quarters using the method of ordinary least squares (OLS). The estimation equation (1) is as follows:

TRIAL_MEDDELAY = C(1)*TRIALS(-4) + C(2)   (1)

Table 1 shows the output from the OLS regression model described above. There is only one predictor variable in the model: the variable, *trials(-4)*, represents the number of cases awaiting trial (pending caseload) at the end of each quarter, one year previous to the period for which delay is estimated. A constant term is also included in the model.

The coefficients in Table 1 are substituted into equation (1) to yield the following mathematical representation of the relationship between the trial pending caseload and median delay.

TRIAL_MEDDELAY = 0.091*TRIALS(-4) + 118.376   (2)

From equation (2), we see that median delay during the period January 1997 to December 2000, may be modelled as a base (constant) level of 118 days, plus 9.1 per cent of all cases awaiting trial four quarters previously. The elasticity of median delay relative to backlog is calculated as 6.2 per cent.[4] That is, a 10 per cent increase (or decrease) in backlog, results in a 6.2 per cent increase (or decrease) in median trial delay one year hence.

The fit of the model is reasonable (R-squared = 0.46) and there are no problems with the underlying assumptions of OLS regression, as shown by the model diagnostics in Table 1. The error terms from the regression model are approximately normally distributed, and there is no serial correlation evident in the model. Examination of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the residuals, however, shows a spike at lag four. Though not significant, this suggests that the seasonal pattern in the data could be modelled.[5]

Table 1: OLS Regression output for model A: Predicting median trial delay from pending caseload lagged by one year

Dependent Variable: TRIAL_MEDDELAY
Method: Least Squares

Variable	*Coefficient*	*Standard error*	*t-statistic*	*Probability*
TRIALS(-4)	0.0913	0.0262	3.4864	0.0036
Constant	118.3763	60.2851	1.9636	0.0698

Model diagnostics	
R-squared	0.4647
F-statistic	12.1549
Prob (F-statistic)	0.0036
Prob (Jarque-Bera)	0.5166
Durbin-Watson statistic	1.9344

Figure 2: Actual versus fitted median delay - Model A
NSW District Court, 1997 to 2000, by quarter

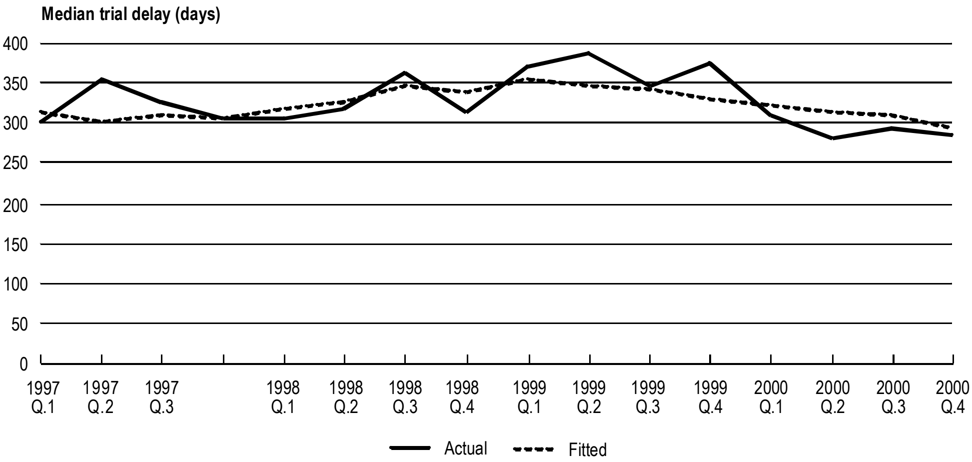


Figure 2 shows the graph of actual median trial delay compared with that which is estimated by equation (2). There is a reasonably close correspondence between the two series, though the modelled series is smoother. This suggests that the inclusion of a seasonal variable may be beneficial. We approach this task in the next section.

MODEL B

To incorporate a seasonality term into model A we include a quarterly moving average term into equation (1) as follows:

TRIAL_MEDDELAY = C1*TRIALS(-4) + C(2)
+ [MA(4)=C(3)] **(3)**

Table 2 again shows the output from an OLS time series regression model fitted to the data plotted in Figure 1, but this time with an additional term, *MA(4)*, which represents a seasonal adjustment factor (moving average) to allow for potential quarterly serial dependence in the residuals.

The coefficients in Table 2 are substituted into equation (3) to give the following representation of the relationship between trial pending caseload and median delay.

TRIAL_MEDDELAY = 0.073*TRIALS(-4) + 156.303
+ [MA(4)=-0.971] **(4)**

From equation (4), we now see that median delay during the period January 1997 to December 2000, may be modelled as a base (constant) level of 156 days, plus 7.3 per cent of all cases listed for trial four quarters previously. The inclusion of a moving average term (which includes in the model a weighted average of recent random disturbances) has changed our specification of the relationship. By assuming that there is some other seasonal process at work, we in fact decrease our reliance on backlog as an indicator of delay.

Table 2 shows that the fit of the model is much improved with the inclusion of the MA term (R-squared = 0.80) and again there are no problems with the underlying assumptions. The error terms from the

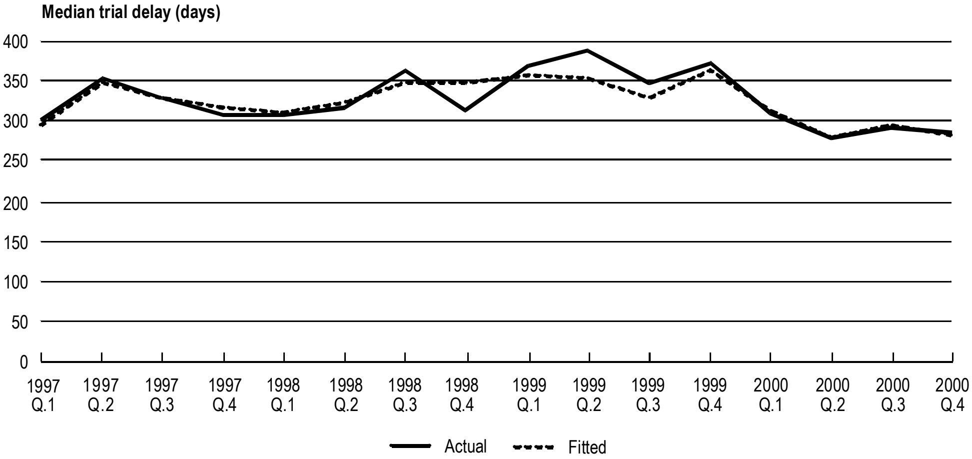
Table 2: OLS Regression output for model B:
Predicting trial delay from pending caseload (lagged) and MA term

Dependent Variable: TRIAL_MEDDELAY
Method: Least Squares

Variable	*Coefficient*	*Standard error*	*t-statistic*	*Probability*
TRIALS(-4)	0.0731	0.0229	3.1927	0.0071
Constant	156.3027	54.2471	2.8813	0.0129

Model diagnostics	
R-squared	0.8011
F-statistic	26.1743
Prob(F-statistic)	0.0000
Prob (Jarque-Bera)	0.3105
Durbin-Watson statistic	1.8226

Figure 3: Actual versus fitted median delay - Model B
NSW District Court, 1997 to 2000, by quarter



regression model are again approximately normally distributed, and there is no serial correlation evident.

Figure 3 shows the graph of actual median trial delay compared with that which is modelled by equation (4). Not surprisingly, the quarterly movement is better reproduced by model B.

It should be noted that while the simple models developed here clearly fit the data reasonably well, the usefulness and stability of the specified relationship suffers from the small number of data points which were available to develop the model.[6] Furthermore, the inclusion of the MA term with such a short time series is not entirely desirable, and model B should best be fitted again once more data is gathered.

CONCLUSION

Large delays in the time between committal and finalisation of trials in the NSW District Criminal Court have been a significant problem in the past. While trial delay is now less of a problem, it is nevertheless useful to develop a simple mechanism which will assist in monitoring future levels of trial delay. In practical terms the most useful model is one which is based on a readily measurable factor, such as the current pending caseload, as described here.

The modelling strategy in this paper is therefore a first step towards the development of a planning indicator for delay. The specification of the actual models themselves will benefit from additional data and, potentially, from the inclusion of further variables. A more sophisticated time series analysis could incorporate those interventions (i.e. changes to court processes) which have been introduced to decrease trial delay, as noted in the introduction.

For now, it suffices to observe that, on current evidence, a rise in the pending trial caseload of the District Court is followed 12 months later by an increase in delay. The nature of the relationship is such that a 10 per cent increase in backlog leads to a 6 per cent increase in median delay.

NOTES

1 Weatherburn, D. and Baker, J. (2000), 'Managing Trial Court Delay: An analysis of Trial Case Processing in the NSW District Criminal Court', NSW Bureau of Crime Statistics and Research, Sydney.

2 For a discussion of the recent drop in trial delay see Doak, P. (2001), 'Recent trends in criminal court delay'. Crime and Justice Statistics Bureau Brief, NSW Bureau of Crime Statistics and Research, Sydney.

3 Note that trial median delay refers only to those cases which end as a trial. On the other hand, pending trial caseload includes some cases which may eventually go to sentence only, due to a change in plea. Note also that the pending caseload is measured at the end of each quarter, while median delay refers to trials finalised during the entire three-month period.

4 The elasticity was estimated by regressing log(delay) on log(backlog) using OLS regression. The estimated regression coefficient for the predictor variable is defined as the elasticity.

5 Note that the Ljung-Box Q-statistics are not significant at any lag up to lag 12 for either the residuals or squared residuals, indicating that serial correlation is not present.

6 Initially, a model using months rather than quarters over the five years was trialed, but the data were too volatile and the model residuals therefore contain too much 'noise' or unexplained variation.