



PERS 341

COMMENTS ON THE REPORT PERS 335:  
DETERMINANTS OF BLACK STRIKE  
ACTIVITY IN SOUTH AFRICA, 1976/81  
BY A GALIN

CSIR Special Report PERS 341  
UDC 658.389(680-963)  
Johannesburg, Republic of South Africa, July 1982

NATIONAL INSTITUTE FOR PERSONNEL RESEARCH  
Council for Scientific and Industrial Research

NASIONALE INSTITUUT VIR PERSONEELNAVORSING  
Wetenskaplike en Nywerheidsnavorsingsraad

001.3072068 CSIR NIPR PERS 341

# BIBLIOTEK LIBRARY

## RGN

RAAD VIR  
GEESTESWETENSKAPLIKE  
NAVORSING

## HSRC

HUMAN  
SCIENCES RESEARCH  
COUNCIL



RGN · HSRC



RGN-BIBLIOTEEK  
HSRC LIBRARY

VERVALDATUM/DATE DUE

1983 -01- 3 1		T PERS 335:
1983 -08- 3 0 1983 -08- 2 0		: STRIKE
1986 -12- 0 7		: ICA, 1976/81
1988 -06- 0 4		
02 JUL 1988 30 NOV 1991 16 MAR 1992 03 AUG 1992		

CSIR Special Report PERS 341 (i-v ; 1-13)  
UDC 658.389(680-963)  
Johannesburg, Republic of South Africa, July 1982



1929933866



ISBN 0 7988 2327 5

CSIR Special Report PERS 341

Published by

National Institute for Personnel Research

Council for Scientific and Industrial Research

PO Box 32410

Braamfontein 2017

Republic of South Africa

Printed in the Republic of South Africa by  
National Institute for Personnel Research

## ACKNOWLEDGEMENTS

Dr G K Nelson	Director, NIPR
Mr R F Skawran	Leader Personnel and Organisation Development Group, NIPR
Dr J van Rooyen	Head: Management Studies Division, NIPR
Mrs G Godsell	Human Development Division, NIPR

## SUMMARY

This report is a critical assessment of the use of regression for causal analysis in a study of strikes and also of the interpretations placed on the results of this study.

## OPSOMMING

Hierdie verslag is 'n kritiek van die gebruik van regressie-metodes in die ontleding van oorsake van stakings en ook van die gevolgtrekkings.

## C O N T E N T S

	<u>Page</u>
COMMENTS ON THE METHODOLOGY	1
GENERAL COMMENTS ON THE VARIABLES USED	4
COMMENT ON THE HANDLING OF THE DEPENDENT VARIABLES	5
COMMENT ON THE INDEPENDENT VARIABLES	6
(a) In the Economic Model	6
(b) In the Political Model	8
(c) In the Industrial Relations Model	9
(d) In the Integrative Model	11
COMMENTS ON THE CONCLUSIONS	13

## COMMENTS ON THE METHODOLOGY

The main criticism of this work is that the wrong methodology has been used. Correlation does not necessarily imply causality and even where a causal model implies correlation, multiple regression is an instrument for prediction and is unsuitable for determining the interrelations of causes. Anyone who has worked extensively with multiple regression knows how the values and signs of regression weights can change, and variables drop out, as new variables are introduced or the period over which the regression is calculated is extended or decreased. Thus regression weights are highly unstable and can in no way be regarded as measures of 'pure' impact.

The reason is that the relationship between an independent and the dependent variable includes a unique part and a part shared with other independent variables in complex ways. The computational process balances the different shared relationships to produce the closest overall fit of the regression surface to the data. At times the best fit is obtained by allocating weights which, though maximising some parts of common relationships, are too large for other parts which then have to be offset by allocating negative weights to other variables. For example, when multiple regression is used for job evaluation, variables which have a positive correlation with the criterion variable, sometimes appear in the equation with negative weights. This is the reason for many of the puzzling correlations that have been observed in this study.

A further distorting factor is that if, as is common in multiple regression studies, linear regression has been used, the regression weights in effect represent straight lines run through the actual data curves. This may eliminate the detail which is of great importance for determining causal relationships. For example, the number of strikes declined sharply in 1976 and then remained stationary until 1979 when a change in the Industrial Relations Act led to a large increase in the number of Black workers joining trade unions and the founding of many new unions. This increased



the opportunity for industrial action and the power struggles typical of union growth have produced the greatly increasing number of strikes in the last half of the study. Thus the data curves for strike activity are sharply U-shaped and many of the other variables show a roughly linear rate of increase over the period, a linear approximation to the true relationship is a gross distortion. The analysis in Table 8 shows how this can cause relationships to differ in different parts of the study (the division between low and high union membership coincides with the low point of the strike activity data) and this must apply to other independent variables.

Another difficulty with the method is that there are invariably complex interactions between sociological variables and unless interaction terms have been incorporated in the equation, which has not been done in this study, the interactions are parcelled out among additive combinations of variables which introduces still further distortion.

In practical applications the method has to rely on data for which there are statistical series, and with this limited set of variables, which are not the most important, it attempts to represent the complex factors involved in strike activity by simple mathematical equations. The unquantifiable psychological and sociological variables that are the major determinants such as management-worker relations, the effectiveness of downward and upward communications, how militant union leaders are, the proportions of contract workers in the factory work-force and agricultural conditions in the areas from which they originate, have no part.

The correlation of data over time is a frequent source of spurious correlation, that is variables are positively correlated because each increases or decreases over time for reasons other than a direct causal relationship. This does not affect prediction, but

it creates serious problems for causal analysis. An attempt has been made to overcome this difficulty by correcting for linear regression on time, but this is a blind and inappropriate procedure and likely to remove causal as well as coincidental relationships. Variables with an almost linear trend will be virtually eliminated as appears to be the case with NBWDET. Where the data are curvilinear the effect of the correction is merely to change the position and orientation of the data curve without changing its shape. If these changes result in the adjusted curves lying closer together their intercorrelations and very likely their regression weights will increase. This is probably what has happened in Table 9. LBLEG is an exception as it is a step function. The effect of the correction in this case depends very much on the positioning of the steps; the resulting curve is a saw-toothed curve and coincidental matching of sections can produce various correlations. In Table 9 the effect has been to reduce the three weights. Thus the fact that the regression weights show similar patterns before and after correction for regression on time does not mean that spurious correlations do not exist. In fact, most of the deductions from the different models in this study are invalid because they have been based on spurious correlations.

Another incorrect application of statistical methodology occurs in the tests of significance. The use of moving averages, while a legitimate technique for smoothing out random fluctuations, invalidates the application of the usual statistical tests used in regression. The deviations from the regression surface are no longer independent but are autocorrelated. Further, the test on the multiple correlation has been used as a measure of significance. It is well-known that this is unreliable as the ratio of the mean square due to fitting the curve to the residual mean square may increase when variables which do not make a significant reduction to the residual mean square are added to the equation. Ultimately  $N-1$  variables will produce a perfect fit to  $N$  points even if they

are all uncorrelated with the dependent variable. Reliance on this test has led to the inclusion of variables which logically cannot be determinants of strike activity, for example, RULPAR. It should also be noted that where one model appears superior to another because of the larger percentage of explained variance, this percentage is a function of both the magnitudes of the regression weights and the number of variables used. The Integrative Model shows this clearly. The large number of significant weights make this model appear better than the other models, but on closer examination it is seen that the values of some weights are little different from their values in the other models. They are now significant because the residual variance has been further reduced by using six variables.

#### GENERAL COMMENTS ON THE VARIABLES USED

In order to facilitate understanding there should be no doubt as to the meaning of key words in the report. For example, the report has no definition of the following words: Strike, Employment, Black. It is unclear whether stoppages are included in the strike data. Furthermore, if comparisons are to be made with data from countries overseas it is important that similar definitions of employment are used.

The data used in this report have been collected from a variety of sources. There has been no discussion as to whether the data can be properly integrated. It is also clear that the data cannot be perfectly reliable. The report should contain some comment on the possible effects of unreliability. Obvious problems have not even been mentioned. For example, a number of homelands gained independence during the period of discussion. This would effect the number of employed as well as the number of economically active members of the population.

The report discusses a number of approaches that explain strike activity. The author has seen fit to select particular variables

as representative of each approach. There has been little explanation as to why some of the variables have been included and why others have been left out. One should also note that it is rash to conclude that one approach is better than another on the basis of a few selected variables.

The author has used the number of strikes per 1 000 employees as the dependent variable in the analysis. However, a number of the independent variables have not been standardised in this way. Variables used in the Industrial Relations model include NTRUM -: the number of Black trade unions, NCOM -: the number of work and liaison committees, NBWDET -: the number of Blacks regulated by wage determination and so on. In addition the author has used UNIMEM -: the strength of Black trade unions as measured by the ratio of union members to the economically active Black population. Why the divisor has been changed from the number of employees to the number of economically active, has not been explained.

Lastly, the data have been pushed into a regression sausage machine without checking whether it is internally consistent. Outliers, especially in this sort of data, can have strange effects.

#### COMMENT ON THE HANDLING OF THE DEPENDENT VARIABLES

Criticism can also be levied against the treatment of the dependent variables. While it is correct to standardise for working populations when making international comparisons, it requires care if this is to be done over time, as variation in the standardising variable can introduce spurious fluctuations. Consider what happens to the relation between the standardised frequency of strikes and unemployment. As unemployment increases the standardised value is inflated by the decrease in the divisor, i.e. the number employed, and vice versa. This introduced a spurious positive correlation between unemployment and the standardised frequency.

Further, the lag of one and half months placed on the strike activity data is an arbitrary adjustment and can modify the original correlations in unpredictable ways. If a delayed interaction is suspected a series of lagged correlations should have been calculated with each independent variable.

#### COMMENT ON THE INDEPENDENT VARIABLES

##### (a) In the Economic Model

The variables used in the study are real average wage (average wage divided by the Consumer Price Index), the cost of living (the Consumer Price Index itself) and the proportion of Black unemployment (the monthly average unemployment divided by the total economically active population). The Consumer Price Index is highly correlated with strike activity, but this is obviously spurious as a real correlation between the two would imply no pause in the demand for wage increases, which is not true. It is the rate of change in the cost of living that is likely to be the important factor.

The relation between real wages and strike activity is very much more complex than suggested in the Report. Strikes would be expected if real wages fell below some unknown value related to workers' expectations. A negative rate of change in real wages could have the same effect. Strikes may also occur if strike activity had raised workers' real wage expectations. However, increases in real wages may occur without strike activity, for instance, after negotiations where there has been no strike, where employers try to pre-empt strike activity, or where they have a deep concern for their employees' welfare. Again, less than half the strikes that take place are for wages, the rest are for a variety of other reasons such as recognition, objections to pension schemes, reinstatement of workers and the like. Thus, while the level of real wages can be the main precipitator of a

strike in a particular situation, the real overall correlation should be small and negative. Over the whole period, the simultaneous increase in real wages and strike activity that occurred would produce a large positive but spurious correlation. However, using the CPI as deflator eliminates the correlation between the CPI and wages, and if both the CPI and real wages are used in the regression the computational process will allocate all the correlation with the dependent variable shared by the CPI and wages to the CPI giving it a large and real wages a smaller weight. This has occurred.

The third variable, percentage unemployment, is one of the most difficult to study for reasons of definition and enumeration. For example, it is difficult to know whether a man engaged in rural agriculture while waiting for another work contract is unemployed or not. Also it is known that Black urban populations are under-enumerated and that a large proportion of those missed are unemployed. Further, the impact of unemployment varies - there is a perpetual shortage of skilled labour and none of unskilled. In the semi-skilled field the main constraint is the cost, which is not very great, of training a new worker. The great majority of Black union members are semi- or unskilled, thus the scarcity factor does not work in their favour. In addition, strikes often have a strong emotional content which appears to override the fear of unemployment, and which in many cases union leaders are either reluctant to or do not have sufficient authority to control. In such a complex situation merely considering the correlation between unemployment and strike activity overlooks extremely important modifying influences.

The comment that "... it seems that the economic model may best explain the extent of Black involvement in strikes and less the decision to ...." exemplifies the rationalising that is characteristic of this Report. It is difficult to see why if the cost of living is a cause of decisions to strike, it is a stronger cause

of the numbers involved. There is in fact no explanation, the process has merely indicated relationships between sets of numbers which may or may not have causal origins and which are in most cases statistically not significant.

(b) In the Political Model

This section illustrates the errors that can be perpetrated by blindly transferring a technique into a new environment.

Black unions have no representation in parliament, have no pressure groups and are generally hostile to the political system. The official policy towards trade unions is one of self-government tempered by considerations of security. There is thus virtually no interaction between labour and the political system. The correlations of ELEC, CABSIZ, PMINS, RULPAR with strike activity are therefore purely computational artefacts.

There is no reason to assume that different laws will have a linear effect in predicting various strike variables. The LBLEG variable should have been split into 3 separate variables. The construction of the variable LBLEG also reveals little understanding of recent labour legislation. Only one of these acts, the Industrial Conciliation Amendment Act, because it extended freedom of association to all employees, has any connection with the increase in strike activity. The Black Labour Relations Regulation Act allowed very limited strike action, the Black Labour Regulations Amendment Act gave negotiating powers to liaison committees, and the Black Labour Act was merely a consolidation of influx control regulations. To illustrate how meaningless are the results obtained for this variable, four other acts of the same years, but totally unrelated to labour matters, could have been substituted and identical values would have been obtained for the regression weights.

The apparently greater explanatory power of the political model is thus a mirage. The larger percentage of the variance explained is, as explained earlier, due largely to the use of five variables instead of three.

It is correct to say that political variables cannot be ignored, but multiple regression is not the way to handle them. Such variables change the environment and thus demarcate quite different periods. The results of analyses from one period should not be assumed to be directly applicable to another, as shown by the results in Table 8.

(c) In the Industrial Relations Model

Somewhat different explanations than those given in the report can be provided for the results obtained in this section. The number of works and liaison committees increased steadily over the period of the study as employers, becoming aware of the increasing possibility of strikes, set up such committees to improve communications with their workers. The greatest growth in these committees took place in 1974/5; a period of very few strikes. This period is outside the study. Sometimes these committees have averted strikes, sometimes not, and some unions are hostile to them and regard them as an obstacle. It is therefore difficult to see any real connection between strike activity and the existence of such committees. However, the parallel growth of strike activity and the number of committees has resulted in a high correlation and large regression weights. Had the period of study been extended backward to cover the 1973 strikes, the results would have been different as it would then have included periods of much strike activity and few committees and low strike activity and a rapidly increasing number of committees. This also applies to the variables NTRUN and UNIMEM.

The number of unions and membership have also grown very rapidly, but as some unions tend to call fewer strikes and an appreciable



number of Black workers have joined the largest, most powerful and well-established unions which are responsible for very few strikes, the connection between striking and union membership is not as direct as it might seem. Also what seems to happen in many cases is that when membership is still small a union instigates a strike in order to impress the work force with its power. When the union has established itself and been recognised by the employer strikes are less frequent although further striking, ostensibly for higher wages or better conditions, may be precipitated by the appearance of a rival union. Thus there are important differences in policies and tactics among trade unions with Black membership and any study that does not take this into account can have little practical value. Further, factory membership of Black unions is roughly only 20% to 40% although there are cases of much higher percentages. There is some resistance to paying union dues and Black unions do not find it as easy to recruit members as might be expected since workers find they enjoy the benefits unions have gained without being members and that they gain benefits for themselves without union assistance. This last was very apparent in 1973 when there was a large amount of strike activity but union membership was virtually non-existent. Complicating the relationship still further is that when a strike is called the number striking always greatly exceeds union membership because nearly all the Black factory work force will support the strike.

It happens that the correlation between membership and the number of strikes is also large. To balance the inclusion of the variables NCOM and UNIMEM which have large weights, the computational process has given NTRUN a negative weight. On the other hand since the correlations between membership and the other strike activity variables (workers involved and days lost) are somewhat lower, the process has shared out what is left of the common relationship with the dependent variable fairly equally between UNIMEM and NTRUN so both have fairly small but equal weights.

In the case of INCAG (the number of Black workers covered by industrial council agreements) the most simple explanation is that either economic growth or changes in the scope of agreements have caused this figure to increase in parallel with strike activity. It is difficult to see any causal connection as only a small proportion of the number of Black workers covered by the agreements are members of trade unions, hence an increase in INCAG will not necessarily result in an increase in union membership and so possibly strike activity. Another factor is that an appreciable proportion of the unionised Black workers covered by these agreements belong to the older established unions or to unions affiliated to them and these unions have a record of a low frequency of strikes. Among the other unions drawing their membership from the same group of workers are some strongly opposed to the Industrial Council system and these use strike tactics to set up their own agreements.

The correlation between strike activity and the number of workers covered by Wage Determinations is obviously spurious as the determinations apply to workers which by nature of the work or its location are difficult to organise, and the numbers have risen in parallel with strike frequency, not fallen as the regression weight suggests. This is another example of what was pointed out earlier about the behaviour of regression weights. To come to any conclusion on the influence of determinations, it would be necessary to analyse the strike data of those subject to determinations separately.

(d) In the Integrative Model

The intention with the Integrative Model is to obtain a better insight into the interactions between and the relative importance of the variables that have shown up in the preceding models. The conclusions regarding correlation between strike activity and prosperity while having prima facie validity are unfounded as the study does not cover a period in which there is simultaneously

a government policy of allowing virtually unhindered freedom of association in trade unions for Black workers and a recession. The comment on the connection between strike activity and Black workers becoming more aware of their power is certainly correct, but the latter unfortunately is not measurable and thus cannot be used as a determinant.

The explanation of the relationship between unemployment and strike activity may be valid for the well-established trade unions of Europe or America, but Black trade unionism is going through a period of extremely rapid growth and many of the natural controls on strike activity are ineffective. This is revealed by the course of events after the end of the study when strike activity has reached even higher levels although the country has entered a recession.

A possible explanation for the results in Table 8 showing the relations between unemployment and strike frequency in periods of low and high union membership is that membership would have been low in the period 1976-1979 when the number of strikes first declined sharply then slowly increased giving a downward trend. Employment showed a similar trend and unemployment should have had the opposite trend producing a negative correlation with the number of strikes. The number of unemployed had to be estimated for 1976/7 and, as earlier figures are educated guesses, it is highly unlikely that they would be consistent with those from the Current Population Survey. If they had been over-estimated they would have produced a downward trend in unemployment and so a positive correlation with strike activity. The large negative correlation for the later period appears to confirm the connection between strike activity and prosperity, but see the final remark in the preceding paragraph. The conclusions drawn from Table 8 regarding the influence of strong and weak unions are thus invalid.

Other invalid conclusions from this section are the following. As mentioned earlier, the increasing number of works and liaison

committees does not create an ability to organise strikes on a larger scale, and, the connection between union membership and strike activity is anything but direct.

The conclusion that "the number of Blacks subjected to wage determination, the greater the reduction in their tendency to strike" has also been dealt with in the previous section.

#### COMMENTS ON THE CONCLUSIONS

There are few comments on the statements in this section as for the most part they are not derived from the preceding analysis. There are two exceptions:

The first are the references to the behaviour of weak and strong Black unions. These statements would appear to be true of the established White and Mixed unions, but there are not as yet any Black unions who are as large and well organised, and recent events have shown an increasing number of strikes by the larger Black unions over discharges for redundancy.

The second is the statement that "multivariate quantitative analysis is essential in order to gain insight into the pattern and development of Black strike activity." If by multivariate quantitative analysis is meant multiple regression, then the preceding comments have shown that the method is unsuitable and has led to incorrect conclusions.

