

# The striking Netherlands: time series analysis and models of socio-economic development and labour disputes, 1850-1995

Velden, Sjaak van der; Doorn, Peter

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

# The Striking Netherlands: Time Series Analysis and Models of Socio-economic Development and Labour Disputes, 1850-1995

*Sjaak van der Velden & Peter Doorn*[*]

**Abstract:** In the past 150 years, about 15.000 strikes took place in The Netherlands. This article gives an overview of the trends in aggregate strike activity from 1850-1995. A composite index is constructed, which reflects the number of strikes, the number of strikers and number of strike days. On the basis of the literature, a theoretical model is constructed which seeks to explain variations in strike activity over time. In order to test the model, the properties of the time series are inspected. Despite a number of shortcomings in the data, an attempt is made to test the model using correlation and regression analysis. The theoretical model can however not be substantiated by the empirical data. The strike activity appears to be an almost random series (after log-transforming and removing the trend from the data). It is suggested that catastrophe theory and non-linear or chaotic models cast light on the patterns observed, but the series is considered too short an empirical testing.

## 1. Introduction

### 1.1. Purpose of the research

Although many books and articles have been written about strikes in the Netherlands, an overview of the subject was lacking until recently. Recently, an extensive research was finished, in which the available information on almost

* Address all communications to: Sjaak van der Velden, International Institute for Social History (IISG), Cruquiusweg 31, 1019 AT Amsterdam, E-mail: svv@iisg.nl. Peter Doorn, Netherlands Institute for Scientific Information Services (NIWI), PO Box 95110, 1090 HC Amsterdam, The Netherlands. Tel. 31 (20) 4628606, E-mail: peter.doorn@niwi.knaw.nl.

15.000 strikes that took place in the Netherlands between 1830 and 1995 was compiled on the basis of books, magazines, newspapers, archives, etc.[2] This article is based on the data related to strikes for the period 1850-1995. We will try to describe the patterns of aggregate strike behaviour over the past century and a half, and to investigate the complex interrelations between economic, political and institutional developments, and the willingness of workers to strike.

A strike is the temporary stoppage of work by employees in order to enforce certain demands. It was noticed that until the end of the nineteenth century Dutch people often used the English word 'strike' or French '*grève*'. Only slowly the word '*staking*' was adopted, although as far back as the seventeenth century the Dutch had used the very Dutch word '*uitscheiding*' (cessation). The fact that this word was not used again in modern capitalism illustrates the decline of the Dutch economy during the eighteenth century. Different forms and goals of a strike exist. Contrary to most existing research 'lock-outs' are not included in this research, because these are not regarded as actions of workers[3], whereas, on the other hand, 'occupations of companies' are treated as a specific form of a strike.

Because wage-earners sell their ability to work, they try to get the best price possible. At the same time they want to work under the most pleasant working conditions. These are the two most important strike reasons, but because the state has gained increasing control over both, more and more strikes are directed against the state. Besides these, there are also altruistic reasons, e.g. support of sacked colleagues, the unions or fellow-strikers.

Whereas in the nineteenth century a strike was often begun by a group of workers, roaming the streets, singing and waving a red banner, at the end of the twentieth century most strikes are triggered off by a union decree. These two opposites should not lead to the assumption that there is a linear development, because a constant struggle exists between the 'workers' movement' and the 'movement of the workers'. Every time when the unions seem to be in control of the workers, there occurs an outburst of wildcat strikes. On the other hand such an outburst can lead to a radicalisation of the unions.

Kerr and Siegel suggested that the workers most prone to strike are those who do unpleasant work and live in an isolated mass. In the Dutch situation this only goes for nineteenth century navvies, peat-cutters and workers in governmental employment programs (*Werkverschaffing* and DUW). For the rest which workers are more strike-prone turns out to be dependent on the situation

---

[2] Sjaak van der Velden, *Stakingen in Nederland: arbeidersstrijd 1830-1995* (Amsterdam, 2000). With CD-ROM.

[3] In 1993 labour statisticians of the ILO agreed on taking strikes and lockouts separately in statistics. See: International Labour Organization, The Fifteenth International Conference of Labour Statisticians, *Resolution concerning statistics of strikes, lock-outs and other action due to labour disputes* (January 1993).

of the labour market or their importance in the technical production. However, the differences between groups of workers tend to diminish with time. At the end of the twentieth century we envision police-officers, bank-clerks and teachers on strike. A century ago this was unthinkable.

In the Netherlands in the early 1920's the average number of strikes exceeded 300 per year while in the 1990's the average was only 44. Looking at the number of strikers the impression is the reverse: in the 1920's 47.000 whereas in the years 1990-95 we counted more than 135.000 workers per year on strike. In order to measure and investigate aggregate strike activity, it is necessary to construct a composite index. There are three or four indicators for strike-activity: frequency (the number of strikes and number of affected companies), intensity (the number of strikers) and duration (the number of working-days lost due to strikes). Strike-activity is measured as an average of the indicators related to the national totals. The construction of the index is discussed in more detail in paragraph 2.2.

The index shows that at the end of the period of our research strike-activity reached higher peaks than ever before. Although the frequency and duration declined, this was off-set by for the enormous growth in the number of strikers. During fewer and shorter strikes the modern wage-earners reach very high levels of strike-activity because of their vast numbers.

In this article we try to explain the fluctuations in strike activity. Many explanations are in circulation and some authors emphasise economic factors while others point out that the fluctuations are caused by political or institutional developments. In the economic tradition we composed a model in which thirteen independent variables play a part over the years 1850-1995 (paragraph 2). To test the model we applied correlation and multiple regression-analysis. Before doing so, we first had to take into account the statistical particularities of the time-series data available (paragraph 3 and 4).

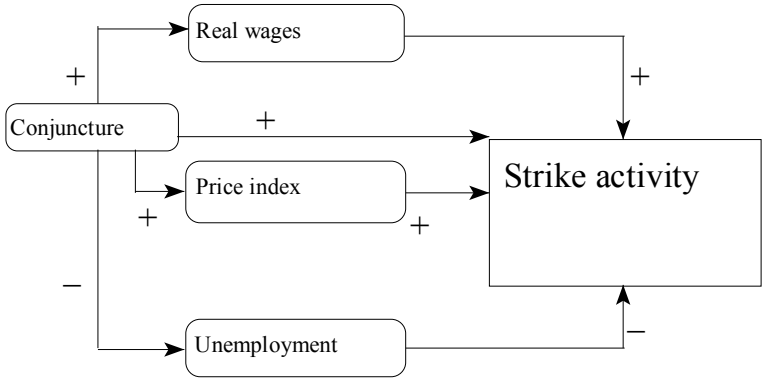
# 2. A theoretical model of strike activity

## 2.1. Presentation of the model

Is it possible to explain strike behaviour economically, politically and from the development of unionisation? Many authors have tried to make such explanations and we will try to fit these into a model. It is commonly accepted that workers are more inclined to strike when the economic situation is booming. It is then that labour can get a bigger piece of the cake. The reverse is supposed to be true during a slump; strike activity declines because the chance of winning is small.

To test this general idea it is necessary to select proper variables. Variables that not only represent the general state of the conjuncture (e.g. the Dow Jones index), but also the direct results for working class life. In accordance with existing research, Dutch and international alike, we selected: unemployment, real wages and the price-index for consumer goods. Besides these three we also took two general variables: the development of National Income per capita and the more subjective vision on the conjuncture. So we selected five independent variables in order to try to explain the development of strike activity, the dependent variable (see Model 1).

Model 1. Five economic variables explaining strike activity



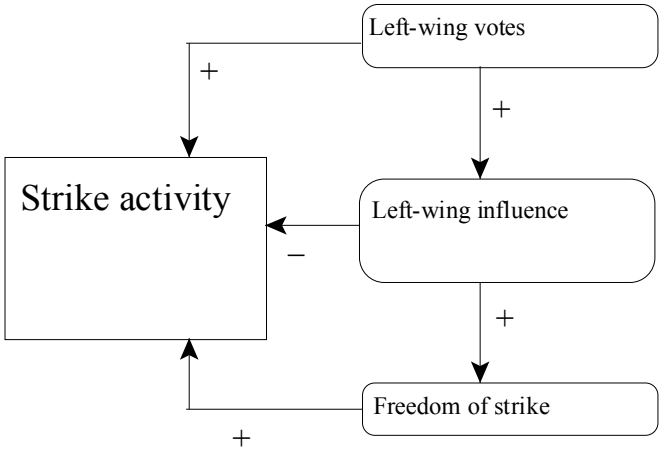

The expectations in the model are as follows. Next to the arrows we find the expected direction of the correlation: '+' means that the expected correlation is positive, which means that the dependent variable is supposed to rise if the independent variable rises and vice versa; '—' means a negative correlation: if one variable goes up, the other goes down.

Economic factors are of course not the only variables explaining strike behaviour and therefore we also created a model with political and legal factors as independent variables. In this sub-model we include the percentage of votes for left-wing parties during parliamentary elections and the influence these parties had in the respective governments. This influence is measured by calculating the percentages of ministers supplied by the left-wing parties. The third variable is the freedom of strike connected with the socio-economic measures taken by the governments.

We expect first of all the left-wing influence in the government to rise with growing numbers of left-wing votes. Second the leftist government is supposed to enlarge the freedom of strike which in turn should lead to a growth in strike activity. Thirdly it is plausible to expect that workers are more eager to strike in times of more powerful left-wing parties, while on the other hand unions will

be more reluctant to engage in strike-activities when left-wing parties are represented in the government. After all, the unions for a great part share a common history with the social-democratic party. The above expectations are represented in Model 2.

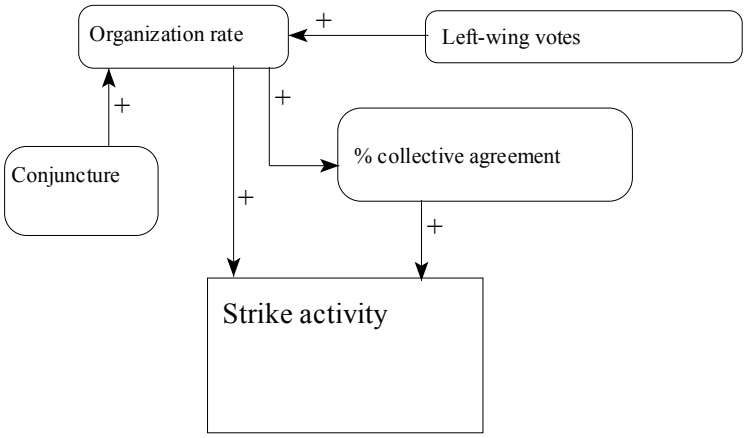
Model 2. Political variables explaining strike activity



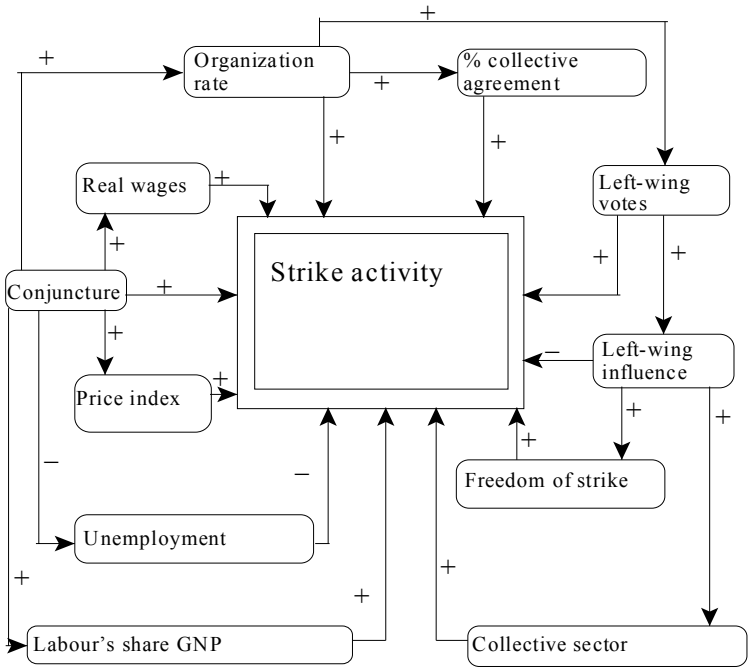

As a result of the considerations we end up with two conflicting expectations. More votes for left-wing parties can correspond with more strike-activity, but because of the influence of labour-unions these same left votes can also correspond tot less strike-activity. This brings us to the over-all influence of unions on strike behaviour, the third part of our model. It is agreed that unions flourish in times of economic booms and we can measure this by looking at the organisation rate (union-members as a percentage of the dependent labour force). In following the Dutch labour-historian Th. Van Tijn we also take the percentage of workers working under collective agreement as a measure of the success of unions. This leads us to the third part of our model (Model 3, next page).

This third part contains two elements, which connect this sub-model with the other two sub-models. If we actually connect them we end up with one extended and complicated model to which we can add two other variables connecting the economic and political parts. These are the collective sector (government expenditure and social insurances) and labour's share in National Income (see Model 4, next page).

Model 3. Variables reflecting influence of labour unions explaining strike activity




Model 4. Strike Activity Model composed of the variables in Models 1-3



## 2.2. Measuring strike activity

So far we have chosen the dependent and independent variables for constructing and testing a model on strike behaviour. More is to be said about how these variables are measured. First we will look at strike activity. Most international statistics concerning strikes include not only strikes but also lock-outs. If one wants to calculate how much working-time is lost by labour disputes it makes sense to incorporate them. We are looking for an explanation for the inclination to strike by wage-earners, so for us it does not make sense to incorporate lock-outs.

Apart from the problem of lock-outs there is the fact that most official statistics include workers not directly involved in strike activities too. These workers are the ones that are unable to work due to the strike. Again, in researching strike behaviour this figure must be omitted. After eliminating lock-outs and workers not directly involved, we were left with 14.500 strikes. For each year we have four indicators: number of strikes (N), total number of strikers (S[4]), number of affected companies (C) and the total of working-days the workers were on strike (D)[5].

In 1966 P. Galambos and E.W. Evans tried to connect three of these indicators in one index in order to compare in the incidence of strikes in several industries[6]. In this index they related S and D to the total labour force (W, all workers). Good old Knowles in the same Bulletin wondered why Galambos and Evans did not also relate N to W. We from our part wonder why one should relate strikes and lost working-days to people at all. People ought to be related to people, days to days and so forth. We therefore adjusted the formula as follows:

$$I_t = 100 \times \{[ ( N_t + C_t \times s_t) + S_t / S_b + D_t / D_b] / [ (N_b + C_b \times s_b) + W_t / W_b + V_t / V_b ]\}$$

In this formula, s stands for the number of strikers per company[7] and V for the total labour volume (the number of days to be worked by all workers) in a given year. The subscripts indicate this given year ( $_t$ ) and a basis year ( $_b$) that is to be chosen by the researcher.

By using this formula we could not solve one specific problem: the indicators are not weighted. Weighting them can only be subjective: an adherent of peaceful industrial relations will place more emphasis on the diminishing num-

---

[4] Workers that struck more than once during one year were nevertheless counted only once.

[5] This not necessarily equals the working-time lost for the employers, because the latest can go on producing during the strike with the help of blacklegs.

[6] P. Galambos and E.W. Evans, Work-stoppages in the United Kingdom, 1951-1964: a quantitative study, in: *Bulletin of the Oxford University Institute of economics & statistics*, vol. 28, 1966, p. 33-55

[7] C x s S, because we counted workers that struck more often only once in S

ber of strikes while a Marxist may point to the growing number of strikers. We therefore acquiesce in it that we are not able to weight the indicators.

Objections may also be raised that the first component of the formula is of another magnitude than the second and third components. We have experimented with several alternatives to overcome this handicap, which however did not offer significantly different outcomes.

Another disadvantage is the intransparency of the formula. Our answer to that is given by the correlations between the index on the one hand and the original indicators on the other. The correlations between the index and frequency, intensity and duration of the strike-movement are reflected in Table 1. Given these correlations the index seems to give a fair reflection of the original indicators.

Table 1. Correlations between the strike-index and the original strike-indicators

| Index | frequency (number of strikes) | intensity (number of strikers) | duration (number of working days lost due to strikes) | |
|---|---|---|---|---|
| 1 | 0.7 | 0.95 | 0.86 | index |
| | 1 | 0.64 | 0.81 | frequency |
| | | 1 | 0.88 | intensity |
| | | | 1 | duration |

## 2.3. The independent variables

In our model we have eleven independent variables. Most of them need no further explaining: the development of real wages and prices are taken as indexes. Unemployment is measured as a percentage of the dependent labour force. The votes for left-wing parties are also taken as a percentage and the same goes for the influence of these parties in the government. The organisation rate is again a percentage and unsurprisingly the number of workers under collective agreement is a percentage. We need not say anything about the collective sector and labours share in National Income, but there are two variables that cause more problems.

Conjuncture is measured by taking the annual growth rate of the National Income against constant prices, but also as a ‘mirror of the economy’. Economists and historians over the years have judged the economic situation as positive or negative and these judgements gives us a binary list of ones and zeros from 1857 on. This list is the second indicator for the conjuncture.

The other variable with a high rate of subjectivity is ‘freedom to strike’. To enable us to computerise the model we translated the legal development and the development of economic policy into six ordinal data. The extreme values of

these data are: (1) total freedom to strike and a very reluctant government (1873-1902 and 1982-1995) and (6) Prohibition to strike and during some periods even a capital offence and state regulation of labour relations (1940-1945).

## 3. Statistical properties of time-series data related to strikes

Before we can test the hypothetical model developed above, we must inspect the statistical properties of the data available. If we want to estimate the relationships between the variables, we must take into consideration that we are dealing with time series. As the variables that play a role in the model can be considered as interval/ratio variables, correlation and regression analysis may serve as appropriate techniques to measure relationships among them[8]. The strike index is the dependent variable in the model. When using correlation and regression for data that are not time series, we have to be aware of problems connected to the normality of the data (e.g. the effect of outliers) and of multicollinearity (independent variables that statistically explain the same variance in the dependent variable; such variables display a high degree of mutual correlation).

When applying correlation and regression analysis to time-series data, the effects of trend, autocorrelation, and time lags may also be of influence. In addition to this, not all data are available for all years. In the analysis presented here, we concentrate on the period 1850-1995, a total of 146 years. For only 63 years (mainly after World War I), not one variable is missing. In earlier quantitative studies, which seek to explain strike activity in The Netherlands, almost nothing is found on these statistical complications. We think, however, that without taking into account the statistical properties of the time series data, the interpretation of the results is virtually impossible.

[8] The variable representing economic boom or depression is a binary variable, which may be treated as an interval variable. The variable representing institutional strike restrictions is an ordinal variable, which we will however treat as interval. The percentage of votes for left-wing parties and of cabinet ministers from left-wing parties change only at election years and have therefore limited numbers of values.

Table 2. Independent variables used to explain strikes in the Netherlands, 1850-1995

| Variable | Explanation |
|---|---|
| Nominal wage | Index number of nominal wages, 1990=100 |
| Real wage | Index number of real wages, 1990=100 |
| Price index | Index number of consumer prices, 1990=100 |
| Unemployment | Unemployment as percentage of dependent labour force |
| Left government | Percentage of cabinet ministers from left-wing parties |
| Freedom of strike | Institutional strike restrictions |
| Union membership | Union members as percentage of dependent labour force |
| Collective agreement | Percentage of employees under a collective labour agreement |
| Wage quota | Labour income quota as percentage of national income |
| Collective sector | Collective sector as percentage of national income |
| National income | Index of national income per capita, 1990=100 |
| Economic level | Economic boom (1) or depression (0) |
| Left-wing votes | Percentage of votes for left-wing parties |

## 3.1. Trends, cycles and the effect of outliers

Two time series that display a trend will always be correlated. The stronger the trend is in comparison to other components of the time series (such as cyclical and irregular effects), the stronger the correlation will be. Two series with either a positive or a negative trend correlate positively, two series with opposite trends correlate negatively. A trend that we observe may however be distorted by outliers and/or extreme values.

All variables to be taken into consideration in the models to explain strike activity, except two, show a positive trend. The variable representing the strike restrictions has a negative trend, and for unemployment the trend is virtually absent. Therefore, practically all simple (bivariate) correlations are influenced by the trend effect.

A trend in a time series can be stochastic or deterministic. In a deterministic trend model, the average level of a series grows or declines according to a specific function of time. Many time series do not exhibit the stability of a deterministic trend, but have trends with variable growth. In a stochastic model the parameters randomly change over time.[9]

[9] N.R. Farnum & L.W. Stanton, *Quantitative forecasting methods* (Boston, 1989), 451-452.

The pattern of strike activity can be seen as transitory fluctuations around a deterministic trend line; alternatively, the trend can be a non-stationary stochastic process akin to a random walk. Such processes contain a unit root in their autoregressive representation and require first differencing for stationarity.

The unit root test, originally developed by Dickey and Fuller, takes as its null hypothesis a unit root in the autoregressive representation of the time series.[10] The first differences then follow an autoregressive process of known and finite order. The alternative hypothesis is that the largest autoregressive root is smaller than one, so the series is a stationary autoregressive process added to a deterministic trend.

Research on the development of the American GDP since 1870 shows that the presence of outliers can lead to the spurious rejection of the unit root hypothesis: 'Outliers that alter the level of output for only one period reliably trigger false rejections of the unit root hypothesis when it is true and signal the presence of permanent shifts in trend that did not occur'.[11] Nelson and Murray go so far as to propose to move away from the unit root issue completely. They call the method fragile, stating that 'determinism is not an hypothesis that is supported either in economic theory or in history [...] A statistical model implies a conditional distribution of future observations given the data, not simply an accounting of past events.'

Not all variables to be correlated can be regarded as normally distributed. Some variables display a high skewness or kurtosis, which is partly caused by outliers and/or extreme values.[12] Such values tend to have a major effect on trends in time series. The strike index itself appears to be the most problematic in this respect (Figure 1). The kurtosis of 71.7 of the strike index indicates that the series is extremely peaked. The skewness of 7.8 shows that there are many years where the index is lower than the average. There are 7 extreme values and 13 outliers (all positive) in the strike index (this amounts to 14% of the 146 values). The most extreme value is found in the year 1970, where the strike index is over 1500; the other extremes are over 200, the first of which occurs in 1943; the remaining five are all after the year 1979. The 13 outliers are all above 70 and they all occur since the 1920's.In short, , the pattern in the series is heavily influenced by a few years, especially after World War II, when the strike index was extremely high).

---

[10] D.A. Dickey & W.A. Fuller, 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association*, 74 (1979), 427-431.

[11] Ch.R. Nelson & Ch.J. Murray, 'The Uncertain Trend in U.S. GDP', *Working Papers in Economics* (Research paper, University of Washington), January 21 (1997), 28 pp. Electronic publication at http://netec.wustl.edu/WoPEc.html

[12] Outliers are defined here as values of between 1,5 and 3 times the interquartile distance away from the mean; extremes have values of more than 3 times this distance.

Figure 1: Strike index (original data and regression line), 1850-1995

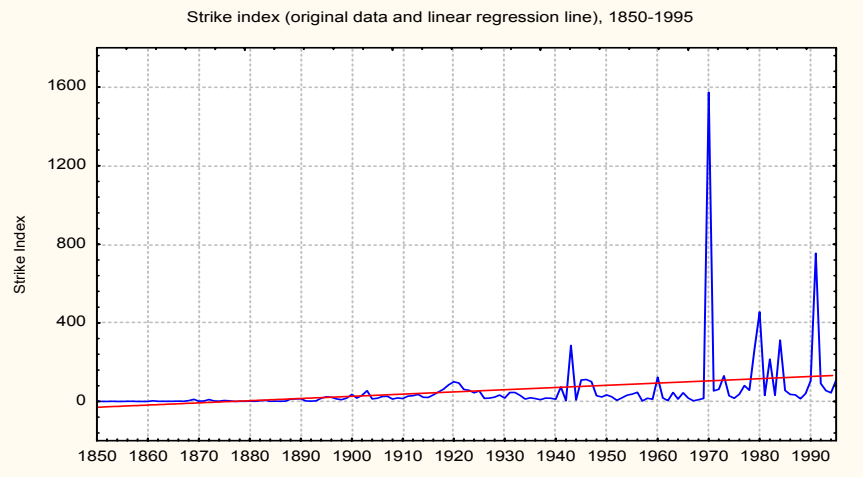
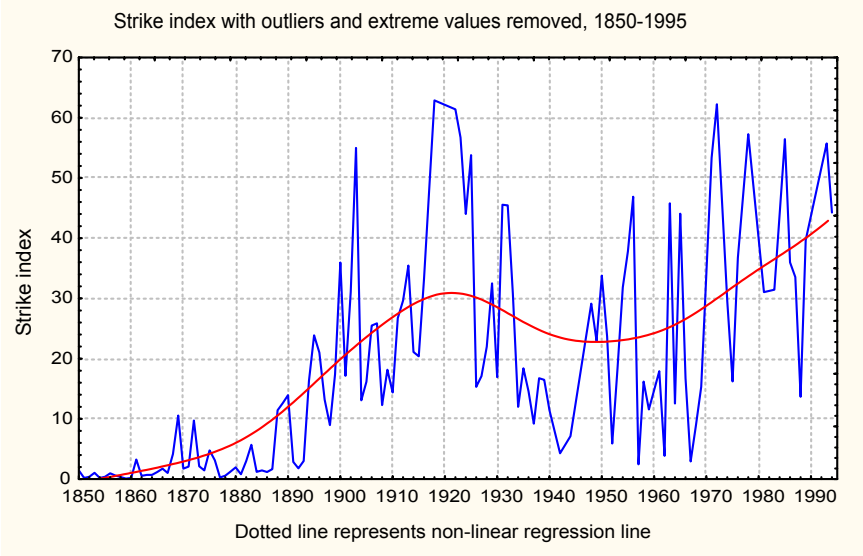



Figure 2: Strike index with outliers and extreme values removed, 1850-1995




When we remove outliers and extremes, the overall impression of the strike index time series changes considerably (Figure 2). The characteristics are:

- there is still a clear overall positive trend.
- there are indications that the best line to fit the data is not a linear regression line, but a polynomial curve with a local maximum around 1920 and a local minimum around 1950.

Deleting the outliers would severely restrict the period for which we can construct models of strike activity: before 1911 many independent data are missing; after 1940 many outliers and extreme values in the strike index occur. An analysis of the period 1911-1940 only would remain.

On the basis of the above discussion, we decided to remove the trend and to reduce the effect of the outliers in the following ways. One possibility to remove a trend is to calculate the trend line and then to subtract the trend from the data. The disadvantage of this transformation is that deviations from the trend tend to be smaller where the original values are smaller. When a trend is positive, this phenomenon means that deviations in the early years will be of less influence than those in later years. Another widely used detrending technique is to take the so-called 'first differences'. The method of first differences means that we calculate the change from year to year. This is the method we use in this study.

We decided to reduce the effect of the extremes and outliers by taking the logarithm of the original data instead of removing them manually. The skewness and kurtosis are then reduced to around zero (-0,37 and -0,11 respectively). After taking the logarithm of the strike index, the positive trend in the data becomes clearer. A linear trend in the logarithmic series gives a good fit, but a polynomial represents the data even better. The curve indicates a slowdown in the strike intensity from the 1930s to the 1960s (Figure 3).

Except for the trend, there are also a number of variables that display a more or less cyclical pattern. This is most clear in the variable representing the economic progress (influenced by business cycles) and is less clear in the variables representing votes for left-wing parties, membership of the unions and presence of left-wing parties in the government. Trend and/or cycle, either together or apart, are present in all variables eligible for the model. All variables have some form of autocorrelation (meaning that an observation in a given year is influenced by the previous year). The only exception is the (untransformed) strike index: here the autocorrelations for the first fifteen time lags are negligible (the highest autocorrelation is 0,24 at 10 lags). After taking the logarithm, the autocorrelation is substantial (gradually decreasing from 0,74 at lag 1 to 0,40 at lag 15; see Figure 4).

Figure 3: Lognormal strike index and non-linear regression line (1850-1995)

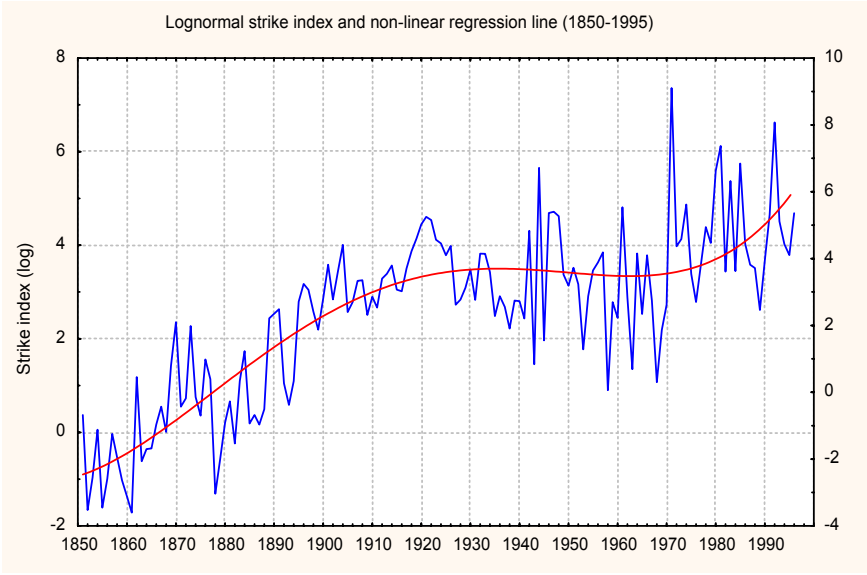
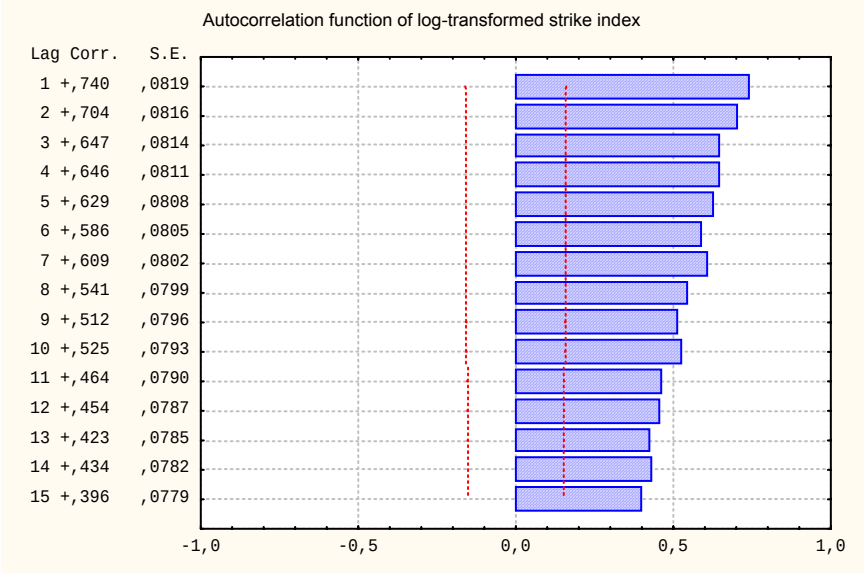


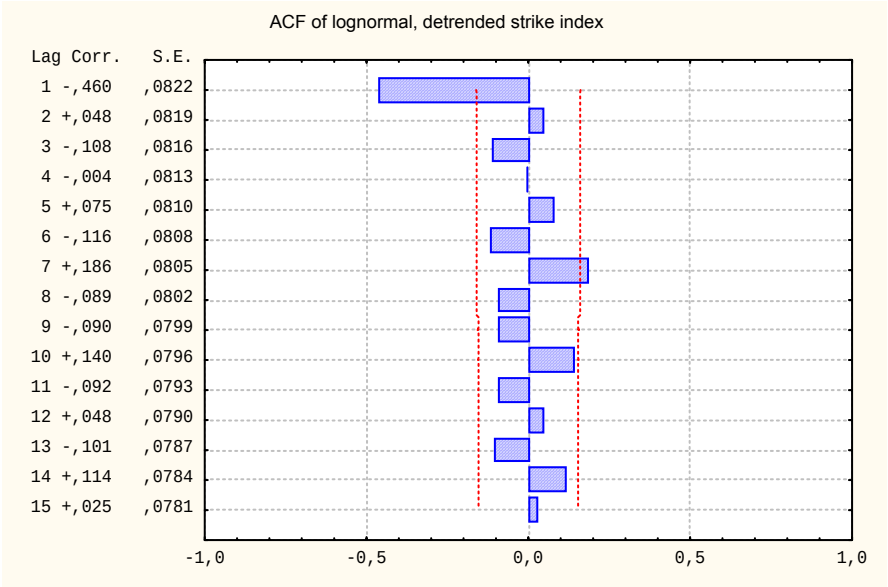

Figure 4. Autocorrelation function of log-transformed strike index



## 3.2. Autocorrelation and spectral analysis

Although the number of observations is limited, it is interesting to inspect the characteristics of the time series of the strike index in its own right. It has already been noted that the strike index series in its original form does not have a distinct statistical pattern. The series is strongly influenced by a number of peaks in 1970, 1991, the first half of the 1980s, and 1943. After taking the logarithm of the series, an upward trend is the most remarkable characteristic (Figure 3). But what remains after we remove the trend from the data by taking the first differences of the log-transformed strike index? Calculating the autocorrelation function (ACF) of the now stationary series can shed light on the fluctuations in the strike activity from year to year. A rather strong negative autocorrelation of -0,46 appears to exist at lag 1: this means that there is a tendency that a year with a high strike index is followed by a year with a low strike index. In a diagram of the (detrended, log-transformed) strike index over time, a rather strong zigzag pattern prevails. The autocorrelations at higher time lags are fluctuating around zero. They are not strong enough to indicate a cyclical pattern (Figure 5).

Figure 5. ACF of log-transformed, detrended strike index

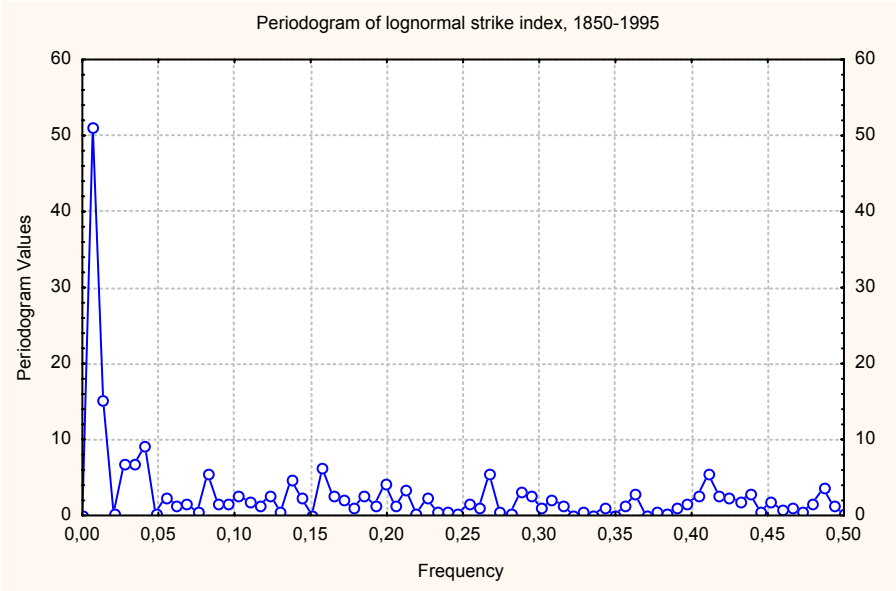



The best indication for the presence (or absence) of cycles in the strike index can be obtained with the help of spectral or Fourier analysis. We must however

keep in mind that the series is relatively short (146 years or observations) to get definite results. For spectral analysis, the series to be analysed must be detrended. If the observations in the series are independent of each other (i.e., there is no periodicity), and they follow the normal distribution, such a time series is also referred to as a white noise series (like the white noise one hears on the radio when tuned in-between stations). A white noise input series will result in periodogram values that follow an exponential distribution.

We carried out the spectral analysis on the (detrended) log-transformed data, since the outliers would too heavily bias the analysis of the original strike index. The impression of the periodogram is indeed one that resembles a (negative) exponential or logarithmic distribution (Figure 6). The highest periodogram peaks and spectral densities appear to occur at the lowest frequencies. The periodogram and spectral density diagram resulting from the Fourier analysis do therefore not indicate any clear cyclical pattern of strike activity.

Figure 6. Periodogram of log-transformed strike index, 1850-1995




## 3.3. Consequences for the model

To summarise, in any analysis to test the relationships between explanatory variables and the strike index, we must take into account the following statistical properties of the time series:

- The strike index has some strong outliers; the influence of these outliers can be reduced significantly by taking the logarithm of the original series.
- The (log-transformed) strike index shows a positive trend; the trend effect can be removed by taking the first differences.
- The strike index has a negative autocorrelation at one time lag, meaning that a zigzag pattern prevails.
- There is no cyclical pattern in the strike index: the detrended, normalised strike index series resembles a series produced by chance (a 'white noise' series).
- Some explanatory variables are mutually correlated and may cause multicollinearity.
- Trend and/or cyclical patterns are present in practically all explanatory variables. The trend can be removed by taking the first differences of these variables.

This means that the prospects of finding a meaningful and statistically significant explanatory model of the behaviour of the strike index series over time by a combination of the independent variables are not very good. Nevertheless, an attempt was made to construct and test a multiple regression model on the basis of the theoretical considerations outlined in paragraph 2.

# 4. Testing the model

## 4.1. Correlation and regression analysis

Before we attempted to construct a multiple regression model, we first carried out simple correlation analysis between the strike index and the thirteen explanatory variables. The directions of the relationships are generally in line with our theoretical expectations. Because of different measurement methods (and length of observed period) for some variables, some discrepancies with the calculations by earlier researchers exist (notably the correlation with unemployment and the price development by Van Dam van Isselt[13]; when we conform to his definitions, the differences are negligible).

[13] E.W. van Dam van Isselt, "Over het verband tusschen werkstakingen, werkloosheid en prijzen", in: *De Economist* 1914, p. 1-9.

Table 3. Correlation of strike index and thirteen explanatory variables

| | | | Correlation with | | | | | |
|---|---|---|---|---|---|---|---|---|
| Independent variable | Data available for years | Expected correlation | Strike index (original series) | | Log-transformed strike index | | First differences of log-transformed strike index | |
| Nominal wage | 1880-1995 | + | 0,26 | ** | 0,44 | ** | 0,04 | |
| Real wage | 1880-1995 | + | 0,30 | ** | 0,49 | ** | 0,22 | ** |
| Price index | 1880-1995 | + | 0,26 | ** | 0,44 | ** | -0,08 | |
| Unemployment | 1911-1995 | - | -0,08 | | -0,03 | | -0,10 | |
| Left government | 1921-1995 | + | 0,07 | | 0,23 | * | -0,08 | |
| Strike restrictions | 1850-1995 | - | -0,09 | | -0,30 | ** | 0,07 | |
| Union membership | 1911-1995 | + | 0,08 | | -0,01 | | -0,15 | |
| Collective agreement | 1911-1981 | - | 0,09 | | -0,06 | | -0,19 | |
| Wage quota | 1921-1995 | + | 0,25 | * | 0,35 | ** | -0,03 | |
| Collective sector | 1921-1994 | + | 0,21 | | 0,27 | * | -0,12 | |
| National income | 1900-1994 | + | 0,28 | ** | 0,35 | ** | 0,20 | |
| Economic level | 1857-1995 | + | 0,06 | | 0,21 | ** | 0,11 | |
| Leftwing votes | 1880-1995 | + | 0,17 | | 0,47 | ** | -0,04 | |

*Note:* * = significant at 95%; ** = significant at 99%. Although the correlations were calculated on the basis of population numbers, we can use the significance levels as an indication for the relative strength of the coefficients.

However, as discussed above, the outcome of the correlation analysis will be influenced by outliers and by the trend effect. We removed both influences in two steps. First we recalculated the correlations after taking the natural logarithm of the strike index. It is remarkable that practically all correlation coefficients are now considerably higher than before. It is easily demonstrated that this is to a large extent due to the trend present in most variables. After removing the trends by taking the first differences of all variables, the correlations are reduced notably. The highest remaining correlation coefficient is that with real wage (0,22). Half of the correlations are 0,10 or less (Table 3).

Given the characteristic of the strike index time series, anything else was hardly to be expected. The chances of finding a multiple regression model explaining a significant part of the variance in the dependent variable are therefore slight. The interrelations between the de-trended independent variables are often much higher than those between independent and dependent variables, so that those relations may furthermore be distorted by multicollinearity.

A variety of regression models have been estimated, but not one of them was satisfactory. Variations attempted were:

- to delete missing values pairwise (per pair of variables all cases are included in the analysis for which data are available) or casewise (all variables included in the model must have non-missing values); in total, no more than between 48 and 62 cases of the total series of 146 years were included in the analysis).
- for the trendless non-interval variables we used the original values as well as the first differences
- to apply (forward) stepwise regression and standard regression techniques.

Although we are confident that the strike index is a fair reflection of the original indicators (frequency, intensity and duration), we nevertheless also tested these indicators. The results were as disappointing as testing the index.

In conclusion, not one model had an explained variance of more than 20%, the regression parameters were barely significant (except for the real wage variable, which already appeared from the correlation analysis), and the results were rather sensitive to the regression technique used.

### 4.2. Breakdowns by periods

Hitherto we have not been able to compute satisfying results for the long run. However, it may be possible that the theoretical insights are true for the short run. Maybe, what cannot be proven for a period of 150 years is valuable for shorter periods.

We therefore broke the whole period down into 20-year periods. For some relations we now calculated much higher correlations, while changing directions partly explain the low values for the whole 150-year period. The highest correlation was R= -0.75 for unemployment and the strike index in the years 1911-1920. More surprising is the positive correlation R= +0.33 during the years 1921-1940, because this outcome is contradictory to most literature on the relationship between strikes and unemployment. The second highest is R= -0,65 for strike restrictions in the 1901-1920 period. The third highest is R= -0.55 for union membership during the sixties and seventies of the twentieth century. A result that is in line with left-wing criticism of union-leaders: while the unionisation rate grew, strike proneness diminished. No higher correlations were found. So, breaking down the whole model in shorter periods also leads us nowhere.

## 5. Conclusions and directions for further research

So far, we have not been able to statistically explain variations in strike activity in The Netherlands between 1850 and 1995. There are various reasons for this

lack of statistical success, a number of which we have already looked into. The theoretical model that was tested, which was discussed in paragraph 2, was derived from the literature on strikes in The Netherlands and abroad. Data has been gathered from a multitude of sources to operationalise the central processes as well as possible. Of course, there were flaws in the data and missing values were a serious problem. Moreover, the statistical properties of the time series data were duly taken into consideration. Correlation and regression analysis did bring along a number of handicaps when applied to the data. Perhaps other (nonparametric or non-linear) techniques could have been used, but it is our contention that those would not have led to different insights.

Is the model on strike activity wrong (or too simplistic)? Are we generally driven by false hope in socio-economic historical research, when we expect to find significant relationships between variables over long periods of time?[14] It is of course so that over a time period of a century and a half, both structural changes and disrupting events take place, that are extremely difficult to disentangle. When looking for structural relationships over time, the implicit postulate is that in the long run, the effects of incidents tend to level each other out. This assumption can however not be proven. The influence of events such as wars and crises may be much more important than the structural effects. The whole meaning of what is structural and what is not may even change over time. Does it make sense to compare, for instance, votes for left-wing parties in the 1990s to those in the 19th century? Are we still talking about the same phenomenon? Statistically the answer is affirmative, but the whole idea of what constitutes a left-wing party has changed over time. And is it realistic to expect that the factors influencing strike activity do so in a consistent way? For instance, the economic policies of governments with left-wing influence have so much changed over time, that it is unlikely that the variable 'government influence of left-wing parties' works out in a similar way over the decades.

Or maybe we are simply expecting too much. As we discussed above, most variables in the strike model (including the strike index itself) displayed a trend. Perhaps discussing which trends co-occur should satisfy us, although then we will never be able to say anything further that can be measured.

Alternatively, striking may be heavily influenced by accidental factors; factors, that cannot be easily linked to broad socio-economic and political structures. The fact that the strike index had practically no cyclical pattern and, except for a negative autocorrelation at lag 1, mainly had the appearance of a white noise series, raises the question how such a pattern emerges. It does not seem likely that strikes take place by chance alone.

[14] Peter Doorn, 'The old and the beautiful: a soap opera about misunderstandings between historians and models", *VGI Cahier* 10 (1998), pp. 11-30.

## 5.1. Explaining the inexplicable: catastrophe theory and chaos models

Decisions to strike take place in stress situations. The many different reasons for strikes to 'break out' are discussed in van der Velden's dissertation. Certain conditions must be met, but there will be a tension between the options to strike or not to strike. In this sense, the outbreak of a strike can be considered as a specific case of a bifurcation model or a so-called cusp-catastrophe.

Zeeman has described the cusp-catastrophe[15] with the help of an example from animal ethology. Rage and fear in animals (and humans) are conflicting factors influencing aggression. When an animal is angry, it will fight; when it is scared, it will flee. But what happens when it is enraged and afraid? It will attack or run away, and a sudden jump (the attack or retreat catastrophe) between both forms of behaviour (originally described by Konrad Lorenz) may occur. Catastrophe theory, later extended to chaos theory, may offer a model for forms of behaviour that show sudden jumps. These models have been applied to describe a variety of phenomena, ranging from turbulence in meteorology to the outbreak and spread of contagious diseases in medicine. In the social sciences, chaotic models serve more as metaphors than as testable theories.

Andreev, Borodkin and Levandovskij attempted to apply chaos theory to strike activity in Russia between 1895 and 1908[16]. Their series of 168 monthly observations is however much too short to offer definite results (according to Nijkamp, over 10,000 observations are necessary to apply chaotic models responsibly)[17]. Our series of 146 observations is even shorter than the Russian series, and it was not attempted to fit non-linear dynamic models to the data. The cusp-catastrophe may be an appropriate model to describe the outbreak of individual strikes, in which dissatisfaction with labour or political conditions competes with the fear of loss of income or job. The pattern observed in the strike series can be seen as the summation of the thousands of strikes that have taken place over the past one and a half century or so. Strikes sometimes spread like the proverbial wildfire, and chaos theory has been applied to describe diffusion processes successfully.

---

[15] E.C. Zeeman, *Catastrophe Theory: selected papers 1972-1977* (Addison-Wesley, 1977), pp. 3-7.

[16] A. Andreev, L. Borodkin & M. Levandovskij, "Using methods of non-linear dynamics in historical social research: application of Chaos Theory in the analysis of the Worker's movement in pre-Revolutionary Russia", *Historical Social Research* 22 (1997) 3/4, pp. 64-83.

[17] P. Nijkamp, "Voorspelbaarheid en chaos", *Mededelingen van de Koninklijke Nederlandse Akademie van Wetenschappen*, Afdeling Letterkunde. Nieuwe reeks 55:4 (1992), pp. 5-29.

## 5.2. Directions for further research

In summary, our research seems to be disappointing because we did not come up with a valid model. We think however that the outcome of the above is very instructive. Using several statistical techniques we were able to show that most assumptions about strike behaviour simply cannot be verified. Theories on strikes need to be refined by studying in more detail where, when and how strikes occur. The role of worker activists for example may play a more significant role than can be tested in a socio-economic model.

One way forward seems to be to analyse strike activity for several shorter periods. In fact, van der Velden did do this in his dissertation, for the period 1911-1940 and for the period 1945-1995, and for periods of between 15 and 30 years.[18] The correlations between the 'independent' variables and the strike index appear to fluctuate over time, which are an indication for the possible presence of not one, but a variety of 'models' (all different) statistically explaining the strike activity.

Of course, we do not plead for an atomising of strike research. Besides the detailed study of strikes and workers' protest in general, a scientific approach can contribute to a better understanding. After all, the negative results of our research have taught us much about the many easily made assumptions about human behaviour. Without these statistical techniques history will be nothing more than plain story-telling.

[18] Van der Velden, op cit., p. 272-273.