A Loglinear Analysis of Legal Representation Statistics on Hearings of Civil Cases in the District Court of Hong Kong

Wai-Sum Chan and Felix W. H. Chan***

Before 1 September 2000, the District Court's civil jurisdiction was limited to claims of up to HK$120,000. From 1 September 2000, the District Court's civil jurisdiction was extended to claims of up to HK$600,000. With effect from 1 December 2003, the civil jurisdiction of the District Court was further increased from HK$600,000 to HK$1 million. The main objective of this paper is to assess the impact of the enlargement of the District Court's jurisdiction on the usage rates of barristers in the conduct of civil cases in the District Court. Having performed a loglinear analysis of the legal representation statistics on the hearings for civil claims in the Hong Kong District Court over a two-year period from 1 September 1999 to 31 August 2001, the authors discovered, inter alia, that the enlargement of the District Court's jurisdiction has resulted in less use of barristers (relative to the use of solicitors) when all other factors are being controlled.

Introduction

In Hong Kong, the legal profession is divided into barristers and solicitors. There are about 800 barristers,[1] and 5,070 solicitors.[2] The two branches of the profession have their own governing bodies, namely the Hong Kong Bar Association and the Law Society of Hong Kong.

Broadly speaking, a barrister specialises in advice on intricate issues of law, the drafting of pleadings and the conduct of oral and written proceedings before courts on behalf of clients. A solicitor, to a large extent, can be characterised as a general practitioner, dealing with matters such as business negotiations, incorporation of companies, sale and purchase of property, tax planning, the drawing up of wills, the conduct of proceedings in the

* Associate Professor, Department of Statistics and Actuarial Science, University of Hong Kong.

** Associate Professor, Department of Professional Legal Education, University of Hong Kong. The authors are grateful to the Judiciary of the Hong Kong SAR and the Hong Kong Bar Association for the useful data and their kind assistance. The authors are also grateful to the anonymous referees for reading earlier drafts of this article and providing insightful comments.

1 The Hong Kong Bar Association website (http://www.hkba.org) accessed on-line on 10 Feb 2003.

2 The Law Society of Hong Kong website (http://www.hklawsoc.org.hk) accessed on-line on 10 Feb 2003.

Magistracies and the District Court, as well as much preliminary work in cases coming before the High Court and the Court of Final Appeal. Professor David Walker put it this way: "the distinction is not unlike that between medical specialist or consultant (particularly connected with a hospital), and general practitioner, and that between consultants and general practitioners in other professions".[3] Nonetheless, this statement must be interpreted with caution. Some solicitors, especially those in international law firms, are experts in particular areas of law, such as construction, intellectual property, corporate finance, shipping and banking. In contrast, some barristers, especially junior barristers, can handle a wide range of general civil and criminal cases.

Generally, only barristers have the right of audience in Hong Kong's Court of First Instance, the Court of Appeal and the Court of Final Appeal. Both barristers and solicitors share the right of audience before the Magistracies and the District Court. Before 1 September 2000, the District Court's civil jurisdiction was limited to claims of up to HK$120,000,[4] or, where claims were for recovery of land, where the annual rent or rateable value did not exceed HK$100,000. From 1 September 2000, the District Court's civil jurisdiction was extended to claims of up to HK$600,000, or, where claims are for recovery of land, where the annual rent or rateable value does not exceed $240,000. With effect from 1 December 2003, the civil jurisdiction of the District Court was further increased from HK$600,000 to HK$1 million.

The objectives of this paper are:

1 to assess the impact of the enlargement of the District Court's jurisdiction on the usage rates of barristers in the conduct of civil cases; and
2 to investigate the extent to which the use of barristers is dependent on other factors, such as the amount of the claim in question, the length of hearing and the type of hearing.

Methodology

In this section we review the loglinear modelling approach to analysing categorical data that is grouped in the form of contingency tables. We shall restrict the discussion to the points that are necessary for describing the applications in this paper. Further details can be found in many standard reference books in this field.[5]

[3] David M. Walker, *The Scottish Legal System* (UK: Sweet & Maxwell, 2001), p 367.
[4] This covers claims for contracts, quasi-contracts and torts. The limits that apply to claims under equity jurisdiction and claims for arrears of rent are different.
[5] For example, Ronald Christensen, *Log-Linear Models* (New York: Springer-Verlag, 1994); and Alan Agresti, *Categorical Data Analysis* (New York: John Wiley & Sons, 2002).

The purpose of most research is to assess relationships among a set of variables (factors). Choosing an appropriate statistical technique depends on the type of variables under investigation. If each element of the set of variables may lie only at a few isolated points, we have a *categorical* data set. In other words, a categorical variable is one for which the measurement scale consists of a set of categories; examples are gender, race, counts of events, or some sort of grouping.[6]

We focus on the analysis of categorical data summarised in a tabular form and the eventual aim concerns *multivariate* problems when at least three variables are involved. The loglinear analysis is the most appropriate statistical technique for modelling this type of data. A simple hypothetical example is employed to illustrate the key steps of loglinear modelling. Table 1 classifies 300 automobile accidents by gender of the driver (G), seat belt use (S) and injury (I).

Table 1 Automobile Accident Example

	Gender of the Driver			
	Male		Female	
	Injury		Injury	
Seat belt	Yes	No	Yes	No
Yes	4	120	1	22
No	31	80	18	24

Given the categorical data in a multi-dimensional contingency table, we can have three types of statistical inferences:

1 to test for a specific model; for example, we may wish to know whether two specified factors are conditional independent while controlling all the effects of other factors in the model;
2 if two factors are not conditional independent, a statistic (called the odds ratio) can be computed to measure the degree of their association; and
3 to search for a model that can best explain the relationship(s) found in the observed data.

We shall use the three-way data table in the automobile accident example (see Table 1) to illustrate these three types of statistical inferences from the loglinear analysis. Denote the cell expected frequencies in the data table by

[6] In *McClesjey v Zant*, 481 U.S. 270 (1987), the petitioner, a black man who had been sentenced to death for the murder of a white policeman, argued that the State of Georgia's death penalty was unconstitutional because it was applied in a racially discriminatory manner. Categorical data of 594 cases was analysed. Data for each case grouped the information about the race of killer and victim, the crime, the evidence, aggravating and mitigating circumstances, etc.

μ_{ijk}; where {i = male or female} for the gender (G) factor; {j = yes or no} for the seat belt (S) factor; and {k = yes or no} for the injury (I) variable. The homogeneous association loglinear model is defined as:

$$\log \mu_{ijk} = \lambda+\lambda_i^G+\lambda_j^S+\lambda_k^I+\lambda_{ij}^{GS}+\lambda_{jk}^{SI}+\lambda_{ik}^{GI}.$$

For convenience, the subscripts of the symbols are dropped. The model contains a constant λ, terms representing main effects (λ^G, λ^S, λ^I) and terms representing two-factor interactions (λ^{GS}, λ^{SI}, λ^{GI}).

Now, suppose we are interested in testing the null hypothesis that injury (I) and gender (G) are conditional independent:

$$H_0: \lambda + \lambda^G + \lambda^S + \lambda^I + \lambda^{GS} + \lambda^{SI}.$$

It should be noted that the interaction term λ^{GI} is omitted from the null model. If this model is adequately fitted by the data, then it implies that the relationship between factors G and I is not significant. On the contrary, after omitting this interaction term, if the null model is not supported by the data, it indicates that factors G and I could be related. Further investigation should be carried out.

When testing the goodness-of-fit in loglinear models, we often rely on the Pearson's χ^2 (chi-squared) statistic. Using the above null model, we have:

χ^2 = 3.10 with degrees of freedom (df) = 2, and p-value = 21.19%.

A large χ^2 value (or equivalently, a small p-value)[7] indicates a poor fit. Now, the computed χ^2 value is relatively small and its corresponding p-value is significantly larger than 5 per cent. We conclude that the null model is supported by the data and that factors G and I are conditional independent. This result means that there is no difference in the odds of having an injury during an automobile accident between male and female drivers, regardless of whether he or she was wearing a seat belt during the accident.

Suppose we are interested in testing another hypothesis that injury (I) and seat belt use (S) are conditional independent:

$$H_0: \lambda + \lambda^G + \lambda^S + \lambda^I + \lambda^{GS} + \lambda^{GI}.$$

Applying the above model we obtain the results:

7 Conventionally, it means that p-value is less than 5%.

$\chi^2 = 43.93$ with degrees of freedom (df) = 2, and p-value < 0.01%.

The data does not support the null model and so we need to carry out further analysis. Eight models are fitted and the results are summarised in Table 2. We attempt to search for the best model for the data.

Table 2 Possible Models for the Automobile Accident Example

Model	Description	χ^2	Degrees of Freedom	p-value	Conclusion
A	$\lambda + \lambda^G + \lambda^S + \lambda^I$	56.65	4	<0.01%	Reject the model
B	$\lambda + \lambda^G + \lambda^S + \lambda^I + \lambda^{GS}$	50.42	3	<0.01%	Reject the model
C	$\lambda + \lambda^G + \lambda^S + \lambda^I + \lambda^{GI}$	50.17	3	<0.01%	Reject the model
D	$\lambda + \lambda^G + \lambda^S + \lambda^I + \lambda^{SI}$	9.34	3	2.51%	Reject the model
E	$\lambda + \lambda^G + \lambda^S + \lambda^I + \lambda^{GS} + \lambda^{GI}$	43.93	2	<0.01%	Reject the model
F	$\lambda + \lambda^G + \lambda^S + \lambda^I + \lambda^{GS} + \lambda^{SI}$	3.10	2	21.19%	
G	$\lambda + \lambda^G + \lambda^S + \lambda^I + \lambda^{GI} + \lambda^{SI}$	2.85	2	24.01%	
H	$\lambda + \lambda^G + \lambda^S + \lambda^I + \lambda^{GS} + \lambda^{GI} + \lambda^{SI}$	0.09	1	76.52%	

Model G and Model H are two competing models for the data. Both of them are supported by the observations. Model H incorporates one more term λ^{GS} than Model G. The corresponding χ^2 value is reduced from 2.85 to 0.09 at the expense of one degree of freedom. Such reduction is not statistically significant.[8] Therefore, Model G is the best model to explain the automobile accident data set. Since the term λ^{SI} is included in the final model, it indicates that the seat belt (S) and injury (I) factors are related. Furthermore, the odds ratio implied from the fitted Model G for λ^{SI} can be computed. The result is 13.38. It means that the odds of having an injury for those who did not use a seat belt is about 13.38 times the odds of having an injury for those who used a seat belt during an automobile accident.

In this section we briefly explain what kinds of research questions can be answered using the loglinear analysis. Although we describe the method using an example with three factors, the procedure can be easily generalised to higher dimensional data tables. The computations required in the loglinear analysis can be carried out by many commonly used statistical computer packages.[9]

8 In order to be statistically significant, we require that the reduction is at least 3.84, which is the fifth percentile of a chi-squared random variable with one degree of freedom.

9 For example, Statistical Analysis System (SAS) and Statistical Package for the Social Sciences (SPSS).

The Data

We consider the legal representation statistics on hearings for civil claims in the Hong Kong District Court for a two-year period from 1 September 1999 to 31 August 2001.[10] The data were divided into two sub-periods for the purpose of studying the effects of the enlargement of the District Court's jurisdiction which took place on 1 September 2000.[11] Table 3 presents the data in a contingency tabular form. There are six categorical variables in the data set:

1 claim amount distribution (C);
2 length of time set down for the hearing (L);[12]
3 type of hearing (H);[13]
4 before / after the enlargement of the District Court's jurisdiction (E);
5 plaintiff / respondent (P); and
6 representation: counsel / solicitor (R).

Table 4 Variable Codes for the Legal Representation Statistics

Variable	Codes
C:	1 – less than HK$120,000 2 – HK$120,001 to HK$250,000 3 = HK$250,001 to HK$400,000 4 = HK$400,001 to HK$600,000
L:	1 – less than or equal to 2 hours 2 = more than 2 hours
H:	1 = Hearings before Masters 2 = Hearings before Judges 3 = Trial
E:	1 = before the enlargement 2 = after the enlargement
P:	1 = plaintiff 2 = respondent
R:	1 = counsel 2 = solicitor

10 The data was collected by the Hong Kong SAR Judiciary.

11 The period before the enlargement (1 Sept 1999 to 31 Aug 2000) and the period after the enlargement (1 Sept 2000 to 31 Aug 2001).

12 Originally, there were four response categories for this variable: (a) "less than or equal to 15 min"; (b) "more than 15 min but less than or equal to 2 hours"; (c) "more than 2 hours but less than or equal to 1 day"; and (d) "more than 1 day" The observed frequencies in the first two categories are fairly homogeneous and the same is found in the last two groups. Therefore, for convenience, the four categories are condensed into two groups: (a) "less than 2 hours" and (b) "more than 2 hours"

13 This variable has three possible categories of response: (a) "hearings before Masters (exclude taxation hearings)"; (b) "hearings before Judges (exclude taxation hearings)"; and (c) "trial hearings"

Table 3 Contingency Table for the Legal Representation Statistics

Claim Amount	Length	Types of Hearings	1 Sept 1999 to 31 Aug 2000 Plaintiff Counsel	1 Sept 1999 to 31 Aug 2000 Plaintiff Solicitor	1 Sept 1999 to 31 Aug 2000 Respondent Counsel	1 Sept 1999 to 31 Aug 2000 Respondent Solicitor	1 Sept 2000 to 31 Aug 2001 Plaintiff Counsel	1 Sept 2000 to 31 Aug 2001 Plaintiff Solicitor	1 Sept 2000 to 31 Aug 2001 Respondent Counsel	1 Sept 2000 to 31 Aug 2001 Respondent Solicitor
Less than HK$120,000	Less than /equal to 2 hrs	Masters	0	0	0	0	12	2,036	7	760
		Judges	31	1097	44	510	40	299	172	207
		Trial	1	1	0	1	1	1	2	2
	More than 2 hrs	Masters	0	0	0	0	0	0	0	0
		Judges	0	1	0	4	4	5	1	8
		Trial	106	74	103	58	97	68	143	49
HK$120,001 to HK$250,000	Less than /equal to 2 hrs	Masters	0	0	0	0	5	905	10	558
		Judges	0	27	0	5	35	141	163	112
		Trial	0	0	0	0	0	1	0	1
	More than 2 hrs	Masters	0	0	0	0	0	0	0	0
		Judges	0	0	0	0	4	5	4	2
		Trial	1	1	0	2	35	20	27	12
HK$250,001 to HK$400,000	Less than /equal to 2 hrs	Masters	0	0	0	0	4	487	4	30
		Judges	0	12	2	8	24	39	62	29
		Trial	0	0	0	0	0	0	0	0
	More than 2 hrs	Masters	0	0	0	0	0	0	0	0
		Judges	0	0	0	0	0	0	0	0
		Trial	0	5	0	3	14	8	21	1
HK$400,001 to HK$600,000	Less than /equal to 2 hrs	Masters	0	0	0	0	0	376	0	291
		Judges	0	31	1	3	2	40	4	32
		Trial	0	0	0	0	0	0	0	0
	More than 2 hrs	Masters	0	0	0	0	0	0	0	0
		Judges	0	0	0	0	1	4	3	2
		Trial	0	0	0	0	4	36	7	2

Note: Structural zeroes are shown in italics.

For the convenience of statistical analysis, the variables are numerically coded.[14] The definitions of the codes are given in Table 4.

The data table (see Table 3) in this study contains many zero counts. It is often called *sparse* in loglinear analysis. A zero for a sparse cell in which it is theoretically impossible to have observations is called a *structural zero*. On the other hand, if a zero counted cell event is theoretically possible (but it might be a rare event which we did not observe in this sampling period), we call the empty cell a *sampling zero*.

A structural zero is not part of the data set, and loglinear models should be fitted without such cells making any contribution to the estimation. In many statistical computer packages for analysing categorical data, it is necessary to explicitly identify structural zeroes. It should be noted that the category of "Hearings before Masters in the District Court" did not exist before the enlargement of the District Court's jurisdiction. Therefore, the zero counts in these cells are structural zeroes.[15]

Model Building

There are six dimensions in the data table for the legal representation statistics. It is difficult, if not impossible, to consider all possible combinations of models. To begin with the analysis, we look for a starting model, one that is able to explain the data well. A general model, which contains all combinations of terms up to three-factor interaction, is considered:

$$\begin{aligned} H_3:\ & \lambda + \lambda^{C} + \lambda^{L} + \lambda^{H} + \lambda^{E} + \lambda^{P} + \lambda^{R} \\ & + \lambda^{CL} + \lambda^{CH} + \lambda^{CE} + \lambda^{CP} + \lambda^{CR} + \lambda^{LH} + \lambda^{LE} + \lambda^{LP} \\ & + \lambda^{LR} + \lambda^{HE} + \lambda^{HP} + \lambda^{HR} + \lambda^{EP} + \lambda^{ER} + \lambda^{PR} \\ & + \lambda^{CLH} + \lambda^{CLE} + \lambda^{CLP} + \lambda^{CLR} + \lambda^{CHE} + \lambda^{CHP} + \lambda^{CHR} + \lambda^{CEP} + \lambda^{CER} + \lambda^{CPR} \\ & + \lambda^{LHE} + \lambda^{LHP} + \lambda^{LHR} + \lambda^{LEP} + \lambda^{LER} + \lambda^{LPR} + \lambda^{HEP} + \lambda^{HER} + \lambda^{HPR} + \lambda^{EPR}. \end{aligned}$$

Fitting the above model, we obtain:

χ^2 = 39.83 with degrees of freedom (df) = 71, and p-value = 99.90%.

This indicates that the model fits the data very well. Higher orders of interaction terms might not be useful for explaining the observations.[16] Therefore, we use the H_3 as the starting model.

[14] *See*, for example, M. O. Finkelstein and B. Levin, *Statistics for Lawyers* (New York: Springer, 2nd edn, 2001), p 464.

[15] There are 32 structural zeroes in the data table and they are shown in italics.

[16] Refer to four-factor, five-factor and six-factor interaction terms.

Next, we search for the best model to represent the data. The structure of the final model and its implied odds ratios are useful in interpreting the associations among the factors.

There are 42 terms in the starting model (H_3). We shall use the method of backward elimination to deduce the best model for the data table. In the first step, we fit the starting model. Also, the significance of each term in the model is calculated.[17] We perform the backward elimination by deleting the most insignificant term in the model at a time and re-calculating the χ^2 statistics for the remaining terms. The process continues until no term can be deleted. The resulting model is called the best model for the data.

The backward elimination procedure is carried out for the starting model. The final (best) model contains 26 parameters and all of them are statistically significant. The final model has the following structure:

$$\begin{aligned} H_{best}: & \lambda + \lambda^{P} + \lambda^{R} + \lambda^{CL} + \lambda^{CH} + \lambda^{CE} + \lambda^{CP} + 1\lambda^{CR} \\ & + \lambda^{LH} + \lambda^{LE} + \lambda^{HE} + \lambda^{HP} + \lambda^{HR} + \lambda^{EP} + \lambda^{ER} + \lambda^{PR} \\ & + \lambda^{CLH} + \lambda^{CLR} + \lambda^{CHP} + \lambda^{CEP} + \lambda^{LHE} + \lambda^{LEP} + \lambda^{LPR} + \lambda^{HER} + \lambda^{HPR} + \lambda^{EPR}. \end{aligned}$$

Statistical details of the final model are given in Appendix A.[18]

Main Results

First research question: has the enlargement of the District Court's jurisdiction resulted in a change of usage rates of barristers in the conduct of civil cases?

The question can be formulated as testing for the conditional independence between factor R (representation: counsel v solicitor) and factor E (enlargement of the District Court's jurisdiction: before v after) in the data table. Under the null hypothesis that the choice of representation has not been influenced by the court enlargement, the λ^{ER} term and its related three-factor interaction terms in the starting model should be zeroes. Therefore, five parameters (λ^{ER}, λ^{CER}, λ^{LER}, λ^{HER}, λ^{EPR}) are deleted from the starting model to form the null model:

[17] The significance of each parameter is indicated by its χ^2statistic and the corresponding p-value. A large χ^2 value (ie p-value < 5%) implies that the term is significantly different from zero. Therefore, this term with non-zero value should be retained in the model. In contrast, insignificant terms should be deleted from the final model.

[18] The SAS (Statistical Analysis System) software is employed to analyse the data. Appendix A contains an extract of the computer output for model fitting of the best model using PROC CATMOD.

$$H_0: \lambda + \lambda^{C} + \lambda^{L} + \lambda^{H} + \lambda^{E} + \lambda^{P} + \lambda^{R}$$
$$+ \lambda^{CL} + \lambda^{CH} + \lambda^{CE} + \lambda^{CP} + \lambda^{CR} + \lambda^{LH} + \lambda^{LE} + \lambda^{LP}$$
$$+ \lambda^{LR} + \lambda^{HE} + \lambda^{HP} + \lambda^{HR} + \lambda^{EP} + \lambda^{PR}$$
$$+ \lambda^{CLH} + \lambda^{CLE} + \lambda^{CLP} + \lambda^{CLR} + \lambda^{CHE} + \lambda^{CHP} + \lambda^{CHR} + \lambda^{CEP} + \lambda^{CPR}$$
$$+ \lambda^{LHE} + \lambda^{LHP} + \lambda^{LHR} + \lambda^{LEP} + \lambda^{LPR} + \lambda^{HEP} + \lambda^{HPR}.$$

Fitting the above null model, we get the results:

χ^2 = 320.64 with degrees of freedom (df) = 77, and p-value < 0.01%.

The null model is rejected by the data. This indicates that factors R and E are significantly related. The enlargement of the District Court's jurisdiction has resulted in a change of usage rates of barristers in the conduct of civil cases. Furthermore, the maximum likelihood estimate of the parameter λ^{ER} in the best model is negative (see Appendix A), *which indicates that the enlargement has resulted in less use of barristers when other factors are being controlled.*

Further analysis can be obtained from the final fitted loglinear model H_{best}. Two related three-factor interaction terms (λ^{HER}, λ^{EPR}) are statistically significant. The significance of λ^{HER} implies that factor H (type of hearing) modifies the *interaction effects of ER (enlargement and representation factors). As we mentioned before, "Hearings before Masters in the District Court" was only allowed after the enlargement of the District Court's jurisdiction in September 2000. This change contributed significantly to the relative reduction in the use of counsel in civil claims after the court enlargement.*

To study the impact of the introduction of "Hearings before Masters" (HBM) in the District Court after September 2000 on the choice of legal representation, we perform an additional analysis. First, the data table (Table 3) is condensed into a simple three-way form:

Table 5: The Condensed Data Table

	Enlargement				
	Before		After		
	Non-HBM	HBM	Non-HBM	HBM	Total
Counsel	289	0	870	42	1,201
Representation Solicitor	1,843	0	1,126	5,714	8,683
Total	2,132	0	1,196	5,756	9,884

After the enlargement of the District Court's jurisdiction, there was a drastic increase in the number of District Court HBM, from 0 (not allowed) to over 5,000 cases. These HBM cases have very low use of barristers

(P = 42/5,756 = 0.73%). It might be due to the fact that HBM cases usually deal with short routine matters for which it might not be worthwhile instructing counsel.

Excluding the HBM cases, the observed proportion of the use of counsel before the court enlargement is P^*_b = 289/2,132 = 13.56% and the proportion is P^*_a = 870/1,196 = 43.59% after the enlargement. The overall proportion is P^* = 1,159/4,128 = 28.08% with sample sizes n_b = 2,132 (before the enlargement) and n_a = 1,196 (after the enlargement). A standard *z-score* test is computed:[19]

$$z = \frac{P^*_a - P^*_b}{\sqrt{P^*(1-P^*)\left[\frac{1}{n_a} + \frac{1}{n_b}\right]}}$$

$$= \frac{0.4359 - 0.1356}{\sqrt{(0.2808)(0.7192)\left[\frac{1}{1196} + \frac{1}{2132}\right]}}$$

$$= 21.46$$

The critical value of the *z-score* test is 1.645 at the 5 per cent level of significance. The test result concludes that the true value of P^*_a is statistically (with a very high significance level) larger than the true value of P^*_b.

Our additional analysis discovers that *the enlargement of the District Court jurisdiction allows HBM that were not allowed before. The cases regarding HBM involved the rare use of barristers. As for the remaining cases other than HBM, the enlargement led to more use of barristers.*

Concerning the significance of the parameter λ^{EPR} in the context of choice of legal representation, the degree of difference between the situations before and after the enlargement of jurisdiction is influenced by whether the plaintiff or the defendant is the client in the matter. This would be easier to explain using the odds ratio concept. From Appendix A, we obtain the maximum likelihood estimates of λ^{ER} and λ^{EPR}. Their values are –0.5161 and 0.0757, respectively. The odds ratio for the ER interaction, given the factor P (plaintiff v respondent), is calculated as follows:[20]

[19] *See* n 16 above, s 5.2.

[20] More technical details of the calculations are available in Appendix B.

(a) Estimates for λ^{ER}:

Representation	Enlargement	
	Before (code = 1)	After (code = 2)
Counsel (code = 1)	–0.5161	0.5161
Solicitor (code = 2)	0.5161	–0.5161

(b) Estimates for λ^{EPR}:

Client	Representation	Enlargement	
		Before (code = 1)	After (code = 2)
Plaintiff (code = 1)	Counsel (code = 1)	0.0757	–0.0757
	Solicitor (code = 2)	–0.0757	0.0757
Respondent (code = 2)	Counsel (code = 1)	–0.0757	0.0757
	Solicitor (code = 2)	0.0757	–0.0757

For plaintiff (code = 1):

log (odds ratio: after v before)
= 2 [(0.5161) – (–0.5161)] + 2 [(–0.0757) – (0.0757)]
= 1.7616.

(odds ratio: after v before)
= exp(1.7616)
= 5.82.

The conclusion derived from this odds ratio figure can be stated as: *for plaintiffs, barristers are more likely (5.82 times) to be used after the enlargement of the District Court's jurisdiction than before the enlargement.*

The second research question: to what extent is the use of barristers dependent on the other factors (eg the claim amount distribution, the length of hearing and the type of hearing)?
From the best model, we understand that the parameters (λ^{CR}, λ^{HR}, λ^{PR}) are statistically significant. In other words, in addition to the enlargement of jurisdiction, these three factors (the claim amount, the type of hearing, and

whether the plaintiff or the respondent is the client in the matter) are related to the choice of legal representation. On the other hand, λ^{LR} is not significant and is excluded from the final model. The duration of hearing is not directly related to the choice of representation. However, this factor influences the other relationships through the three-factor interaction terms. For example, λ^{CLR} is a significant term in the best model. Again, we illustrate the influence of factor L (length of hearing) on the CR relationship by the concept of odds ratios. The relevant maximum likelihood estimates for the parameters λ^{CR} and λ^{CLR} are extracted from Appendix A.

(c) Estimates for λ^{CR}:

	Claim Amount[21]			
Representation	C1	C2	C3	C4
Counsel (code = 1)	0.0990	0.2263	0.3522	–0.6775
Solicitor (code = 2)	–0.0990	–0.2263	–0.3522	0.6775

(d) Estimates for λ^{CLR}:

		Claim Amount			
Length of hearing	Representation	C1	C2	C3	C4
Less than 2 hours (code = 1)	Counsel (code = 1)	–0.0999	0.0277	0.1627	0.0905
	Solicitor (code = 2)	0.0999	–0.0277	–0.1627	–0.0905
More than 2 hours (code = 2)	Counsel (code = 1)	0.0999	–0.0277	–0.1627	–0.0905
	Solicitor (code = 2)	–0.0999	0.0277	0.1627	0.0905

Odds ratios for comparing different pairs of claim amount situations are obtained in Table 6. We employ the case (C3 v C1) as an example for illustrating the calculation details and interpretation of figures in the Table. Others cases can be computed and interpreted in a similar way.

21 For convenience, we let C1 = {claim amount is less than HK$120,000}; C2 = {claim amount is between HK$120,001 and HK$250,000}; C3 = {claim amount is between HK$250,001 and HK$400,000}; and C4 = {claim amount is between HK$400,001 and HK$600,000}.

Table 6 Computed odds ratios

Claim amount Comparison	Length of hearing ≤ 2 hours	> 2 hours
C4 v C3	0.1	0.1
C4 v C2	0.2	0.1
C4 v C1	0.3	0.1
C3 v C2	1.7	1.0
C3 v C1	2.8	1.0
C2 v C1	1.7	1.0

For the length of hearing ≤ 2 hours in the case (C3 v C1),

$$\begin{aligned}\log(\text{odds ratio: C3 v C1}) &= 2\{(0.3522) - (0.0990)\} + 2\{(0.1627) - (-0.0999)\} \\ &= 1.0316\end{aligned}$$

$$\begin{aligned}(\text{odds ratio: C3 v C1}) &= \exp(1.0316) \\ &= 2.8.\end{aligned}$$

For the length of hearing > 2 hours in the case (C3 v C1),

$$\begin{aligned}\log(\text{odds ratio: C3 v C1}) &= 2\{(0.3522) - (0.0990)\} + 2\{(-0.1627) - (0.0999)\} \\ &= -0.0188\end{aligned}$$

$$\begin{aligned}(\text{odds ratio: C3 v C1}) &= \exp(-0.0188) \\ &= 0.98 \\ &\approx 1.0.\end{aligned}$$

The interpretation of these two odds ratio statistics is as follows:

1 when the duration of hearings is shorter than or equal to 2 hours, barristers are more likely (2.81 times) to be instructed in cases involving claim amounts between HK$250,001 and HK$400,000, than when claim amounts are less than HK$120,000; and
2 when the duration of hearings is longer than 2 hours, there is no difference in the choice of legal representation for cases involving claim amounts between HK$250,001 and HK$400,000 or in cases with claim amounts less than HK$120,000.

From Table 6, we conclude that barristers are *more likely* to be instructed in C3 cases when the duration of hearings is shorter than or equal to 2 hours. On the other hand, barristers are *less likely* to be instructed in C4 cases (ie claim amounts greater than HK$400,000) in the District Court.

A Summary of the Key Findings

We have performed a loglinear analysis of the legal representation statistics on the hearings for civil claims in the Hong Kong District Court for a two-year period from 1 September 1999 to 31 August 2001. The key findings are as follows:

1 the enlargement of the District Court's jurisdiction has resulted in *less* use of barristers (relative to the use of solicitors) when all other factors are being controlled;

2 other things being equal, plaintiffs (as compared to respondents) are much more likely to *reduce* the relative use of barristers since the enlargement of the District Court's jurisdiction in the conduct of civil claims;

3 hearings before Masters (HBM) in the District Court are allowed since the enlargement of the District Court's jurisdiction in September 2000. This change has contributed significantly to the relative reduction in the use of barristers in civil claims since the enlargement of jurisdiction. Excluding the HBM cases from the analysis, the conclusion in (1) has been reversed. In other words, the enlargement of the District Court's jurisdiction has resulted in *more* use of barristers (relative to the use of solicitors) *when HBM cases are being discarded*; and

4 in addition to the extension of the District Court's jurisdiction, three other factors (namely, the amount of the claim, the type of hearing, and whether the plaintiff or the respondent is the client) also have impact on the choice of legal representation. For instance, when the duration of hearings is shorter than or equal to 2 hours, barristers are *more likely* to be instructed in cases involving claim amounts between HK$250,001 and HK$400,000, than in other cases. On the other hand, barristers are *less likely* to be instructed in cases involving claim amounts greater than HK$400,000 in the District Court.

Legal representation statistics for land related (Possession and O.88/O.84) cases and non-monetary and unliquidated claims in District Court are also available. However, there are too many structural and sampling zeroes in these data tables and they have not been examined in this paper.

Conclusion and Recommendation for Future Research

The key findings of the present loglinear analysis reveal several interesting associations among the variables for the legal representation statistics. However, it should be noted that a statistical loglinear model only helps detect association between two or more factors. It does not identify the causes of such phenomenon. Whether there is a cause-and-effect relationship among the various factors needs further investigation. The findings could be attributed to other factors that were not considered in the raw data set. For instance, factors such as the relative costs of barristers and solicitors, and the relative financial positions of the plaintiffs and the defendants, may be relevant in explaining the phenomenon.

In the future, researchers can consider conducting a qualitative research to explore the causes explaining the above quantitative findings. Structured interviews could be conducted with the plaintiffs, defendants, solicitors and barristers involved in the District Court hearings. The data, if systematically collected and analysed, could provide a rich source of qualitative insights.

APPENDIX A

SAS Computer Output for the Best Model

Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr>ChiSq
P	1	21.36	< .0001
R	1	724.47	< .0001
C*L	3	152.38	< .0001
C*H	6	278.71	< .0001
C*E	3	409.78	< .0001
C*P	3	12.72	0.0053
C*R	3	58.96	< .0001
L*H	2	222.21	< .0001
L*E	1	198.63	< .0001
H*E	2	197.53	< .0001
H*P	2	28.18	< .0001
H*R	2	688.73	< .0001
E*P	1	15.61	< .0001
E*R	1	208.53	< .0001
P*R	1	12.13	0 .0005
C*L*H	6	854.98	< .0001
C*L*R	3	11.79	0 .0081
C*H*P	6	48.11	< .0001
C*E*P	3	10.73	0 .0133
L*H*E	2	464.07	< .0001
L*E*P	1	12.15	0 .0005
L*P*R	1	4.29	0 .0383
H*E*R	1	70.10	< .0001
H*P*R	2	10.04	0 .0066
E*P*R	1	9.68	0 .0019

SAS Computer Output for the Best Model (continued)

Analysis of Maximum Likelihood Estimates

	Parameter			Standard Estimate	Chi- Error	Square	Pr>ChiSq
P	1			0.2829	0.0612	21.36	<.0001
R	1			-1.4797	0.055	724.47	<.0001
C*L	1	1		2.1873	0.1906	131.75	<.0001
	2	1		-0.2377	0.2337	1.03	0.3091
	3	1		-0.2153	0.2475	0.76	0.3844
C*H	1	1		-5.4301	0.4894	123.08	<.0001
	1	2		2.9144	0.2669	119.23	<.0001
	2	1		-0.1939	0.4667	0.17	0.6779
	2	2		0.7285	0.2744	7.05	0.0079
	3	1		3.7803	0.587	41.47	<.0001
	3	2		-3.315	0.6916	22.97	<.0001
C*E	1	1		1.0551	0.0558	357.41	<.0001
	2	1		-0.7523	0.0966	60.63	<.0001
	3	1		-0.0467	0.0864	0.29	0.589
C*P	1	1		-0.1255	0.0556	5.1	0.0239
	2	1		0.0911	0.0989	0.85	0.3572
	3	1		-0.2582	0.0875	8.71	0.0032
C*R	1	1		0.099	0.0456	4.71	0.0299
	2	1		0.2263	0.056	16.34	<.0001
	3	1		0.3522	0.0687	26.26	<.0001
L*H	1	1		-0.9146	0.0909	101.16	<.0001
	1	2		3.19	0.2524	159.71	<.0001
L*E	1	1		-2.4815	0.1761	198.63	<.0001
H*E	1	1		2.9497	0.2265	169.54	<.0001
	2	1		-1.7343	0.1373	159.54	<.0001
H*P	1	1		0.1344	0.059	5.19	0.0227
	2	1		-0.2259	0.0425	28.18	<.0001
H*R	1	1		-2.0247	0.0772	687.33	<.0001
	2	1		0.4601	0.0411	125.02	<.0001
E*P	1	1		0.2196	0.0556	15.61	<.0001
E*R	1	1		-0.5161	0.0357	208.53	<.0001
P*R	1	1		-0.1392	0.04	12.13	0.0005
C*L*H	1	1	1	5.0648	0.3546	203.96	<.0001
	1	1	2	-2.9189	0.3973	53.97	<.0001
	2	1	1	0.1411	0.3443	0.17	0.6819
	2	1	2	-0.516	0.3657	1.99	0.1583
C*L*H	3	1	1	-3.5137	0.6288	31.23	<.0001
	3	1	2	3.1339	0.6094	26.45	<.0001
C*L*R	1	1	1	-0.0999	0.0428	5.44	0.0197
	2	1	1	0.0277	0.0544	0.26	0.6109
	3	1	1	0.1627	0.0699	5.43	0.0198

SAS Computer Output for the Best Model (continued)

Analysis of Maximum Likelihood Estimates

	Parameter	Standard Estimate	Chi- Error	Square	Pr>ChiSq
C*H*P	1 1 1	0.2218	0.033	45.2	<.0001
	1 2 1	–0.00169	0.0386	0	0.9652
	2 1 1	0.00129	0.0407	0	0.9747
	2 2 1	–0.0201	0.0475	0.18	0.673
	3 1 1	0.0012	0.0512	0	0.9814
	3 2 1	0.081	0.0618	1.72	0.1895
C*E*P	1 1 1	–0.1209	0.0552	4.79	0.0286
	2 1 1	0.1293	0.0974	1.76	0.1846
	3 1 1	–0.2241	0.0872	6.6	0.0102
L*H*E	1 1 1	–5.5248	0.298	343.8	<.0001
	1 2 1	3.2067	0.2328	189.67	<.0001
L*E*P	1 1 1	0.0853	0.0245	12.15	0.0005
L*P*R	1 1 1	–0.1407	0.0679	4.29	0.0383
H*E*R	1 1 1	–0.4367	0.0522	70.1	<.0001
	2 1 1	–	–	–	–
H*P*R	1 1 1	0.1921	0.0714	7.23	0.0072
	2 1 1	–0.0657	0.0544	1.46	0.2276
E*P*R	1 1 1	0.0757	0.0243	9.68	0.0019

APPENDIX B

The following shows the technical calculation of the odds ratio for plaintiff.

Consider only factors E, P and R and let μ_{ijk} denotes the frequency for cells when E, P and R take levels *i*, *j*, and *k* respectively. For plaintiff ($j = 1$):

log (odd ratio: after v before)

$$= \log \left[\frac{\mu_{211} / \mu_{212}}{\mu_{111} / \mu_{112}} \right]$$

$$= \log \left[\frac{\mu_{211} \cdot \mu_{212}}{\mu_{111} \cdot \mu_{112}} \right]$$

where μ_{211} / μ_{212} is the odds of the use of barristers relative to solicitors after the enlargement, and μ_{111} / μ_{112} is the odds of the use of barristers relative to solicitors before the enlargement.

Now, $\lambda_1^E = 0$, $\lambda_1^P = 0.2829$, $\lambda_1^R = -1.4797$, $\lambda_{11}^{EP} = 0.2196$, $\lambda_{11}^{ER} = -0.5161$, $\lambda_{11}^{PR} = -0.1392$, $\lambda_{111}^{EPR} = 0.0757$ for the best model. Here

$$\log \mu_{211} = 0 + 0.2829 - 1.4797 - 0.2196 + 0.5616 - 0.1392 - 0.0757$$
$$\log \mu_{212} = 0 + 0.2829 + 1.4797 - 0.2196 - 0.5616 - 0.1392 + 0.0757$$
$$\log \mu_{111} = 0 + 0.2829 - 1.4797 + 0.2196 - 0.5616 - 0.1392 + 0.0757$$
$$\log \mu_{112} = 0 + 0.2829 + 1.4797 + 0.2196 + 0.5616 + 0.1392 + 0.0757$$

and

log (odds ratio: after v before)

$$= 2[(0.5161) - (-0.5161)] + 2[-0.0757 - (0.0757)]$$
$$=1.7616.$$