

中央研究院中山人文社會科學研究所專書(41)

社會科學計量方法發展與應用

楊文山 主編

中華民國八十六年九月
臺北 南港

目 錄

編者序

The Structure of the Chinese CES-D Scale:

An Application of Confirmatory Factor Analysis

to Higher-Order Factor Structure.....周雅容 1-26

兄弟姊妹教育相似性的探究：

兩種「迷迷模型」的應用薛承泰 27-60

「對數線性模型」與「對數相乘模型」的發展與

運用：以社會流動的分析為例蔡瑞明 61-102

團體比較研究中結構差異的考慮與處理：

論迴歸標準化及對數線性消除法的應用

.....謝雨生、鄭宜仲 103-137

Covariance Structure Analysis with

Intraclass Dependent Observations... 翁儷禎、P. M. Bentler 139-183

大陸臺商工廠待獲利期間之分析：

存活模型之應用林祖嘉、李紀珠 185-209

貨幣數量學說的競爭性：

遺傳規畫在知識發掘上的應用陳樹衡、葉佳炫 211-262

社會共識程度之衡量施俊吉 263-287

模糊統計在社會調查分析的應用吳柏林、楊文山 289-316

作者簡介 317-320

索 引 321-325

附 錄 327-329

編 者 序

記得二十多年前剛開始社會學研究生的生涯時，在社會科學實驗室門外看到一部龐然大物，雖然常常經過它，卻發現從來沒有人使用過這一部機器。有一次好奇走近這一個龐然大物，看看它究竟是什麼東西？機器的側面印著國際商務機器公司 (IBM) 的商標，但是並沒有任何的說明。直到有一天上研究法時，教授才道出這是社會科學邁向量化的一個重要里程碑的紀念物，也就是著名的人工半自動讀卡機，可是這時已經落伍而沒有人使用。而社會科學實驗室中所陳列的昂貴儀器，也只是王安電腦公司所出的類比式電子計算器而已。此一計算器從資料輸入到結果輸出，僅可以做一些最基本的統計運算而已，其功能不如當前的一部手上型電子計算機。今天電子計算機與套裝軟體的普遍運用，一部桌上型的電腦以及其所附帶的功能，可能比當時的電腦主機具有更多的運算能力。

電子計算機與套裝軟體在社會科學中的普遍運用，以及社會科學計量方法的發展，是今日社會科學的一項重要趨勢。在此一背景與前提下，中央研究院中山人文社會科學研究所，於民國八十四年三月九日至十日舉辦了「社會科學計量方法發展與應用」學術研討會。此一研討會的主要目的不僅對於計量方法應用提出適切的討論，並且對於計量方法的未來發展趨勢提供社會科學家彼此一次切磋琢磨的機會，而且對於國內近年來計量方法在社會科學中的應用與前瞻，做一次成果的發表與展現。

由於過去國內社會科學界對於研究法與計量方法已有多次的研討會，對於計量方法在社會科學的應用，已有相當精闢入裡的討論與見地，我們不敢專美於前。所以此次的研討會在籌辦過程中，我們嘗試另闢蹊徑，以社會科學計量方法晚近的發展與未來的運用，作為此次研討會的主題。社會組同仁在數次的籌備會議中集思廣益，並且參照

近年來國內外社會科學期刊論文計量方法的應用情形，與社會科學計量方法發展的趨勢，作為此次研討會的主題。再三研議之下，社會組同仁決定以歷年所出版的社會學方法 (Sociological Methodology) 專書以及社會學研究與方法 (Sociological Research and Methods) 期刊中所討論的計量方法為經，並以社會學研究與方法期刊中數次對於計量方法的發展趨勢專題為緯，作為此次會議邀稿的目標與研討會的主題。

此次會議共發表十六篇論文，這些論文環繞在下列六個主題議程中。這些主題分別為：一、結構方程式模型；二、對數線性模型；三、迴歸方法之應用；四、事件史分析；五、時間數列；六、計量方法的發展。每篇論文均獲得與會學者的廣泛迴響與熱烈討論，在彼此知無不言，言無不盡的問辯及詰難中，與會學者對於社會科學計量方法的發展與應用，均有相當豐碩的收穫與視野的開拓。本書所蒐集的九篇論文是以大會主題為綱，依論文發表的先後次序所編排的論文集。此一論文集的另一特色是將各篇發表論文中重要的詞彙與特殊名詞，以中英文依序置於書末之索引中，以利於讀者的檢視與查詢。

本書所發表的論文其內容摘述於下。周雅容教授在“the Structure of the Chinese CES-D Scale: An Application of Confirmatory Factor Analysis to Higher-Order Factor Structure.” 為題之論文中，以驗證性因素分析模式來探討美國流行病研究中心所發展的憂鬱量表 (CES-D) 之因素結構。研究資料來自兩個不同年齡層樣本，659位老人及389位大學生。CES-D 的因素結構包含四個因素：正向情緒、負向情緒、身體化相關症狀及人際關係因素。分析結果指出，CES-D 量表在台灣地區樣本資料並不支持四個因素結構模式，反而是三個因素結構較符合實證資料，原因是負面情緒與身體化症狀兩因素相關度太高，不符合成為兩個獨立因素條件，應合併比較恰當。

薛承泰教授則在〈兄弟姊妹教育相似性的探究：兩種「迷迷模型」的應用〉一文中，以兄弟姊妹教育的相似性為課題，來說明「一

般迷迷模型」(MIMIC Model)與「等級迷迷模型」(Ordered MIMIC Model)的應用。研究發現家庭社經地位以及潛藏的共同因素(可能源自於手足之間的互動),對手足配對的成就之影響,並不會因性別或長幼之別而有差異。至於對手足配對的教育取得之影響,社經地位的作用仍然並不會因性別或出生序而有差異;但潛藏共同因素在控制了家庭社經地位情形下,其影響力以長男較長女強些。

蔡瑞明教授所發表的〈「對數線性模型」與「對數相乘模型」的發展與運用:以社會流動的分析為例〉論文,主要是討論對數線性分析(log-linear analysis)的最新發展趨勢。最新發展的關連模型結合了對數線性與對數相乘的參數,它們不僅擁有 Goodman 的關連模型中項目排列的優越性,適合分析垂直(vertical)關係,而且又可加入其他非垂直(nonvertical)的結構性參數。另外,文中使用「台灣地區社會變遷基本調查」的資料,以對數線性與對數相乘模型來分析台灣男性與女性的代間傳承之相對社會流動形態,分析的結果發現女性的階級受其父親的影響比男性來得高。

謝雨生教授與鄭宜仲先生共同發表〈團體比較研究中結構差異的考慮與處理:論迴歸標準化及對數線性消除法的應用〉一文,提供研究者兩種結構差異的處理方法。一是迴歸標準化(regression standardization)或是迴歸調整法,是以多元迴歸為主的調整法;另一則是屬於非連續性變項的處理方法,即對數線性消去法(log-linear purging method),是以對數線性模式(log-linear model)為主的調整法。團體粗差異的分解實例更闡明及釐清了團體比較研究中團體結構差異的考慮及處理的必要性和重要性。

翁儷貞教授(Li-Jen Weng)與 P.M.Bentler 教授合著的“covariance Structure Analysis with Intraclass Dependent Observations.”這篇論文則在探討共變結構分析模型中相依觀察值的問題。觀察值之間相互獨立是共變結構分析模型的基本假設之一,但在很多環

境下很難辦到，因此作者試圖解決此一困境。資料被假定為常態分配的矩陣，相依結構是方陣對角線內部群，並假設每一群體的觀察值相關性相等。母體共變矩陣的最大近似估計值及內部群相關性具一致性、漸進常態與漸進有效性。作者利用母體共變矩陣提出一個二階過程，以電腦模擬與真實資料來研究相依觀察值。

林祖嘉、李紀珠二位教授所著之〈大陸台商工廠待獲利期間之分析〉一文，採用存活模型 (survival analysis model) 或事件史分析模型對於廠商待獲利期間加以分析，探究影響赴大陸廠商投資待獲利期間的可能因素為何？以及待獲利期間是否具有時間依賴性 (time dependence)？研究發現廠商員工人數、自台灣進口原料比例、自台灣進口機器比例、成本降低比例、投資型態、投資地區及產業型態等因素，都對大陸台商的投資待獲利期間有顯著的影響。另一方面，模型估計結果發現待獲利期間出現明顯的正時間相依效果，顯示赴大陸投資的獲利風險會隨著待獲利期間的增加而降低。

陳樹衡教授與葉佳炫先生發表的〈貨幣數量學說的競爭性：遺傳規劃在知識發掘上的應用〉論文則指出，遺傳規劃 (Genetic Programming) 提供了我們一個環境，在此環境中，任何現今的理論或模型，都將面臨由演化所產生潛在競爭者的挑戰。如果在演化的過程中，它被其他的競爭者所擊敗，那我們說它缺乏競爭力，否則稱它具有競爭力。作者使用在經濟學中非常著名的貨幣數量學說為例子，以美國的資料而言，貨幣數量學說是非常具有競爭力的，但是對於台灣的資料來說，卻並不具有競爭力。文中藉著具體的應用來試圖說明遺傳規劃如何可能成爲一個知識發掘的有利工具。

施俊吉教授在〈社會共識程度之衡量〉一文中，探討如何利用實地訪談問卷調查所獲得的資料，衡量社會對於某一特定問題之共識程度。此一論文所研究的問題是社會之共識能否分出等級與程度？目的則是建立一項測量共識程度的指標。由於共識程度乃是一個含意深邃

且複雜的概念，既有的不均度指數以及其他用於衡量離散程度的指標，都無法精確反映「共識」這項概念。因此，作者發展出一個新的指標，將之命名為「 s^* 指標」，此一指標意義既清楚又相當明確，不會由概念的測量到資料的轉換而失真，故作者建議社會學者在衡量共識程度時應當採用 s^* 指標。

至於吳柏林與楊文山二位教授，在共同發表的〈模糊統計在社會調查分析的應用〉一文中則是認為，當實地調查中所使用的問卷問題本身的經驗指標模糊或不一定的狀況下，如果用結構性的問卷設計，並延用清楚 (crisp) 的數學模式作為分析依據時，就可能導致誤差，產生信效度的偏誤，進而影響研究的可能結果。故提出模糊統計調查與分析模糊資料的方法，並將此一模糊統計與分析方法實際應用在社會調查中，從社會與經濟問題來分析人類思考與行為模式。希望藉模糊問卷調查及模糊統計分析，對人類在思想認知不易表達之問題，能有更合理的測度與解釋。

綜觀本書所彙集의九篇論文，對於晚近社會科學計量方法的發展與應用，均提出了相當精闢的見解與實地應用的方法。本書在出版過程中曾要求論文作者，對於估計各種計量模型的套裝統計軟體一併發表，因此各位讀者在仔細閱讀各篇論文之際，若有相同資料並估計類似模型，可以參照論文之後所列出的套裝統計軟體程式，作為估計模型之用。

大會所宣讀的論文由於某些特殊原因，並未完全收錄於本書之中，為免造成遺珠之憾，所以本書附錄中收錄此次會議發表的所有論文題目以及論文作者之職稱，若讀者對於本書未收錄的會議論文有參考需要，可直接向作者索閱。值得一提的是，本書所未收錄的論文均極具參考價值，例如：吳齊殷教授對於潛藏曲線的發展與應用，提出了相當精闢的見解與實際運用的實例。李隆安、陳君厚與李克昭諸教授所提出的反切迴歸在社會科學領域的應用，可能是此一晚近發明的

統計方法在世界社會科學界中所發表的第一篇論文。鄭讚源教授所發表的多層級分析策略，是國內第一篇關於多層級模型理論與實務的文章。陳信木與林佳瑩教授所發表的事件史分析模型之應用，是一篇由簡入繁、由淺入深關於事件史分析的代表力作。田芳華教授所發表的文章對於事件史分析在教師升等的應用上，提出了精闢的見解與實際操作此一方法的實例。林建甫教授與盧陽正先生所提出的 LSE 經濟計量學的分析方法，相當具有開創性與實用性。邱昌泰教授則對匯合分析法做出了整理與如何實地應用的討論。

此次會議能夠順利進行，首先要感謝本所前所長彭文賢先生的鼎力支持，以及本所社會組前主任蔡淑鈴女士的規劃與居中協調。蔡淑鈴女士在會議籌備過程中，展現了許多過人的見解與分工並整合的領導風格，讓人耳目一新。此外，社會組同仁伊慶春女士、朱瑞玲女士、李安妮女士、陳寬政先生、覃怡輝先生以及吳齊殷先生在籌備過程中所提供的意見，以及在大會中的付出與辛勞，在此致上特別的敬意與謝意。

此外，也要感謝參與此次會議諸位學者以及擔任會議主席、評論人諸位女士、先生的辛勞與為國內社會科學計量發展的付出。此次會議共有一百廿十餘國內學者並各校社會科學研究生與會，其中不乏遠道由中南部趕來赴會的先生、女士，在此特別致上謝意與敬意。更要感謝的是在會議期間，社科所行政人員所付出的辛勞，使會議得以順利、圓滿進行。社會組組助理陸昭陽先生不辭辛苦通知開會、安排會議細節，以及彭玉屏小姐往後蒐集論文所付出的努力，他們兩位是促成此次會議順利成功以及本書出版的兩位重要幕後功臣。此論文集的出版從論文的蒐集、送審、編排到完整出書，要感謝中研院社科所出版室同仁：任淑華小姐、林容如小姐、張芝蘭小姐、崔承蕙小姐以及協助校對的小姐、先生們，在工作繁忙之餘協助完成本書的出版事宜，印製精良的本書算得上是他們辛勞的成果。這裡值得一提的是，

出版室同仁林蓮治小姐在協助本書出版時不幸因病辭世，她的專業精神以及對於本所出版室的付出，特別值得懷念。

社會科學計量的發展可以說是一日千里，但是計量方法如何配合社會科學理論而對人類行為有更深入的瞭解，這可能才是計量方法在社會科學中的真正目的。本書從研討會到完整出書歷經了比較長久的時間，看到本書的出版，也算得上是國內社會科學界近年來對於計量方法的應用與發展的成果展示。希望藉此一論文集的出版能夠開拓國內社會科學界對於計量方法的認識與瞭解，同時也能夠引發社會科學學者討論如何正確地使用計量方法，並且發展新的社會科學計量方法，期望未來國內的社會科學計量研究成果，能夠在世界社會科學計量方法研究中攻佔一席之地。

楊文山於中研院社科所
中華民國八十六年九月

《社會科學計量方法發展與應用》 楊文山主編
中央研究院中山人文社會科學研究所專書 (41) · 頁1-26
民國86年9月，臺灣，臺北

The Structure of the Chinese CES-D Scale: An Application of Confirmatory Factor Analysis to Higher-Order Factor Structure

Yah-Jong Chou*

I . Introduction

Psychological distress, especially depression, is an important construct in research on stress, social support and coping process. There is compelling evidence in western as well as Chinese society that high level of psychological distress is significantly associated with stressful life events, chronic life strain, and lack of support (Chou, 1993; Cohen & Syme, 1985; Kessler, Price & Worthman, 1985). However, meaningful comparisons of research findings from different studies are based on the assumption that psychological distress in various social groups and across different cultures is in fact the same construct. And many researchers have argued that this assumption may not be warranted. For instance, there may be cultural differences in the manifestation of depressive

* National Chung Cheng University. The author thanks the anonymous reviewer for comments on an earlier version of this paper.

symptoms or in the expression of psychological distress (Aneshensel, Clark, & Frerichs, 1983; Kleinman, 1986).

The major purpose of the study is to investigate whether a depressive-symptom scale developed in the United States is assessing the same underlying psychological distress in Taiwan. In particular, it aims to examine the factor structure of a Chinese version of the Center for Epidemiological Study Depression (CES-D) scale. The CES-D scale is widely used as a measure of mental health in research on stress and social support, as well as on social gerontology.

The 20-item CES-D(Radloff, 1977) scale was developed based on items tapping depressed mood, feeling of guilt and worthlessness, feeling of helplessness and hopelessness, psychomotor retardation, loss of appetite, and sleep disturbance. Four items were worded in the positive direction to break tendencies toward response set as well as to assess positive affect. According to Radloff and Locke (1986), factor analysis of the scale usually revealed a four-factor structure. It included a positive affect factor, a depressed affect factor, a somatic-retarded activity factor and a fourth factor reflecting interpersonal difficulty. Although the scale was not designed to discriminate various subtypes of clinical depression, it was expected to identify the presence and the severity of depressive symptomatology (Radloff & Locke, 1986). This scale has been used in many community studies as an indicator of well-being. Several studies have attempted to investigate the factor structure of the CES-D scale.

Krause and Markides (1985) using maximum likelihood exploratory factor analysis found four factors that correspond to the factor structure found by Radloff (1977) and Roberts (1980) of the CES-D scale. Ensel (1986) also factor analyzed the CES-D scale and reproduced the four-

factor model. Lin (1989) using principal components factor analysis with orthogonal rotated solution to identify a 16-item Chinese version of the CES-D scale (four positive affect items were excluded from factor analysis) and found three factors (i.e. affective mood, somatic-retarded activity, and interpersonal problem) that were consistent with findings from other studies.

However, these previous research on the structure CES-D scale has generally conducted using exploratory factor analysis (EFA) approach which could not specify a prior factorial patterns. And orthogonality of factors was often arbitrarily imposed in the analysis. Thus factor solutions from different studies showed substantial discrepancy for some items. Liang, Van Tran, Krause and Markides (1989) have provided a succinct comparisons among different factor solutions of the CES-D. The limitations of EFA approach in construct validation have been well documented. These include limitations on defining a testable model, yielding unique factorial solutions, assessing alternative models, and adequately test factorial invariance across multiple groups (Long, 1983). Unlike exploratory factor analysis, confirmatory factor analysis approach (CFA) does not have these limitations, therefore, is considered more powerful test of factorial validity.

Only recently, there were studies using CFA approach to investigate the factorial structure of the CES-D scale. In a study comparing generational differences in the structure of the 12-item CES-D scale among Mexican American, Liang et al. (1989) found that a three-factor model (i.e. depressed affect, somatic-retarded activities, and positive affect factor) to fit observed data adequately. The 12-item CES-D scale examined by Liang et al. did not include the two interpersonal problem items, thus their results gave partial support to the factor structure of the

CES-D proposed by Radloff. In a study of elderly women, Thompson and Heller (1990) briefly reported a three-factor (i.e. positive affect, negative affect, and behavior symptoms) structure of the 20-item CES-D to have adequate fit. Their findings were consistent with other studies with the exception that they combined the two interpersonal items with the negative affect items to form a negative affect factor. However, Thompson and Heller did not include detail information on the process of model-fitting in their article.

Based on findings from these studies on the CES-D, one can conclude that the stability of the four-factor structure was generally supported. However, no study has yet demonstrated the original 20 items of the CES-D can be accounted for fully by the four-factor model using confirmatory factor analysis strategy. One major study of the CES-D scale using covariance analysis approach focused on 12 items of the original 20 items because some items did not pass reliability test (Liang et al., 1989). Because many empirical studies were and are still using the summated score of the 20-item CES-D as an indicator of mental health (Ensel & Lin, 1991; Levy, Derby & Martinkowski, 1993; Roberts & Bengtson, 1993). More investigations using confirmatory factorial analysis are needed to establish the factorial validity of the full 20-item CES-D scale. Therefore, the present study will use confirmatory factor-analytic approach to examine the factor structure of the Chinese version of the 20-item CES-D scale.

In addition, the present study proposes a single second-order factor to account for the correlations among four primary factors. Since previous dimensionality research on the CES-D has generally conducted using exploratory factor analysis approach, it focused on primary level components and often assumed independent (orthogonal) factors

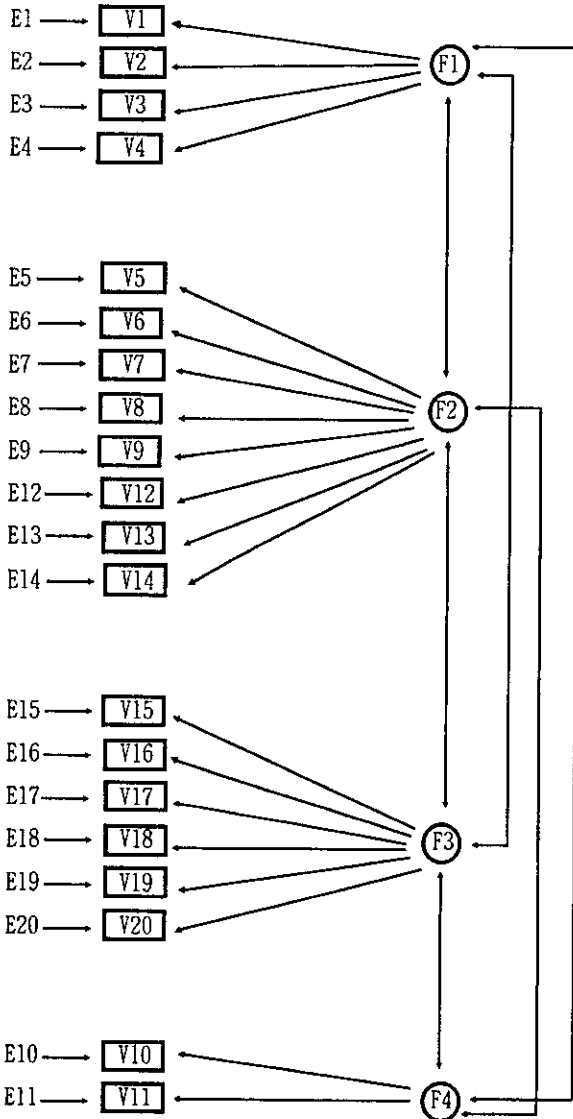
structure to explain the covariation of depressive symptoms. Tanaka and Huba (1984) have pointed out that incorrectly imposed orthogonality can produce results that the levels of simple structure after varimax rotation are poor. This may be part of reason that research findings did not reveal complete consensus concerning factorial structure of the full 20-item CES-D scale. With the hierarchical confirmatory factorial analysis one can test the appropriateness of the factor independence assumptions. If these primary level factors of the depression scale are found to be highly intercorrelated, then this indicates the possible presence of higher order constructs. In practice, a single sum score from the CES-D is often used to represent the degree of depressive symptoms. Thus, it is critical to establish empirically a framework of the CES-D that has a single second-order construct accounting for interrelationships among primary level factors of depressive symptoms.

The present study will address three major research questions: (a) Is the Chinese version of the CES-D scale measuring the same underlying dimensionality of psychological distress among the elderly in Taiwan? that is, can the four-factor primary level model of the CES-D obtained in the United States be replicated in Taiwan? (b) can a single second-order factor model explain the interrelationships among four primary level factors, and (c) if so, can the second-order factorial model be validated successfully across another independent sample of the elderly?

II . Analysis

This study uses confirmatory factor-analytic (CFA) approach to assess both primary and secondary factor structure of the Chinese version of the CES-D scale. The CFA model specifies the relations of the

Figure 1. The Hypothesized Four-Factor Structure of the CES-D Scale (F1: positive affect; F2: depressed affect; F3: somatic activities; F4: interpersonal difficulty).



observed measures to their posited underlying constructs, with the constructs allowed to intercorrelate freely. Figure 1 presents a diagram of the proposed four-factor model of the CES-D scale. The four factors are positive affect (F1), depressed affect (F2), somatic/retarded activities (F3), and interpersonal difficulty (F4). The specific items associated with these four factors can be found in Table 1. In the present study, the schematic presentation of structural equation model follows the Bentler-Weeks representation system (i.e. EQS notation, Bentler, 1992). Thus, in Figure 1 the Vs indicate the observed variables while residuals associated with the measurement of the observed variables are designated as Es. Fs refer to the latent constructs (i.e. factors).

In addition, one-way arrows represent structural regression coefficients and thus indicate the impacts of F latent constructs on the observed V variables. The sourceless one-way arrows pointing from the Es indicate the impact of random measurement error on the observed Vs. And finally, the curved two-way arrows linking all possible pairs of factors suggest that the factors are intercorrelated. Since in the CFA model one is more interested in testing the significance of each factor loading than the significance of the factor variances. Therefore, the factors, rather than one of the factor loadings, were given unit variances for the purpose of setting the scale metric of the latent variables (Bollen, 1989). Anderson and Gerbing (1988) have also recommended that in confirmatory measurement models it is more appropriate to fix the variances of the latent construct to unity than to fix the pattern coefficient for one indicator of each latent construct at 1.0.

Factorial validity of the CES-D was tested using analyses of covariance structure. The maximum likelihood estimates of the models

were obtained from the EQS program (Bentler, 1992). It is important that the evaluation of model fit in the analyses of covariance structure should base on multiple criteria (Anderson & Gerbing, 1988; Cudeck & Browne, 1983; Mulaik et al., 1989; MacCallum, Roznowski, & Necowitz, 1992; Tanaka, 1993). That is, the model fitting strategies should reflect the meaningfulness of the substantive theory, the significance of statistical fit as well as parsimony of the model.

Therefore, the hypothesized four-factor model was proposed based on past research on the CES-D scale, several criteria were then used to evaluate the overall model fit in this study. These include the χ^2 likelihood ratio statistic, the relative χ^2 likelihood ratio (χ^2/df ratio), the Bentler-Bonett's Normed Fit Index (NFI), the Bentler's Comparative Fit Index (CFI). The chi-square likelihood ratio statistic measures the fit between the sample covariance and the fitted covariance matrix. A nonsignificant chi-square indicates the hypothesized model fits the observed data adequately. However, the χ^2 statistic is very sensitive to sample size and deviation of multinormality assumption. In practice, it is often found that a good fit based on a large sample size may result in a significant chi-square. Hence, alternative indices of goodness-of-fit should be used. One alternative is the χ^2/df ratio which attempts to take sample size into consideration.

The goodness-of-fit indices of the NFI and CFI are based on comparisons to a null model. They range from 0 to 1. Usually, values of NFI and CFI greater than .90 are desirable. According to Bentler (1990), the CFI is not influenced by sample size and avoids the underestimation of fit found with the normed fit index. In addition to the overall goodness-of-fit indices, assessment of model adequacy should also be examined by evaluating parameter estimates. Convergent validity of the

factorial model can be assessed by determining whether each indicator's estimated pattern coefficient on its posited underlying construct is significant (Anderson & Gerbing, 1988; Byrne, 1994).

Recently, MacCallum et al. (1992) have call attention to the problem of capitalization on chance in model modifications in covariance structure analysis. They argue that stability and cross-validity are critical issues in model fitting process. Cudeck and Browne (1983) also argue for the importance of cross-validating the covariance structure models. Therefore, to provide more stringent test of the proposed structure of the CES-D, the final factorial model obtained from the analysis will be cross-validated using another independent sample.

III. Method

1. Respondents

The data for the present investigation came from a study of social support and psychological well-being among older people conducted in 1992 in Taiwan. A stratified three-stage area probability design was used to draw a sample of noninstitutionalized adults 60 years of age and older that reflected the different geographic regions and degree of economic development in the island of Taiwan. Face to face interviews were conducted by trained interviewers in respondents' home. The CES-D scale was used as one of indicator of psychological well-eing in that study (Chou, 1993). A total of 1460 respondents were interviewed.¹

Since more than 95% of respondents were either married or widowed, the present study will not include those respondents who were divorced or single. The final sample for this analysis consisted of 1379

elderly who were married or widowed at time of interview. Among 1379 elderly, seventy-two percent were married and 28% were widowed. Fifty-three percent were men and 47% were women. The ages of the respondents in this final sample ranged from 60 to 95, with a mean age of 69.4. Most respondents (43%) were illiterate, another 35.5% had no more than a grade school education. Thirty-two percent had monthly family income less than NT\$ 5,000, twenty percent between NT\$ 5,000 and 9,999, 17% between NT\$ 10,000 and 15,000, and the remaining 31% had more than NT\$ 15,000.

To establish the cross-validity of the CES-D factor structure, the sample was randomly split into two subsamples. After excluding cases with missing data, the effective sample size was 662 for the calibration sample and 661 for the validation sample.

2. Instrumentation

The Center for Epidemiological Studies Depression (CES-D, Radloff, 1977) scale is a 20-item scale which measures symptoms of depression over the past week on a four-point scale. The scale was intended to assess depression in general population. The CES-D was translated into Chinese using back-translation procedure. Because the most common spoken language in daily life in Taiwan is Taiwanese rather than mandarin Chinese, and majority of current cohort of older adults in Taiwan either can not understand or are not comfortable to communicate with mandarin Chinese. Therefore, two stages of translation were involved. First, the CES-D was translated into Chinese, then the Chinese version was translated into Taiwanese. Same procedure was followed in each of the two translation stages.

The CES-D was translated into Chinese by the author. The Chinese version was then back-translated by another bilingual professional into English, which was compared with the original version. This back-translation process continued until these two professionals both agreed that the accuracy of Chinese translation was achieved. Next, the Chinese version was translated into Taiwanese by members of research team (including the author) who were proficient in Taiwanese, and version was finalized after agreement was reached among the members. One remark is in order regarding back-translation process. Because some items were more difficult to translate into Chinese and Chinese into Taiwanese, the agreement on accuracy of translation was based on semantic equivalence rather than literal translations of the original items. Lin's (1989) study of depressive symptomatology in China using the CES-D scale also adopted semantic equivalence criterion.

IV. Results

1. Test of the hypothesized four-factor model

The hypothesized four-factor model with independent measurement errors (M1, Figure 1) was examined using data from the calibration sample. As indicated in Table 1, the χ^2 likelihood ratio was highly significant with a value of 613.103 ($df = 164$). Since the χ^2 likelihood ratio is highly sensitive to sample size, with a sample size of 662 it is expected that this ratio will be significant which indicates that the proposed model does not perfectly fit the observed data. However, other fit indices suggested that the proposed model had an adequate fit. The relative χ^2/df ratio was 3.738. The NFI and CFI for this model were .913 and .935 which suggested more than 91% of variance in the observed data was explained by the proposed model.

Table 1. Standardized Maximum Likelihood Estimates for CES-D Scale Items and Goodness-of-Fit indexes (The Four-Factor Model, M1)

parameters	Calibration sample (n=662)
Variances and covariance of first-order factors	
[F1, F1]	1.000*
[F2, F2]	1.000*
[F3, F3]	1.000*
[F4, F4]	1.000*
{F1, F2}	-.277
{F1, F3}	-.243
{F1, F4}	-.221
{F2, F3}	.873
{F2, F4}	.743
{F3, F4}	.752
First-order factor loadings	
F1 Positive Affect	
v1 Felt as good as others	.485
v2 Happy	.837
v3 Enjoyed life	.868
v4 Hopeful	.693
F2 Depressed Affect	
v5 Bothered	.646
v6 Blues	.756
v7 Sad	.882
v8 Fearful	.808
v9 Lonely	.719
v12 Failure	.587
v13 Depressed	.819
v14 Crying spells	.768
F3 Somatic Activity	
v15 Poor appetite	.578
v16 Hard to concentrate	.726
v17 Talked less	.575
v18 Restless sleep	.599
v19 Couldn't get going	.816
v20 Everything an effort	.740
F4 Interpersonal Difficulty	
v10 People disliked me	.812
v11 People unfriendly	.611
Goodness-of-fit indexes	
χ^2 likelihood ratio	613.103 (df=164)
χ^2/df likelihood ratio	3.738
Bentler-Bonett NFI	.913
Bentler's CFI	.935

* denotes a parameter fixed to 1.0 in the analysis.
note: all parameter estimates are significant at .05 level.

In addition to the overall goodness-of-fit indices, assessment of model adequacy should also include the examination of parameter estimates. Ideally, the specified parameters should be statistically significant. Table 1 showed that standardized factor-loadings were all significant and were all greater than .48, indicating each indicator did in fact have substantial nonzero coefficient on its posited latent construct.

As expected, Positive Affect (F1) had significant negative correlations with the other three factors while Depressed Affect (F2), Somatic Activity (F3), and Interpersonal Difficulty (F4) factors were positively correlated. Correlations among factors ranged from -.221 to .873. To establish discriminant validity among factors with high correlations (i.e. among F2,F3,F4), each of the estimated correlation parameters among these factors was constrained to 1.00, then three constrained models were reestimated separately. According to Anderson and Gerbing (1988), a chi-square difference test on the values obtained for the unconstrained and constrained model can be evaluated. If the unconstrained model has a significantly lower chi-square, then the factors are not perfectly correlated and discriminant validity is achieved. The results showed that, compared to the unconstrained model, all three constrained model had a significantly higher chi-square values. For example, the model with correlations between Depressed Affect and Somatic Activity factors constrained to 1.00 had a χ^2 value of 755.079 which yielded a large χ^2 difference score of 141.976 ($df = 1$) with the unconstrained model. The values of NFI and CFI for this constrained model were .893 and .914 as compared to the .913 and .935 of the unconstrained model.

Therefore, although there were high correlations among three of the four proposed factors, above analysis demonstrated the discriminant

validity among these factors. However, moderate to high covariations among these four factors indicates that there is possibility of the presence of a general second-order factor to account for these covariation. The higher-order factorial model will be examined in the next section.

Although researchers generally agree that model-fitting process should be justified on substantive as well as statistical criterion, in practice there is no one agreed upon guideline. For instance, some researchers suggest that confirmatory factor analysis models involving psychological constructs often require the researcher to specify correlated error variances in order to obtain a well-fitting model. Correlated error variances can be substantively meaningful in reflecting minor, possibly sample-specific data covariation not explained by the target factors in the model (Byrne, Shavelson & Muthen, 1989; Tanaka & Huba, 1984).

On the other hand, Anderson and Gerbing (1988) consider that the use of correlated measurement errors can be justified only when they are specified as a prior. MacCallum et al. (1992) also raised the issues of generalizability of models resulting from data-driven modification of an initial model. They strongly recommended that the final model obtained from modification process should be cross-validated to independent samples.

Since the assessments of the initially hypothesized four-factor model (M1) already showed an adequate fit, and the examination of the modification indices based on the Lagrange Multivariate (LM) test revealed that no substantial gain could be obtained from respecification of M1 model by including correlated measurement errors. To avoid the problem of capitalization on chance, no further modification was

proceeded. Therefore, the hypothesized four-factor model is accepted as the CES-D structure model.

2. Validating second-order factor model of the CES-D

In this section, a single second-order factor, called Depression, is further proposed to explain the covariation among the four first-order factors obtained in the previous section (M2, Figure 2). This implies that the structure of the CES-D scale is multidimensional at the first-order but unidimensional at the second-order level. And the variance of each first-order factor was decomposed into the variance accounted for by the second-order factor and a residual variance (Ds in Figure 2) not accounted for by the higher-order factor. To establish the unit of measurement for the higher order factor, the variance of the second-order factor (i.e. F5) was fixed to unity. With this model specification, we expect to find empirically the model to have a satisfactory overall goodness-of-fit with all first-order as well as second-order factor loadings to be statistically significant.

The second-order factorial model was rejected for this sample, with a chi-square value of 614.637 ($df = 166$, $p < .001$), as fully adequate. However, NFI and CFI were .913 and .935 which indicated that the model was compatible with the observed data. The relative χ^2/df ratio was 3.703. Table 2 presents the standardized maximum likelihood estimates of parameters. All of first- and second-order factor loadings are significant at the .05 level. Loadings from the second-order factor, Depression, to the four first-order factors vary between .280 to .936, with the Depressed affect and Somatic Activity having highest loading. The Interpersonal Difficulty factor also had high loading(.800) on the Depression factor. The residual variances of first-order factors ranged from .032 to .323.

Figure 2. The Second-order Factorial Model (F1: positive affect; F2: depressed affect; F3: somatic activities; F4: interpersonal difficulty; F5: second-order general depression).

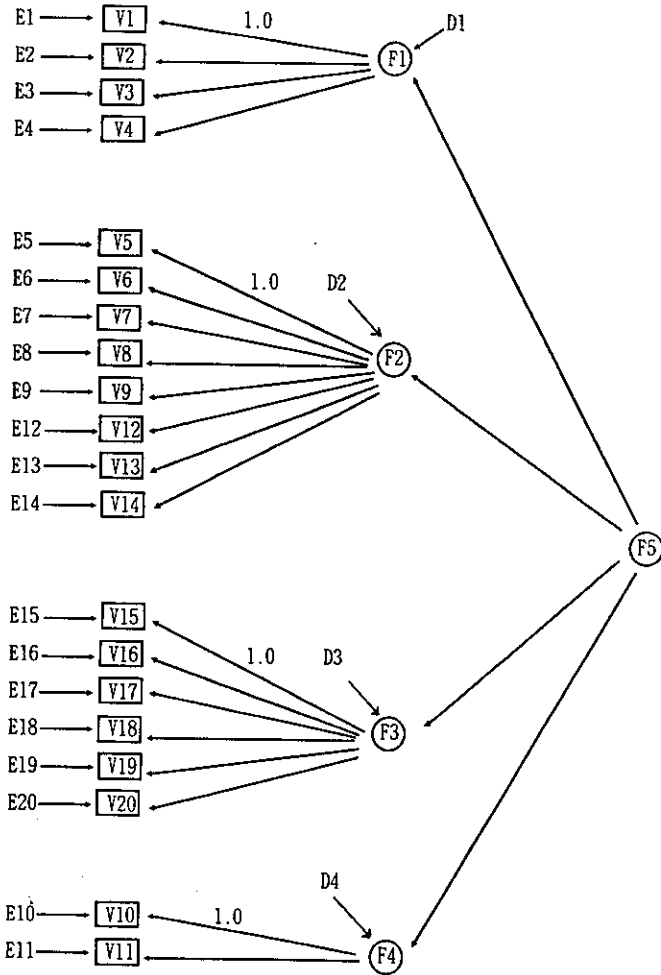


Table 2. Standardized Maximum Likelihood Estimates in the Second-Order Factor Model (M2)

parameters	Calibration sample (n=662)
Second-order Factor Loadings	
F5 General Depression Factor	
F1 Positive Affect	-.280
F2 Depressed Affect	.934
F3 Somatic Activity	.936
F4 Interpersonal Difficulty	.800
First-order Factor Loadings	
F1 Positive Affect	
v1 Felt as good as others	.485*
v2 Happy	.835
v3 Enjoyed life	.869
v4 Hopeful	.693
F2 Depressed Affect	
v5 Bothered	.646*
v6 Blues	.757
v7 Sad	.882
v8 Fearful	.807
v9 Lonely	.719
v12 Failure	.588
v13 Depressed	.819
v14 Crying spells	.769
F3 Somatic Activity	
v15 Poor appetite	.579*
v16 Hard to concentrate	.726
v17 Talked less	.575
v18 Restless sleep	.600
v19 Couldn't get going	.815
v20 Everything an effort	.740
F4 Interpersonal Difficulty	
v10 People disliked me	.810*
v11 People unfriendly	.612
Goodness-of-fit indexes	
χ^2 likelihood ratio	614.637 (df=166)
χ^2/df likelihood ratio	3.703
Bentler-Bonett NFI	.913
Bentler's CFI	.935

* denotes a parameter fixed to 1.0 in the original solution.
note: all parameter estimates are significant at .05 level.

Overall, based on various fit assessment indices, the second-order factorial model can be concluded to well represent the observed data on the CES-D scale. Thus, the results of the second-order factorial model are consistent with the conceptualization of depression as a general construct with multiple interrelated yet distinct dimensions.

3. Cross-validation of the CES-D Factor Structure

To address the issue of cross-validation of the findings from the calibration sample in the previous analysis, the four first-order factors model was first tested for its invariance across a second independent sample (i.e. validation sample) then followed by the cross-validation of the second-order model. In the first-order structure, all factor-loadings, measurement error variance, and factor correlations were constrained equal across calibration and validation samples. For the higher-order structure cross-validation, all factor loadings (including first- and second-order), measurement error variance (variance of E), and the residual variances of first-order factors (Ds), were constrained equal across these two independent samples, and then tested statistically in a simultaneous analysis of the data using multigroups procedures of the EQS program.

The results of the first-order factorial model showed the constrained model to be well-fitting with values of .903 and .928 of NFI and CFI. The chi-square likelihood ratio was 1329.591 with 374 degree of freedom. The relative χ^2/df ratio was 3.555. The LM test statistics in the EQS program provides information on whether the hypothesized equality of the specified constraint can be held statistically. Although the multivariate LM χ^2 statistics indicated the five specified equality constraints did not hold (i.e. [F1, F3]; [V7, F2]; [E5, E5]; [E13, E13]; [E11, E11]), their chi-square values were all rather small (less than

5.815). Based on the above evaluation, the four-factor first-order factorial model is concluded to cross-validate successfully.

The findings of cross-validation of the second-order factorial model revealed a relative χ^2/df ratio of 3.574; the values of NFI and CFI were .902 and .927. This indicates the model fit is adequate. The multivariate LM χ^2 statistics indicated several equality constraints involving measurement errors and residuals of the first-order factor were significant (i.e. [E5, E5]; [E11, E11]; [E13, E13]; [D2, D2]; [D3, D3]). However, their χ^2 values were again all relatively small (less than 5.894). No substantial improvement on model fit can be gained by releasing these equality constraints. Therefore, the second-order factorial model is also considered to cross-validate satisfactorily across validation sample.

V. Summary and Conclusions

Radloff (1977) had proposed that the CES-D consisted four dimensions: positive affect, depressed affect, somatic-retarded activity, and interpersonal difficulty. Although these four factors were generally obtained in several major investigations of the properties of the scale, no study attempts to directly fit all of the 20 items into the hypothesized four-factor model. The purpose of this study was to assess whether the hypothesized four-factor model well represented the structure of the Chinese version of the 20 item CES-D scale. Furthermore, this study proposed a single second-order factor to account for the covariations among primary factors. In addition to use confirmatory factor analysis approach, this study also provided the evidence of the cross-validity of the proposed models.

The results indicated that the four-factor model proposed by Radloff was consistent with the observed data. Not only global fit indices showed adequate model fit, all factor loadings were also significant and greater than .45. Thus, convergent validity of this factorial model for the 20-item CES-D was demonstrated. The moderate relationships between positive affect and other three factors were consistent with the argument in studying mental health that positive and negative affect should not be viewed as an opposite end of a single dimension. Instead, they should be treated as different yet related concepts.

Correlations among depressed affect, somatic activity, and interpersonal difficulty factors were quite substantial. However, further discriminant validity analysis indicated that these three highly correlated factors were indeed different constructs. These moderate to substantial covariations among four primary factors suggested the existence of higher factorial structure. A single second-order factor, called general depression, was proposed to account for the covariation among four primary factors. The second-order factorial model also showed an adequate fit to the data. With exception of moderate loading from the second-order factor to positive affect factor, all the first- and second-order loadings were again substantial.

Several researchers (Aneshensel et al., 1983; Krause & Markides, 1985; Liang et al., 1989) have argued that two interpersonal items in the CES-D should be excluded because conceptually they seemed to confound the lack of social support resources with depressive affect, and empirically they were found to correlate weakly with other factors. Contrary to previous research findings in Western culture showing interpersonal factor having the weakest relationship with other three primary factors and with a second-order factor of depression, the results

of this study indicated that interpersonal factor had strong relationships with depressed affect and somatic factors when the primary level of the factor structure was examined. And when the second-order structure was posited, the interpersonal factor again had strong loading on the second-order general depression factor. This findings suggest that for this sample of elderly in Taiwan these two items of interpersonal difficulty did reflect some degree of respondents' psychological distress. Is it possible that in this culture psychological distress is commonly manifested in the perception of interpersonal relations? This issue merits further study. It is also worth in the future research to examine the differential relationships among these four primary factors and other major concepts in stress and support research.

The presence of a second-order factor implies that the CES-D scale can be conceptualized as a generalized measure of psychological distress with four unique yet related subdimensions. In practice, it implies that the use of total score of the scale as an indicator of general psychological adjustment is justified. Further examination of the second-order loadings found that four primary factors did not contribute equally in defining the second-order factor, with positive affect factor contributing much less than the other three factors. In other words, the second-order factor mainly reflected the respondents' psycho-social distress, though it also assessed the absence of positive affect. Since the CES-D scale was originally developed to include not only the more general measures of distress but also the assessment of positive affect, these results gave the strong empirical support to the validity of the CES-D scale as a general measure of psychological well-being.

The four primary factorial model as well as the single second-order factorial model were all successfully cross-validated in another

independent sample. Therefore, it can be concluded that the Chinese version of the CES-D scale has demonstrated sound psychometric properties as a general measure of psychological well-being. It has a factorial structure with four factor at primary level as originally proposed by the Radloff and a single factor at the second-order level. However, cautions should be taken when making generalizations from findings of this study. Although this study provided a much more stringent test of the CES-D structure than previous research, it is always possible that alternative models other than the one specified in this study will be equally consistent with the observed data. In addition, the model was derived from the elderly sample. More research is needed to substantiate the results of the present study and to investigate factorial invariance of the CES-D among various social groups in Taiwan.

Notes

- 1 Due to some clerical complications, exact response rate could not be computed. However, the demographic comparisons between the present sample and the 1990 census data showed comparable results with one exception that the present sample contained less proportion of older adults aged between 60 and 64, and more proportion of the elderly aged between 70 and 79.

References

Anderson, J. C. & Gerbing, D. W.

- 1988 "Structure Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin* 103: 411-423.

Aneshensel, C. S., Clark, V. A. & Frerichs, R. R.

- 1983 "Race, Ethnicity, and Depression: A Confirmatory Analysis," *Journal of Personality and Social Psychology*, 44: 385-398.

Bentler, P. M.

- 1990 "Comparative Fit Indexes in Structural Models," *Psychological Bulletin* 107: 238-246.
- 1992 *EQS: Structural Equations Program Manual*. Los Angeles: BMDP Statistical Software.

Bollen, K. A.

- 1989 *Structural Equations with Latent Variables*. New York: Wiley & Sons.

Byrne, B. M.

- 1994 *Structural Equation Modeling with EQS and EQS/Windows*. Thousand Oaks, California: Sage.

Byrne, B. M., Shavelson, R. T. & Muthen, B.

- 1989 "Testing for the Equivalence of Factor Covariance and Mean Structure: The Issue of Partial Measurement Invariance," *Psychological Bulletin* 105: 456-466.

Chou, Y.

- 1993 Effects of Social Support from Various Sources on Psychological Well-Being among the Elderly in Taiwan. Paper Presented in the 1993 World Congress of the World Federation for Mental Health, Makuhari, Japan.

- Cudeck, R. & Browne, M. W.
1983 "Cross-validation of Covariance Structures," *Multivariate Behavioral Research* 18: 147-167.
- Ensel, W. A.
1986 "Measuring Depression: The CES-D Scale," in N. Lin, N. A. Dean & W. A. Ensel (eds.), *Social Support, Life Events, and Depression*. New York: Academic.
- Ensel, W. A. & Lin, N.
1991 "The Life Stress Paradigm and Psychological Distress," *Journal of Health and Social Behavior* 32: 321-341.
- Krause, N. & Markides, K. S.
1985 "Employment and Psychological Well-being in Mexican American Women," *Journal of Health and Social Behavior* 26: 15-26.
- Levy, L. H., Derby, J. F. & Martinkowski, K. S.
1993 "Effects of Membership in Bereavement Support Groups on Adaption to Conjugal Bereavement," *American Journal of Community Psychology* 21: 361-381.
- Liang, J., Van Tran, T., Krause, N. & Markides, K. S.
1989 "Generational Differences in the Structure of the CES-D Scale in Mexican Americans," *Journal of Gerontology* 44: s110-120.
- Lin, N.
1989 "Measuring Depressive Symptomatology in China," *Journal of Nervous and Mental Disease* 177: 121-131.
- Long, J. S.
1983 *Confirmatory Factor Analysis*. Beverly Hills, CA: Sage.
- MacCallum, R. C., Roznowski, M. & Necowitz, L. B.
1992 "Model Modifications in Covariance Structure Analysis: The Problem of Capitalization on Chance," *Psychological Bulletin* 111: 490-504.

Mulaik, S. A., James, L. R., Alstine, J. V., Bennett, N., et al.,

1989 "Evaluation of Goodness-of-Fit Indices for Structural Equation Model," *Psychological Bulletin* 105: 430-445.

Radloff, L. S.

1977 "The CES-D Scale: A Self-report Depression Scale for Research in the General Population," *Applied Psychological Measurement* 1: 385-401.

Radloff, L. S. & Locke, G. Z.

1986 "The Community Mental Health Assessment Survey and the CES-D Scale," In M. M. Weissman, J. K. Myers & C. E. Ross (eds.), *Community Surveys of Psychiatric Disorders*. New Brunswick, N. J.: Rutgers University Press.

Roberts, R. E.

1980 "Reliability of the CES-D Scale in Different Ethnic Contexts," *Psychiatry Research* 2: 125-134.

Roberts, R. E. & Bengtson, V.

1993 "Relationships with Parents, Self-Esteem, and Psychological Well-Being in Young Adulthood," *Social Psychology Quarterly* 56: 263-277.

Tanaka, J. S.

1993 "Multifaceted Conceptions of Fit in Structural Equation Models," In K. A. Bollen & J. S. Long (eds.), *Testing Structural Equation Models*. Newbury Park, CA: Sage.

Tanaka, J. S. & Huba, G. J.

1984 "Confirmatory Hierarchical Factor Analyses of Psychological Distress Measures," *Journal of Personality and Social Psychology* 46: 621-635.

Thompson, M. G. & Heller, K.

1990 "Facets of Support Related to Well-Being: Quantitative Social Isolation and Perceived Family Support in a Sample of Elderly Women," *Psychology and Aging* 5: 535-544.

兄弟姊妹教育相似性的探究： 兩種「迷迷模型」的應用

薛承泰*

壹、前言

「迷迷模型」(MIMIC Model) 為 LISREL (LInear Structural RELationships) 的一種，由社會學家 Hauser 和經濟學家 Goldberger 於 1971 年提出，乃為「多重原因—多重指標模型」(multiple-indicators and multiple-causes model) (Hauser and Goldberger 1971) 的簡稱 (圖 1)。

本文主要目的在應用「一般迷迷模型」及其特例「等級迷迷模型」(Ordered MIMIC Model) 於「兄弟姊妹教育取得的相似性」(sibling resemblance in educational attainment) (簡稱手足教育相似性) 的分析。為了重新檢驗 Benin and Johnson (1984) 和 Hauser and Wong (1989) 有關「手足教育相似性」幾個假設的爭辯，作者採用 NLSY (National Longitudinal Survey of the Youth, 1979-1992) 資料，建構手足配對 (sibling pairs) 樣本，然後仿造 Hauser and Wong (1989) 先應用「一般迷迷模型」於「手足教育相似性」的分析。最後，作者提出以

* 臺灣大學社會學系副教授

「學位」(degrees) 成就來取代「求學年數」(years of schooling) 的測量方式，應用「等級迷迷模型」(Ordered MIMIC Model) 來重新驗證「手足教育相似性」的相關假設。

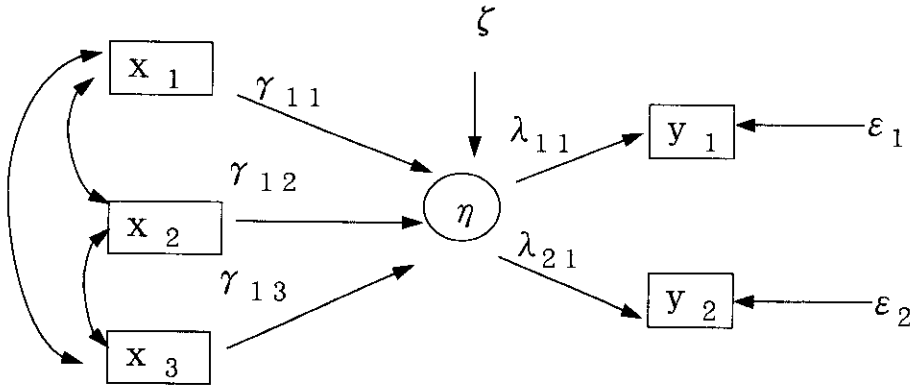


圖 1 一般迷迷模型

文章的主要結構如后：第二節主要敘述家庭背景與手足相似性的研究課題。第三節則專注於手足教育相似性的相關研究之回顧，以及 Benin and Johnson (1984) 和 Hauser and Wong (1989) 有關「手足教育相似性」爭辯的主要內容。第四節介紹「一般迷迷模型」和「等級迷迷模型」的統計原理，主要包括了模型結構和估計方法的簡介；此外也說明本文的模型選擇方法。第五節說明 NLSY「手足樣本」和變項的特性。第六節與第七節為主要分析部分，分別為「一般迷迷模型」與「等級迷迷模型」的分析結果與解釋。最後，第八節為本文之結論。

貳、家庭背景與手足相似性

除了教育學者外，許多社會科學家對於教育現象也都感到興趣；其中社會學家慣常從社會制度、區位環境、個人特質、以及家庭背景來了解教育的不均等 (educational inequality) 現象，他們著手實證研究分析社會文化或政策上可能帶來的教育歧視或障礙，並嘗試對弱勢族群教育機會的增加與教育品質的提昇做出貢獻。儘管學校教育 (schooling) 從20世紀以來不斷擴張，也經常有大規模的改革；很不幸的，教育的危機一直普遍存在於各個社會，甚至不同的時間「危機」以不同的面貌呈現出來 (Hurn, 1993)。社會學家理解到，在民主社會中的許多政策、制度與區位環境是可以改變，以符合新的社會需求或趨勢，以期減輕不均等現象；然而，相對於制度來說，個人特質與家庭背景能透過「外力」來改變的卻是非常有限，尤其兩者之間具相當的互動性，更降低了掌握問題的可能性。

簡言之，要釐清個人特質與家庭環境的影響，某種程度上就是在區別「先天遺傳」(nature) 和「後天教養」(nurture) 的作用，而這也是生物醫學以及精神病理學一直在嘗試的工作。他們常對同卵雙胞胎 (identical twins)、異卵雙胞胎 (fraternal twins)、具收養關係的兄弟姊妹、以及不同環境生長下的同胞兄弟姊妹作研究，比較這些不同遺傳程度與成長環境，以了解「兄弟姊妹」來自「先天」和「後天」的差別。雖然不同學科研究各有其強調的重點，卻一致接受「先天」和「後天」兩者影響力的同時存在，尤其雙胞胎之間顯著的差異 (包括智力、個性、精神病理狀況等) 一再被證實，不但肯定「後天」影響力，更有助於我們對成長環境中非為兄弟姊妹所共享的因素之認識 (Dunn and Plomin, 1991)。

雖然遺傳因素對個人的影響相當重要，但這並非社會學家能力所及；甚至對於「後天」部份，儘管社會學家不遺餘力地探討「家庭背景」(family background) 如何影響子女，到目前為止的認識仍是很有

限的。例如，家庭背景如何影響個人的教育取得 (educational attainment) 此課題，一直是社會學「地位取得研究」(status attainment research) 的重點，Blau and Duncan (1967) 應用路徑分析 (path analysis) 來探討美國男子職業地位取得，他們發現家庭背景指標 (父親職業與教育) 對於個人教育的解釋力並不高 ($R^2 = .26$)。往後的研究除了仿造 Blau and Duncan 的作法之外，也有許多研究者應用多元迴歸或「對數迴歸模型」(或簡稱邏輯模型) (logit model) 來分析教育取得或學位取得的影響因素，基本上家庭背景對於個人教育的變異量的解釋都不超過50%。

從這些研究的發現來看，家庭背景對個人成就影響仍有限，而這些所謂的家庭背景影響能否為家庭中的子女所共享，也值得爭議。Hauser and Featherman (1976) 試圖了解這個問題，以美國全國男性樣本為對象，計算受訪者和其長兄 (oldest brother) 求學年數的相關，發現從這個世紀初各年齡層所顯示的相關介於.52至.60。如果以家庭背景當作兄弟姊妹所共享的影響因素，那麼這個數字只能當作是家庭背景能解釋子女教育取得變異的上限。

爲了更精確來理解家庭背景或成長環境對子女的影響，有學者開始研究家庭的內部過程，早期著名的研究爲 Lindert (1977, 1978)，他探討家庭結構與個人特質 (包括性別、年齡、出生序、兄弟姊妹個數、出生間隔等) 和日後表現 (包括疾病和成就) 的關連。Zajonc (1976)，Zajonc and Markus (1975)，以及 Zajonc et al. (1979) 則探討子女數量和出生序 (birth order) 對智力測驗和學校成績的影響。在社會學界中也有 Blake (1981, 1989) 探討家庭子女數對子女品質 (包括教育、智力等) 的影響；Hauser and Sewell (1985) 分析家庭中子女出生序和教育成就的關係；Mare and Tzeng (1989) 則更進一步考慮家庭的生命週期 (以父親在小孩出生時的年齡為指標) 可能帶給子女的差別影響。

雖然上述這些研究對於家庭環境影響力的分析呈現不同的關照點，它們在資料的應用上卻有一共同的特色即研究的樣本單位為個人 (individual)，因此所分析的家庭背景影響力，所反映的乃為家庭間 (between-families) 差異和個人成就 (如教育取得、收入) 變異之間的關係，而就算是涉及家庭內部過程的研究，也是以受訪者為主並同時提供家庭中其他人的資料，用來理解一般的家庭內部差異。總之，這種研究所採取的觀點，雖然並沒有忽視家庭內 (within-family) 兄弟姊妹的差異、父母對不同子女成就存在可能的差別影響力、以及手足間互動所產生不同的作用 (Dunn and Polmin, 1991; Hauser and Sewell, 1985)，但這些研究因為資料結構的限制，以個人為研究單位，而未能分析家庭間差異的同時也掌握到家庭內部的差異。無論如何，這些研究在強調「家庭間」差異的傳統中，揭示了另一個研究面向——「家庭內」的變異——的重要性，卻是不爭之事實。

參、手足教育相似性研究的回顧

社會學中最早從事手足研究者當屬 Jencks 和其同僚 (1972)，在他們的研究中，探討家庭背景對兄弟智力 (IQ) 和教育取得的影響是主要目的之一。他們認為能解釋子女教育取得的主要因素，包括遺傳、「可觀察」與「潛藏」的共同因素 (通常指家庭背景或環境) 和個人特質等。其中潛藏的共同因素可視為家庭中的兄弟姊妹們所共享卻難以測量的部份，是過去社會學研究想要研究卻又拙於分析的課題；至於潛藏的個人因素就是所謂的「殘差項」，通常被當作隨機而無所不在。Jencks 等人認為，採用手足配對資料即可用來分析家庭內部的差異，因為那些未被控制的共同環境因素，可視為手足間所共享而可以相互抵消掉的部份；所以手足間教育取得的差異，可以由個人的智力、可觀察的個人特質和潛藏的個人特質所決定 (1972)。不久，Olneck and Bills (1979) 針對出生序和家庭規模 (family size) 對兄弟間

的智力、教育、職業以及收入差異作分析，也是採用相同的原則；這些分析對於往後手足社經地位相似性的研究具有相當的影響 (Hauser and Sewell 1986)。

Jencks 等人的研究 (1972, 1979) 注重父母對子女成就的差別影響力，是社會學中手足研究的先期典型。大約就在那時，還有經濟計量學者和教育計量學者如 Behrman et al. (1977), Chamberlain and Griliches (1975), Olneck (1977) 和 Sewell and Hauser (1977) 等人運用不同手足資料來作分析；值得注意的，這些「手足研究」(Sibling studies) 除了研究議題有別於傳統階層或不平等研究，所使用的資料結構也不一樣。在資料使用方面，這些過去的研究通常只從「家戶」(household) 中抽取一人為分析單位，或假定樣本有此特性；而「手足研究」中的受訪者必須存在兄弟姊妹關係，也就是分析單位為「戶」而每一戶中不只一個人。有時候這種以「戶」為單位的資料結構比一般以「人」為單位的資料結構複雜；不但蒐集較為困難，在統計分析技術上也較為煩瑣。然而，為了簡化模型並顧及分析的方便，上述「手足研究」大都採用「手足配對」(sibling pairs) 作為分析單位。¹

「手足研究」雖然在70年代就有相當規模，有關教育取得相似性的研究卻至80年代中，才再度引起不少社會學者的興趣，並競相引介新的統計模型來解析家庭間與家庭內的差別影響力 (Hauser 1988)。這個研究成為「時尚」，一部份源於 Benin and Johnson (1984) (簡稱 BJ) 在「教育社會學」(Sociology of Education) 期刊上的一篇論文，後來被 Hauser and Wong (1989) (簡稱 HW) 指出其不當而引發其他人對此課題的注意。

由於早期手足研究所使用的資料大都是「兄弟配對」(brother pairs)，樣本數量小而且非來自全國的隨機樣本；BJ 指出，以「兄弟配對」樣本來研究手足相似性，忽略了：(1) 家庭對子女教育成就的影響，可能會因子女性別而不同，(2) 兄弟配對的教育相似性程度可能不

同於其它的手足配對，(3) 手足年齡較接近者其相似性也可能較高。BJ 爲了澄清這些問題，在美國 Nebraska 州著手蒐集兩筆手足配對樣本 (NASIS 和 Lincoln data)，他們在論文中提出的主要假設爲：兄弟教育相似性要高於其它類手足配對，而導致高相似性的主要原因爲手足間的交互影響 (cross-sibling effects)。因此，他們認爲兄弟配對迴歸的「餘值共變量」(residual covariance) 應該會比其他手足配對的餘值共變量大。² 最後，因爲兄弟間有較強的互動與學習，BJ 認爲兄弟會比姊妹「獨立」，亦即兄弟在教育地位的取得上，比較不會受到家庭背景的影響。

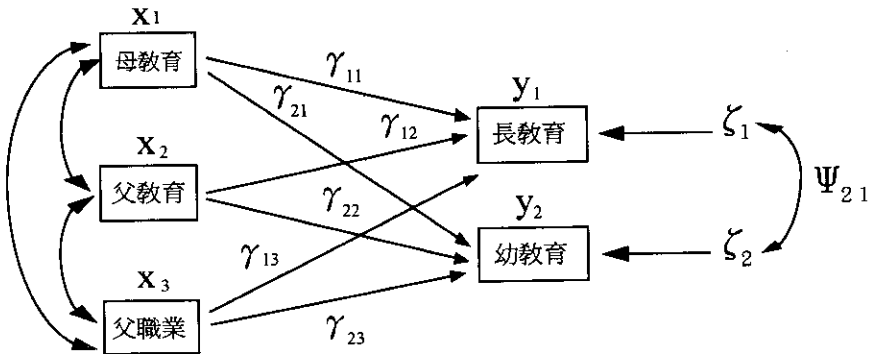


圖 2 Benin and Johnson 聯立迴歸模型

爲了驗證這些說法，BJ 運用聯立迴歸 (見圖 2) 來分析手足配對的教育成就和家庭背景的關係，他們將手足配對分爲兄弟、姊妹、兄妹、姊弟四類，分別以聯立迴歸來分析並比較。果然，他們發現了「兄弟配對」迴歸的餘值共變量大於其他手足配對，而「姊弟配對」迴歸的餘值共變量是四組配對中最小的；然而在另一方面，家庭社經地位對這四種手足配對教育成就的影響卻是相同的，兄弟比姊妹獨立的說法並不能接受。因此，BJ 對手足教育相似性下結論，謂「除了共

同的家庭影響力外，手足教育的相似性以兄弟最高，姊弟為最低；而兄弟教育的相似性主要來自於兄對弟的影響」(1984: 17)，他們並且把兄對弟的影響這個結論當作是手足間「角色模範」(role-modeling) 假設的一個印證。

HW 對於 BJ 上述的結論不以為然，他們認為 BJ 使用聯立迴歸模型來驗證上述的假設，並不適當。他們建議用迷迷模型(如圖 1) 重新來檢證 BJ 的假設。HW 認為，應用迷迷模型在手足相似性的研究要比聯立迴歸適當，其原因除了迷迷模型方便於參數值的假設檢定與模型比較之外，在同時分析四組手足配對樣本時，即所謂「多重群體」(multi-groups) 的比較，更能凸顯 LISREL 模型的方便與精簡。³ 尤其，HW 指出 BJ 最大的錯誤，在於將聯立迴歸餘值的共變量當成了手足的相互影響力；因為迴歸餘值的共變量，除了包含手足互動因素外，還有手足間「潛藏」的共同因素，手足的相互影響力充其量只是其中的一部分。嚴格一點來說，BJ 所應用的迴歸模型並沒有直接分析出所謂的「手足相互影響力」。

於是 HW 應用迷迷模型來重新分析了 Nebraska 樣本，並且發現以下兩點和 BJ 的結論不同：(1) 潛藏共同因素的影響 (Ψ) (即 ζ 的變異量) 在姊弟配對中為最小；但是，兄弟配對的 Ψ 卻不是最高的，它和其它兩組配對的 Ψ 並無差異。這意含著，除了姊弟配對外，潛藏共同因素對其他三種手足配對有一樣的影響力，這和 BJ 的發現有相當的距離。這也同時意含著，用兄弟配對樣本為代表來說明一般手足的相似性，並不會太離譜。(2) 家庭背景對子女教育產生差別影響力，家庭背景對「幼者」(弟或妹) 的影響力約為對「長者」(兄或姊) 影響力的四分之三。最後，HW 為了進一步驗證「手足相互影響力」，以 Kalamazoo 兄弟配對樣本來分析，結果並沒有如 BJ 所言，兄對弟之間有「角色模範」的現象存在，而是兄弟之間存在等量的相互影響力。⁴

肆、分析方法

4.1 「迷迷模型」的簡介

迷迷模型可分解成兩部分，第一個部分有如迴歸分析(見圖 1 的左半部)，包含一個「內衍變項」(endogenous variable)， η ，和多個稱為「因」(causes)的「外衍變項」(exogenous variables)， x 's，以及一個「干擾項」(disturbance)(或稱為「殘餘項」)， ζ ，所組成。這個部分和一般多元迴歸不同之處，在於該內衍變項「無法直接觀察」(unobservable)，也就是一般所謂的「潛在變項」(latent variable)。迷迷模型的另一個部分類似「因素分析」(factor analysis)，由前述的「潛在變項」和其「指標」(indicators)， y 's，以及對應每個指標變項的殘餘項， ε 's，所組成(如圖 1 的右半部)。準此，迷迷模型的數學形式可以按「迴歸」和「因素」兩個部分，分別敘述如下：

$$\eta = \gamma_1 x_1 + \gamma_2 x_2 + \dots + \gamma_q x_q + \dots + \gamma_Q x_Q + \zeta. \quad (1)$$

方程式(1)說明潛在變項 η 乃為 $x_1, \dots, x_q, \dots, x_Q$ 和 ζ 的線性組合；若以矩陣來表示，則將(1)簡化為：

$$\eta = \gamma x + \zeta. \quad (2)$$

其中 η 是一個潛在變項， x 則是 $Q \times 1$ 向量(vector)所構成的外衍變項， γ 是 $1 \times Q$ 個「結構迴歸係數」(structural regression coefficients)， ζ 則是「結構干擾項」(a structural disturbance)。迷迷模型的另一個部分有如因素分析，可表示如下：

$$y_1 = \lambda_1 \eta + \varepsilon_1,$$

$$\cdot \quad \quad \cdot$$

$$\cdot \quad \quad \cdot$$

$$y_p = \lambda_p \eta + \varepsilon_p,$$

$$\cdot \quad \quad \cdot$$

$$y_p = \lambda_p \eta + \varepsilon_p \quad (3)$$

等式 (3) 共包含 P 個線性方程式，每一個 y_p 都是 η 的指標變項，它們是「可觀察」(observable) 或「可被測量」(measured) 的變項。當然，每一個 y_p 都無法完全代表 η 而都有一相對應的殘差項， ε_p ($p = 1, \dots, P$)，通常稱為「測量誤差」(measurement errors)，並且假設它和 x 's 之間是「獨立的」(independent)，以及 ε_p 彼此之間也是「獨立的」。若以矩陣來表示，等式 (3) 可改寫成比較精簡的形式：

$$y = \lambda \eta + \varepsilon \quad (4)$$

其中 y 是 $P \times 1$ 個指標變項， λ 是 $P \times 1$ 個「因素負荷量」(factor loadings)， ε 則是 $P \times 1$ 個殘餘項。就如一般 LISREL 模型的前設 (assumption)，迷迷模型仍假定 $E(\zeta, \varepsilon) = E(\varepsilon, \eta) = E(x, \zeta) = 0$ 。根據這些前設條件，等式 (2) 和等式 (4) 可合併成爲下列的簡式 (reduced form)：

$$\begin{aligned} y &= \lambda (\gamma x + \zeta) + \varepsilon, \\ &= (\lambda \gamma) x + (\lambda \zeta + \varepsilon), \\ &= \Pi x + v. \end{aligned} \quad (5)$$

其共變矩陣則爲：

$$\Sigma = E(v v') = E[(\lambda \zeta + \varepsilon)(\lambda \zeta + \varepsilon)'] = \lambda^2 \Psi + \theta \quad (6)$$

其中 Ψ 是 ζ 的變異量，而 θ 是 ε 的變異量，爲一對角線矩陣。由於等式 (5) 的係數 Π 共有 $P \times Q$ 個元素，可以用 $\gamma_1, \dots, \gamma_Q$ 和 $\lambda_1, \dots, \lambda_Q$ 等共 $P+Q$ 個參數來表示；而等式 (6) 的 Σ 是個對稱 (symmetric) 共變矩陣，共有 $P(P+1)/2$ 個不同元素，可以用 $\Psi, \lambda_1, \dots, \lambda_Q$ ，和 $\theta_{11}, \dots, \theta_{pp}$ ，等共 $1+2P$ 個參數來表示之。換句話說，簡式中共出現了 $(P \times Q) + P(P+1)/2$ (稱爲 A) 個參數，它可以依據結構方程式中 $\gamma_1, \dots, \gamma_Q, \lambda_1, \dots, \lambda_Q, \theta_{11}, \dots, \theta_{pp}$ 和 Ψ ，共 $2P+Q+1$ (稱爲 B) 個結

構參數表示。當 P 和 Q 都大於或等於 2 時 (也就是模型中有兩個「因」及兩個「指標」以上), A 大於 B , 迷迷模型即可「辨識」(identified) — 包括「完全辨識」(just-identified) 與「過度辨識」(over-identified) 兩種情形。例如 $P = Q = 2$ 時, $A = B = 7$, 模型即為「完全辨識」, 沒有留下任何自由度 ($df = 0$); 當 $P = Q = 3$ 時, $A = 15, B = 10$, 模型為「過度辨識」, 自由度為 $A - B = 5$ 。而當 $P = 2$ 與 $Q = 3$ 時, $A = 9, B = 8$, 自由度為 1。當迷迷模型可「辨識」時, 在「多元常態」(multinormality) 的前提下, 即可用「最大概似推估」(maximum likelihood estimation) 來求取參數的「有效估計值」(efficient estimates) (Hauser and Goldberger 1971), 關於迷迷模型的最大概似推估法則與過程, Jöreskog and Goldberger 曾為文詳細說明 (1975: 632-638)。

由於 η 是個「潛在變項」, 其測量單位無從確定 (undetermined), 因此在模型推估時需要將它「標準化」(standardization) 以除去不定性。通常的做法就是將 Ψ (也就是 ζ 的變異量) 設定為 1, 或是將某一項 λ_p 設定為 1, 使 Ψ 的單位和其相同, 其它的 λ 's 則以 λ_p 的乘數解釋之, 後者也正就是本研究所採用的方式。⁵

4.2 等級迷迷模型簡介

通常「教育取得」是以「學年數」來代表, 可是「學位」所代表的教育成就也是不可忽視的; 前者適用於「一般迷迷模型」的分析, 而後者因「學位」屬「順序層級」(Ordinal Level) 的測量, 違反「多元常態」的前題, 並不是典型 LISREL 模型。⁶ 因此, 當迷迷模型中「內衍變項」(如 y 's) 為「順序層級」時, 作者稱它為「等級迷迷模型」(Ordered MIMIC Model)。這種模型的統計原理, 首先由 Muthén (1979) 所提出, 他並且推出一統計軟體 LISCOMP (Analysis of Linear Structural Relations Using a Comprehensive Measurement Model: A Program for Advanced Research) 可供「結構共變模型」(Structural

Covariance Model) 中「內衍變項」為連續、順序或「限控」(censoring) 的情況。然而，該軟體的設計因企圖過強(如其軟體名稱)，使用上沒有 LISREL 方便與易於了解，因此使用的人並不多。雖然如此，對於「階序迷迷模型」的應用仍是很好的選擇，而 Muthén 本人也有特別為文說明(1979)。以下針對 Muthén 所提供的方法，對「等級迷迷模型」做簡單介紹。

等級迷迷模型的設定基本上和一般迷迷模型相似，最大的不同在於手足教育 (y_1 和 y_2) 這個「內衍變項」是非連續性；因此，家庭背景和手足教育之間不再是線性關係。依據 Muthén (1979) 的說法，可假設 y_i 是來自於 y_i^* 這個潛在變項，而且 y_i^* 是呈常態分配 $-\infty < y_i^* < +\infty$ ； y_i 則是 y_i^* 的具體可測量變項， y_i 會等於 j ，如果 $C_{j-1} \leq y_i^* < C_j$ ($j=1,2,3,\dots,J$)，其中 C_1, C_2, \dots, C_{j-1} 是為「門檻」(thresholds)。在本研究的例子中，求學年數 12 和 16 分別來代表「高中畢業」與「大學畢業」兩個門檻。在 LISCOMP 中，用 α 's 來表示這些門檻，至於其它的參數及其前設都和一般迷迷模型相同；因此，階序迷迷模型可以等式 (7) 至 (9) 表示。

$$y_1^* = \alpha_1 + \lambda_{11} \eta + \varepsilon_1, \quad (7)$$

$$y_2^* = \alpha_2 + \lambda_{21} \eta + \varepsilon_2, \quad (8)$$

$$\eta = \gamma x + \zeta. \quad (9)$$

根據上面三個等式，可得出簡式如下(如前將 λ_{11} 設定為 1)：

$$\begin{aligned} y_1^* &= \alpha_1 + \lambda_{11} \gamma x + \lambda_{11} \zeta + \varepsilon_1 \\ &= \alpha_1 + \gamma x + \varepsilon_1^*, \\ \text{Var}(\varepsilon_1^*) &= \text{Var}(\varepsilon_1 + \zeta) = 1 + \Psi. \end{aligned} \quad (10)$$

$$\begin{aligned} y_2^* &= \alpha_2 + \lambda_{21} \gamma x + \lambda_{21} \zeta + \varepsilon_2, \\ &= \alpha_2 + \lambda_{21} \gamma x + \varepsilon_2^*, \\ \text{Var}(\varepsilon_2^*) &= \text{Var}(\varepsilon_2 + \lambda_{21} \zeta) = 1 + \lambda_{21}^2 \Psi. \end{aligned} \quad (11)$$

Muthén 採用「波比模型」(Probit Model) 的原理，先將等式 (10) 和 (11) 標準化，即 $\text{Var}(\varepsilon_1^*) = \text{Var}(\varepsilon_2^*) = 1$ 。準此，在多元常態的前提下「二元波比模型」(Bivariate Probit Model) 能夠以「聯合機率函數」(joint probability function) 方式表示：

$$\Pr(y_1 = 1, y_2 = 1 | \mathbf{x}) = F(\alpha + \Gamma \mathbf{x}, \Lambda \Psi \Lambda' + I). \quad (12)$$

其中 $F(\cdot)$ 是二元常態分配函數，包含了一個「均數結構」($\alpha + \Gamma \mathbf{x}$) 和一個「共變結構矩陣」($\Lambda \Psi \Lambda' + I$)。Muthén (1979) 認為在等式 (6.6) 的 $\alpha, \Gamma, \Lambda, \Psi$ 等參數，能夠由「波比最大概似推估」來做最初的估計，然後再帶入「共變結構矩陣」中，再來獲取等式 (10) 和 (11) 中 $\alpha, \lambda, \gamma, \Psi$ 的估計，同時「門檻」參數的估計，也可由「均數結構」中獲得。詳細過程請參閱 Muthén (1979, 1983, 1984) 或 LISCOMP (Muthén 1987) 使用手冊。⁷

4.3 模型選擇方法

由於實證資料的複雜性，社會科學的量化研究經常以建構模型的方式來「簡化」它，並且藉以說明假設命題中變項間的關係；一般來說，模型乃依據假設所描述變項間的關係來複製資料結構。所以一個好的模型不在於「完全適合」(perfect fitting) 資料，而是在可以接受的機率水準下「適合」即可，這也就是統計上所謂的「適合度」(goodness-of-fit) 的檢驗。由於能夠完全適合資料者，往往就是資料本身或相當複雜的「模型」，而這並非理論或模型所企求的；因此，通常要經過「模型選擇」(model selection) 的過程，研究者嘗試從「適合的」(fitted) 模型中找出較精簡的模型，這個模型被稱為「最可以接受的模型」(the most plausible model)。當然，在精簡的原則下，此模型對資料的「適合度」不見得是「最高」的。

在HW的研究中，採用的模型選擇方式，除了每個模型有「適合度」(以 L^2 表示) 與 p 值的檢驗，並且採用慣常的「概似比測試」(Likelihood Ratio Test) (以 χ^2 表示) 作為模型間的比較；⁸ 最重要的，

他們採用 Raftery (1986) 的說法，使用 *bic* (Bayesian Information Criterion) 作為模型選擇的主要依據。⁹ 簡單地說，他們先在一連串「巢狀」(nested) 模型中以概似比測試做比較，最後選擇 *bic* 值為最小者為「最可以接受的模型」，當然這種模型選取方式不是沒有爭議的。¹⁰ 為了複製 HW 模型以為說明之用，在本文的表三和表四仍沿用 HW 的模型選擇方式，即同時報告 χ^2 ， L^2 和 *bic*，然後再作進一步討論。

伍、資料與變項

5.1 資料來源與手足樣本的建立

本研究採用美國 NLSY (1979~1992) 資料，該資料是由 Ohio State University 「人類資源研究中心」(Center for Human Resources Research) 所蒐集。該調查於1979年首度執行，以全美當年14至21歲青年為對象，共有12,686位青年男女。此後，每年針對此樣本調查，到目前仍未中斷。本研究所使用的資料，為「人類資源研究中心」於1994年首度發行的光碟資料庫系統。由於這筆資料是建立在全美以「戶」(Household) 為單位的隨機樣本，戶中凡屬於該年齡層者均為調查的對象；因此，許多受訪者彼此間存在手足關係，這就成為 NLSY 手足樣本的來源。

本文所使用的手足配對乃作者所建立，這筆資料具有許多過去手足資料所沒有的優點：(1) 它是來自全國隨機樣本，樣本數量多，共有5,856位受訪者彼此間有手足關係，他們分別來自於2,444個家庭。(2) 樣本年齡差距最多7歲，用來分析「手足相似性」可避免時間所造成的干擾。¹¹ (3) 最後，NLSY 包含受訪者逐年教育、職業、婚姻、生育…等資料，對於手足相似性的許多課題，提供了相當具有分析價值的變項。

由於家庭中因兄弟姊妹的多寡 (sibling size) 可以形成不同配對數 (例如，三位可構成三個配對、四位則有六對、五位有十對…)，在分析中需要使用複雜的加權技術，才足以使不同的配對數量呈現「戶」的代表性。本研究抽取戶中只有兩位為手足關係的受訪者，做為分析的對象，主要的原因除了免除加權問題，也為了避免引進太多和主題較遠的課題 (例如兄弟姊妹數、出生序、以及出生間隔的問題)。當然，應用這筆手足配對資料的發現，其推論的範圍只能限於家庭中有兩位子女者。最後，本文分析的對象，共有1,689個配對，分別來自1,689個家庭；其中「兄弟配」有467對、「姊妹配」有397對、「兄妹配」有428對、「姊弟配」有397對。

表 1：手足樣本之基本變項描述 (平均數和標準差)

母親教育	父親教育	父親職業地位*	長教育	幼教育
1. 全體(N=1,689)				
11.02 (3.10)	11.11 (3.78)	3.625 (2.12)	12.92 (2.54)	12.91 (2.42)
2. 兄弟(N=467)				
10.97 (3.30)	10.82 (4.07)	3.641 (2.13)	12.63 (2.68)	12.57 (2.62)
3. 姊妹(N=397)				
11.12 (2.95)	11.25 (3.53)	3.682 (2.12)	13.12 (2.25)	13.23 (2.17)
4. 兄妹(N=428)				
10.85 (3.15)	11.16 (3.67)	3.545 (2.08)	12.64 (2.58)	12.91 (2.34)
5. 姊弟(N=397)				
11.19 (2.96)	11.25 (3.78)	3.635 (2.17)	13.37 (2.54)	12.98 (2.45)

* 父親職業地位採 Duncan 的社會經濟指標 (SEI)，取其十分之一來報告 (後表亦同)。

表 2：迷迷模型變項之相關矩陣：四組手足配對

1. 兄弟

父教	1.0000				
母教	.6009	1.0000			
父職	.5422	.3830	1.0000		
兄教	.4065	.4039	.4193	1.0000	
弟教	.4174	.4039	.4272	.6279	1.0000

2. 姊妹

父教	1.0000				
母教	.5930	1.0000			
父職	.5530	.3803	1.0000		
姊教	.3193	.3485	.3096	1.0000	
妹教	.3415	.3292	.3562	.5051	1.0000

3. 兄妹

父教	1.0000				
母教	.6231	1.0000			
父職	.4989	.3741	1.0000		
兄教	.4323	.3523	.4045	1.0000	
妹教	.3822	.3695	.3490	.6189	1.0000

4. 姊弟

父教	1.0000				
母教	.5938	1.0000			
父職	.5088	.4416	1.0000		
姊教	.4047	.3578	.3240	1.0000	
弟教	.4372	.3662	.3974	.4977	1.0000

5.2 變項的使用

按照 Hauser and Wong (1989) 的做法，本研究以母親的教育程度(以年計)、父親的教育程度(以年計)、受訪者14歲時父親的職業地位(以 Duncan's SEI 來測量)作為家庭社經背景變項，這些變項在迷迷模型中分別以 x_1, x_2, x_3 來代表，做為影響子女(兄弟姊妹)教育的共同家庭環境指標，它們的影響力參數則以 $\gamma_{11}, \gamma_{12}, \gamma_{13}$ 來表示。然而，不同於迴歸分析，子女的教育並不直接成為這些「因」(x's)的「果」。在迷迷模型中是以「子女的教育環境」(η) (family-related educational background)，這個潛在變項當作「果」；並且容許潛藏的共同因素(ζ)的存在，而潛藏的共同因素的影響力則以 ζ 的變異量 Ψ 來表示。同時 η 也可以當作一個「因素」來看，它的兩個指標就是「手」(長者)與「足」(幼者)的教育取得(以 y_1 和 y_2 來表示)。它們的係數參數 λ_{11} 和 λ_{21} 則分別代表前述「子女教育環境」的負荷量。我們通常將其中一個參數固定為 1 (如 $\lambda_{11} = 1$)， λ_{21} 的估計值則成為 λ_{11} 的對比， λ_{21} 就可以解釋為家庭對手足教育的差別影響力。如果 $\lambda_{11} = \lambda_{21}$ ，則表示家庭背景對手足教育的影響力是一樣的，也就是說，沒有家庭內部的差別影響力存在。最後 ε_1 和 ε_2 則分為「長」與「幼」的殘餘變項，分別代表潛藏的個人因素(見圖 1)。本研究中迷迷模型所需要的資料，包含了母教(x_1)、父教(x_2)、父職(x_3)、長教(y_1)、幼教(y_2) 五個變項；有關這些變項的平均數、標準差、和相關矩陣分別報告在表 1 和表 2。

六、一般迷迷模型的分析結果

6.1 全體手足樣本

首先，我們以全體樣本(共1,689個配對)為對象，應用一般迷迷模型於手足教育取得相似性的分析。在模型 A，即基本模型(baseline model)中，除了 $\lambda_{11} = 1$ 是為了賦予潛在變項(η)測量基準外，沒有

對其它參數予以任何「限定」(restrictions)；該模型保留了較多的參數，可作為接下來幾個假設模型的基礎。從表 3 模型 A 的報告，該模型呈現相當好的適合度 ($L^2 = .42, df = 2, p = .809$)。模型 B 乃模型 A 附上 $\gamma_{11} = \gamma_{12}$ 限定；模型 A 和 B 的比較，即可檢定「母親教育和父親教育對子女教育具有相同影響力」的假設。根據表 3 的報告，模型 B 也呈現出相當好的適合度 ($L^2 = 1.15, df = 3, p = .765$)；而且和模型 A 比較的結果，模型 B 並沒有因多了一個自由度而敗壞了模型的適合，至少在慣常使用的統計水準下 A 和 B 在模型的適合上並沒有顯著的差異 ($\chi^2 = 0.7, df_a = 1, p > .05$)。然而，因模型 B 釋出一個參數，換取了一個自由度，模型 B 要比模型 A 精簡。

表 3：「一般迷迷模型」假設檢定與模型選擇：全體手足配對

模型	L^2	df	p	bic	Contrast	χ^2	df_a	p
A. Baseline	.42	2	.809	-14.4				
B. $A + \gamma_{11} = \gamma_{12}$	1.15	3	.765	-21.1	(B)-(A)	.7	1	>.05
C. $B + \lambda_{21} = 1$	1.48	4	.830	-28.2	(C)-(B)	.3	1	>.05

將模型 B 加上 $\lambda_{21} = 1$ 的限定成為了模型 C (見表 3)，該模型用來測試家庭背景對「長」與「幼」的教育影響力是否相等。結果模型 C 也相當令人滿意 ($L^2 = 1.48, df = 4, p = .830$)，和模型 B 比較的結果也沒有達到統計上的差異 ($\chi^2 = 0.3, df_a = 1, p > .05$)。這三個巢狀模型的比較，不論是用概似比測試還是用 bic ，都是以模型 C 為「最可以接受模型」。

6.2 四組手足樣本

前一節的分析是以全體手足配對為對象，只有長幼之分；本節的分析則區別手足間的性別，產生兄弟、姊妹、兄妹、姊弟四組手足樣本，作為分析的對象。這是典型多重群體 LISREL 模型，也可以說是

四組聯立迷迷模型。若按 HW 模型選擇的方法與程序，結果如表 4，在諸多模型比較下，模型 H 因為呈現最小 *bic* 值成爲了「最可以接受模型」。¹² 若根據表 4 的報告，所得到的結論有：(1) 四組手足配對中，家庭背景指標沒有樣本「選擇性偏差」(selection bias) (Hauser and Wong 1989) (模型 B)；¹³ (2) 個人因素對教育的影響，不論性別與長幼別均相同 (模型 C)；(3) 家庭背景中的任一指標對不同手足配對具有相同影響力 (模型 D)；(4) 母親教育和父親教育對子女教育具有相同的影響力 (模型 E)；(5) 沒有家庭內部的差別影響力 (模型 F & G)；(6) 「潛藏的共同因素」對手足教育的影響，在四組配對中均相同 (模型 H)。

表 4：「一般迷迷模型」假設檢定與模型選擇：四組手足配對

模型	L^2	<i>df</i>	<i>p</i>	<i>bic</i>	Contrast	χ^2	<i>df_d</i>	<i>p</i>
A. Baseline	6.6	8	.576	-53				
B. $A + \Phi(I N)$	33.6	26	.146	-160	(B)-(A)	27	18	>.05
C. $B + \theta^E$ (I N & E Q)	48.1	33	.043	-197	(C)-(B)	14	7	=.05
D. $C + \Gamma(I N)$	62.2	42	.023	-250	(D)-(C)	14	9	>.05
E. $D + \gamma_{11} = \gamma_{12}$	63.1	43	.025	-256	(E)-(D)	.9	1	>.05
F. $E + \Lambda(I N)$	66.6	46	.025	-275	(F)-(E)	3.5	3	>.05
G. $F + \lambda_{21} = 1$	71.0	47	.014	-278	(G)-(F)	4.4	1	<.05
H. $G + \Psi(I N)$	81.4	50	.003	-290	(H)-(G)	10.4	3	<.025

IN:invariant, EQ:equal

很明顯的，在表 4 的報告中，概似比測試 (χ^2) 達統計顯著水準.05者有模型 G 與 F 的比較，達顯著水準.025者有模型 H 與 G 的比較，我們究竟應依據 *bic* 還是概似比測試的結果來選擇模型呢？依據

概似比測試則增加 type I (α) 誤差，若依據 *bic* 測試則增加 type II (β) 誤差，這是見仁見智的問題。簡單地說，模型 G 與 H 的假設接受與否，面臨不同統計值的考驗而將有所差別，而比較關鍵的假設在於「潛藏的共同因素」對手足教育的影響，在四組配對中是否均相同(即模型 H)。以下將進一步來探討這兩個假設。

首先，將表 4 基本型中的四組配對加上 $\lambda_{21} = 1$ 的設定，這是驗證「家庭內部的差別影響力是否相等」的假設。從表 5.1 模型 B 的報告得知該模型有良好的適合度 ($L^2 = 11.4, df = 12, p = .50$)，和基本模型比較的結果，該模型雖然多了四個自由度，卻未使模型的適合度有顯著減少的跡象 ($\chi^2 = 4.8, df_d = 4, p > .05$)。因此，可證明「家庭內部的差別影響力」是不存在的，表 4 模型 G 應該是可以接受的。

若將基本模型加上 Ψ 為「恆定」(invariant) 的設定，可用來驗證「潛藏共同因素對四組配對有一致的影響力」的說法。從表 5.1 模型 C 的報告，得知該模型會折損基本模型原有的適合度 ($\chi^2 = 11.2, df_d = 3, p < .025$)。所以，「潛藏共同因素」對四組手足配對的影響力並不相同。至於「潛藏家庭因素的影響力」對兄弟配對是否為最大，而對姊弟配對為最小呢？模型 D 說明，兄弟配對和兄妹配對的 Ψ 相等，姊妹配對和姊弟配對的 Ψ 相等，而該模型並沒有折損適合度 ($\chi^2 = 0.3, df_d = 2, p > .05$)；意指「潛藏共同因素的影響力」，「兄弟」和「兄妹」相似，「姊妹」和「姊弟」相似，但對前者(「兄弟」和「兄妹」)的影響力大於對後者(「姊妹」和「姊弟」)的影響力。

從表 5.2 關於基本模型係數的推估也可知道，家庭內父母對子女的影響力，不分長幼和性別，其差別不大 (λ_{21} 分別為 .991, 1.010, .827, 1.069)；此外，「潛藏共同因素」的影響力當「長」為男性時(如「兄弟」對和「兄妹」對)較大 (Ψ 分別為 2.625 和 2.913)，而以「長」為女性時(如「姊妹」對和「姊弟」對)較小 (Ψ 分別為 1.628 和 1.634)。

表 5.1：「手足教育取得相似性」假設檢定與模型選擇：「一般迷迷模型」

模型	L^2	df	p	bic	Contrast	χ^2	df_d	p
A. Baseline	6.6	8	.576	-53				
B. $A + \lambda_{21}=1$	11.4	12	.500	-78	(B)-(A)	4.8	4	>.05
C. $A + \Psi(I N)$	17.9	11	.084	-64	(C)-(A)	11.3	3	<.025
D. $A + \Psi_{11} = \Psi_{33}$, $\Psi_{22} = \Psi_{44}$	6.9	10	.732	-67	(D)-(A)	0.3	2	>.05

IN:invariant

表 5.2：「手足教育取得相似性」：「一般迷迷模型」基本模型參數估計

	Ψ	θ^e_{11}	θ^e_{22}	γ_{11}	γ_{12}	γ_{13}	λ_{21}
兄弟	2.625	2.735	2.494	.089	.178	.332	.991
姊妹	1.628	2.620	2.219	.062	.151	.210	1.010
兄妹	2.913	2.141	2.387	.154	.123	.287	.827
姊弟	1.634	3.553	2.693	.160	.108	.197	1.069

註：參數估計均達.05顯著水準

綜上而論，關於表 4 中的模型 G 和 H 的比較，經由表 5.1 和 5.2 的說明之後，本研究認為家庭背景對長幼教育取得的影響應該是沒有顯著差異，但潛藏的共同因素效果以兄的配對(包括兄弟與兄妹配對)較大，姊的配對(包括姊弟與姊妹配對)較小。如果按 BJ 的說法，將潛藏的共同因素影響力視為手足的互動效果，那麼上述的發現應指，來自於兄的角色示範作用要強過姊，而且這個示範效果並不因弟或妹而有別。

柒、等級迷迷模型分析結果

傳統地位取得研究以「求學年數」來代表「教育成就」(educational achievement)，在說明與分析上雖較為方便，其隱含著每一年教育是「等同的」(equivalent)說法，畢竟是其缺點。許多學者頗不以為然而嘗試修正，例如 Boudon (1974) 就曾將教育視為一個循序漸進的成就過程；Duncan (1968) 從各年級入學和升級情形來描述不同種族的教育差異；Sewell (1971) 則直接指出「高等教育」和「基礎教育」本質上的不同。事實上，個人在追求教育成就時，通常按照制度的規定，循序漸進來完成各教育階段；因此，許多人並不在乎「在學校共唸了幾年」，而在乎「拿了學位與否」以及「拿了什麼學位」？因為「學位取得與否」和「拿了什麼學位」，不僅象徵個人的能力和成就，也是求職或陞遷，甚至為薪資水準計算所依憑的標準 (Sewell 1971)。此外，教育之所以有階段之分，就足以顯示其在本質與內涵上的差異。所以，將「求學年數」作為「教育成就」的唯一測量並不是很恰當；而以「學位」作為個人教育成就的指標，可能要比「求學年數」來得實際些。因此，本章的分析，將子女的教育分為以「高中畢業或以下」、「大學」肄業 (包括兩年制社區大學或四年大學未畢業)、「大學畢業或以上」三個等級，應用階序迷迷模型來分析四組配對的「手足學位的相似性」，並且重新驗證前節的兩個假設。

在表 6.1 中，模型A為基本模型並呈現很好的適合度 ($L^2 = 6.1$, $df = 8$, $p = .641$)，按前述的程序，分別以模型 B 和 C 來驗證 $\lambda_{21} = 1$ 和 Ψ 在各手足配對組中相等的兩個假設，而這兩個假設也是 Hauser and Wong (1989) 和 Benin and Johnson (1984) 的爭議點。根據概似比測驗結果，模型 B ($L^2 = 7.9$, $df = 12$, $p = .790$) 和模型 C ($L^2 = 6.848$, $df = 11$, $p = .8112$) 都表現了相當不錯的適合度。這兩個模型和基本模型對照的結果，也都未達統計上的顯著 ($\chi^2 = 1.8$, $df_a = 4$, $p > .05$; $\chi^2 = 0.7$, $df_a = 3$, $p > .05$)；因此，對於上述的兩個假設都未能否定，也就是家庭內

的差別影響力並不存在，「潛藏家庭因素的影響力」對於不同手足配對也無不同。後者的結論和一般迷迷模型雖有不同，實際上若以 $\Psi_{11} = \Psi_{33}$ 和 $\Psi_{22} = \Psi_{44}$ 的設定來測試 (如模型 D)，其結果即和一般迷迷模型相符 ($\chi^2 = 0.2, df_d = 2, p > .05$)。

表 6.1：「手足學位相似性」假設檢定與模型選擇：「等級迷迷模型」

模型	L^2	df	p	bic	Contrast	χ^2	df _d	p
A. Baseline	6.1	8	.641	-53				
B. A+ $\gamma_{21}=1$	7.9	12	.790	-71	(B)-(A)	1.8	4	>.05
C. A+ $\Psi(I N)$	6.8	11	.811	-75	(C)-(A)	0.7	3	>.05
D. A+ $\Psi_{11} = \Psi_{33}$, $\Psi_{22} = \Psi_{44}$	6.3	10	.791	-68	(D)-(A)	0.2	2	>.05

IN: invariant

表 6.2：「手足學位相似性」：「等級迷迷模型」基本模型參數估計

	Ψ	τ_1	τ_2	γ_{11}	γ_{12}	γ_{13}	λ_{21}
兄	.557	2.547	3.197	.107	.041#	.016	1.0
弟		2.494	3.283				.985
姊	.489	1.662	2.523	.069	.038	.010	1.0
妹		1.913	2.737				1.182
兄	.659	2.190	2.877	.057	.052	.016	1.0
妹		1.885	2.628				.971
姊	.435	1.826	2.626	.048	.086	.010	1.0
弟		2.078	2.844				1.007

註：參數估計未標示者表示已達.05顯著，#表示未達.05顯著水準

表 7：四組手足配對的各種學位取得機率
(當父母親教育均為12年，父親 SEI 為50時)

	大學或以上	大學肄	高中或以下
兄 / 弟	.19 / .18	.25 / .27	.97 / .97
姊 / 妹	.12 / .05	.37 / .25	.91 / .94
兄 / 妹	.07 / .12	.21 / .31	.92 / .91
姊 / 弟	.17 / .12	.43 / .33	.95 / .95

表 6.2 報告了基本模型的係數推估，雖然該模型所用的子女教育測量為學位而不是學年數；然而，對四組配對中 λ_{21} 之估計值，和表 5.2 十分類似。其它估計值和表 5.2 不同，是因為等級模型的係數屬「波比」(Probit) 型態，再加上模型中有「門檻」參數，其意義沒有一般迷迷模型那麼直接。茲舉一例說明這些係數的意義：假設有一兄弟配對，其父母親教育均為12年，父親職業地位 (SEI) 為50，則兄至少大學畢業的機率為 $F(12 \times 0.107 + 12 \times 0.041 + 5 \times 0.016 - 3.197) = F(-1.34)$ ，查 Z 分配表可知為19%；相同的情形下，弟的機率則為 $F[0.985(12 \times 0.107 + 12 \times 0.041 + 5 \times 0.016 - 3.283)] = F(-1.427 \times 0.985)$ ，其機率為18%。這是按 Probit 算法， $p_j = F(\sum \gamma x - \alpha_j)$ ，其中 $F(\cdot)$ 為標準常態 CDF (cumulative distribution function) (Winship and Mare 1984)。準此條件，四種手足配對的對各種學位取得機率，可整理如表 7。簡單地說，在上述家庭背景下，兄弟完成大學或以上學位的相似性最高，姊弟為其次。

捌、結 論

教育學者、社會學者、以及經濟學者都試圖瞭解家庭背景或環境對子女(教育、職業、經濟等)成就的影響，「手足研究」是其中一個

選擇，它的特色在方法上以解析變異量成分 (variance components) 來呈現共同與個別潛藏因素的影響力 (Chamberlain 1977)，在資料應用上以家戶為分析單位，凸顯「家庭間」與「家庭內」的差異的存在 (Hauser 1988)；此外，還有助於對「遺失變項」在推估所造成偏誤的控制 (Griliches 1977, 1979)。

「一般迷迷模型」最早由 Hauser and Goldberger (1971) 介紹，之後 Jöreskog and Goldberger (1975)，以及 Chamberlain (1977) 都作了進一步的討論；至於應用於「手足教育相似性」的分析，則首為 Hauser and Wong (1989)。本文以「手足研究」中「手足教育相似性」的課題，應用了「一般迷迷模型」與「等級迷迷模型」。應用這兩個模型主要目的是為了澄清 Benin and Johnson (1984) 和 Hauser and Wong (1989) 關於家庭背景對不同子女配對影響的一些說法。在這個研究中，本文除了介紹了「一般迷迷模型」與「等級迷迷模型」及其差異，也使用了 LISREL7 和 LISCOMP 兩個軟體，分別來推估這兩個模型以檢測「手足教育相似性」的幾個假設。

在本研究中「一般迷迷模型」用於內衍變項是連續性時 (如子女教育以學年數來測量)，而「等級迷迷模型」用於內衍變項是順序層次時 (如子女教育以學位來測量)。它們共同的特色在於，能控制家庭社經指標的效應下，還能夠區別「共同的潛藏因素」(ζ)與「個別的潛藏變項」(ε)，而兩者的變異量 (Ψ 和 θ) 分別被視為共同潛藏因素與個別潛藏因素對教育的作用，這是一般迴歸模型所不能的。此外，這兩個模型在 LISREL「多群體」(multi-group) 模型架構下，不同層級的假設檢定可經由「群內」(within-group) 或「群間」(between-group) 參數值的限定來達成。

作者以 NLSY (1979~1992) 最新資料建構手足配對樣本，包括來自家庭中有兩位子女的1,689個配對，相較於過去研究所用的手足資料，在數量、代表性、以及配對的品質與種類來說，都有其優越之

處；然而，本研究的發現僅能推論到部份出生於50年代美國家庭中的兄弟姊妹(也就是只有兩位子女的家庭)。無論如何，這筆資料適合用來檢測 BJ 有關手足教育相似性的說法：雖然家庭社經背景對兄弟姊妹具有相似影響力，四組手足配對中，仍以兄弟的相似性最高，姊弟為最低；兄弟教育的相似性主要來自於兄對弟角色模範的影響。關於這個說法，本研究只支持了家庭社經背景對兄弟姊妹具有相似影響力這部份的敘述，兄弟相似性最高的假設並沒有被支持。

簡單地說，本研究發現，家庭內父母社經地位對子女(不論性別)教育取得(以年為測量)的影響並無明顯的差別；至於家庭中「共同潛藏因素」對手足教育的影響存在些差別，但並不如 BJ 所說的(該因素對兄弟影響為最高)，也不像是 HW 所發現的(該因素對姊弟影響為最弱)，而是對「長」為男性時的影響力大於「長」為女性時的影響力。如果按 BJ 的說法，將潛藏的共同因素影響力視為手足的互動效果，那麼來自於兄的角色示範作用要強過姊，但這個來自「長」的示範效果並不因「幼」的性別而有別。值得注意的，如果是考慮「學位」的相似性時，在家庭社經地位作用恆定下，「共同的潛藏因素」的影響對兄弟姊妹似乎都一樣。

由於本研究採用的樣本，在時間與代表性上和過去研究有所差別，我們不能將這些差異完全歸之於方法或模型的不同。此外，本研究應用於模型的變項並不多，主要原因除了是為了和 BJ 與 HW 兩個研究對話，也便於模型實際運作的解析。事實上，在未來建構手足相似性模型之時，仍有一些變項(如家庭經濟狀況、家庭結構、種族…)值得考慮加入(Hsueh 1992)，畢竟迷迷模型的「因」與「果」指標可以不限本研究所採用的變項。當然，在未來的研究發展也朝向不限於手足配對的分析(Kuo and Hauser 1995, Mare 1993)。

註 釋

- 1 採用手足配對資料的另一個原因，是這種資料結構有助於處理「遺失變項」(omitted variables)在傳統迴歸模型中可能造成的問題(Taubman 1977)。這是由於社會科學的實證研究裡，常會因資料或測量的缺陷，而造成「應該」控制而未被控制的變項未能有效掌握。例如，「家庭背景」常被社會學家認為是影響子女成就的重要因素；然而，「家庭背景」只是理論上存在的「構念」(construct)，它的指標應該是什麼？它們都能被觀察或測量嗎？很遺憾，很難找到令人滿意的答案。許多研究者同意，無論採用哪些變項來代表家庭背景，仍然無法等同家庭背景，而只好將這些遺失歸諸於「未被觀察到」的潛藏因素(Jensen 1980)。這些因素如果沒有妥善控制，將會在迴歸模型中，產生「偏誤的」(biased)估計(Bowles 1972, Griliches 1977, 1979)。
- 2 Benin and Johnson (1984) 認為，弟弟在教育地位的取得上會以哥哥為「模範」(model)；因此，兄弟的相互影響中，哥哥對弟弟的影響要大於弟弟對哥哥的影響。此外，他們也認為同性(same-sex)同胞間的互動會較異性同胞的互動頻繁，也會產生較高的相似性，而其中兄弟配對會比其它配對明顯。
- 3 採用聯立迴歸來處理每一組手足資料(如兄與弟的配對)時，不僅要使用較多的參數，在餘值關連下來比較背景因素分別對兄弟教育取得的影響力時，程序較為複雜；更何況，四組不同手足配對(兄弟、姊妹、兄妹、姊弟)一起考慮時，模型更顯得笨拙。至於從統計的自由度以及參數估計的多寡，亦可看出迷迷模型的精簡(請參考本文第四節或 Hauser and Goldberger 1971有詳細說明)。
- 4 關於手足之間的互動或相互影響力，並沒有在本研究中分析，因為加入這個參數之後，模型在「辨識不足」(under-identified)的情形下，而需加入「工具變項」(instrumental variables)。修改過後的模型已非本研究要探討的迷迷模型。

- 5 本研究所採用的迷迷模型 $P=2, Q=3$ ：其基本模型自由度本應為 1，因為加上 $\lambda_{11} = 1$ 的設定，所以實際的自由度為 2 (見表 3)。
- 6 由於 LISREL8 陸續仍在修改，截至 1995 年底時 SPSS 仍以 LISREL7 為其中之模組發售。而本研究之分析進行於 1994 年底，完成於 1995 年三月，以當時來說 LISREL7 仍是最普遍的版本。按 LISREL7 的應用，「等級迷迷模型」並不符多元常態的前設；不過，若配合 PRELIS 的使用，仍可以進行模型推估。
- 7 簡而言之，使用 LISCOMP 推估「等級迷迷模型」有三個步驟 (Muthén 1984)，第一個階段先求取每個指標變項 (y) 對外衍變項 (x 's) 以波比迴歸作最大概似推估 (univariate maximum likelihood probit regression)。然後，在第二階段加入餘值結構矩陣的最大概似推估。經由這兩個步驟，可得到結構模型簡式中所包含的參數估計 (estimates of the reduced-form parameters of the structural model)。最後，以前面所求得之估計，應用「概化式最小平方適合函數」(GLS fitting function) 來推估模型的參數。
- 8 L^2 往往是以「飽和模型」(saturated model) 和「假設模型」(the hypothesized model) 比較所得之概似比 (likelihood ratio) 的對數值再乘以 -2，它在充量時呈 χ^2 分配 (asymptotically distributed as a χ^2)，其所對應的自由度等於「假設模型」中被釋放出之參數個數，也就是被用來假設檢定的那些參數。根據 Jöreskog and Sörbom (1988)，P 值可用來說明被驗證的模型能夠複製資料結構的機率；換句話說，P 值越大模型「適合」資料結構的情形越好。
- 9 通稱為 *bic* 的統計量是先由 Schwarz (1978) 所介紹，然後經 Raftery (1986) 修飾而成目前的形式， $bic_j = L_j^2 - df_j \times \log N$ ，N 乃是樣本數。根據 Raftery 的說法，在模型選擇時應用 *bic*，來檢測資料和模型的適合，而以值小為佳。Hauser and Wong (1989) 採用 *bic* 為模型選擇主要依據，理由是 L^2 會因樣本數的大小而改變其敏銳性，尤其在大樣本情形下，假設模型和資料間只需少許的不適合即可能在特定統計顯著下被放棄。因此使用 *bic* 可以用來調整 L^2 值，然而問題在於如何調整才算公允。

- 10 請參閱 Hauser and Wong 文表三 (1989: 160)。
- 11 如果手足配對中的長與幼的年齡差距過大，他們的教育程度的差異，不僅會受到個人與家庭環境等因素的影響，不同時代的結構因素 (如教育制度的改變) 也會有干擾，即所謂的「世俗性影響」(secular effects)。因此，在手足相似性的研究主題下，手足配對的年齡差距不宜過大。
- 12 對於表 4 各模型的介紹、假設檢定、與模型比較，其原理和表 3 相同，本段不再重複敘述。
- 13 Φ 代表外衍變項 (x_1, x_2, x_3) 的共變量矩陣。 ΦIN 的假設，說明四組手足配對樣本的社經指標分配沒有選樣偏誤。IN (invariant) 在 LISREL 程式的慣常用法，指參數值在群體間的一致性；而 EQ (equal) 指群體內參數值的相等。

參考資料

Behrman, J., P. Taubman and T. Wales

1977 "Controlling for and Measuring the Effects of Genetics and Family Environment in Equations for Schooling and Labor Market Success," in P. Taubman (ed.), *Kinometrics: The Determinants of Socioeconomic Success within and between Families*. Amsterdam: North-Holland.

Benin, M. H. and D. R. Johnson

1984 "Sibling Similarities in Educational Attainment: A Comparison of Like-sex and Cross-sex Sibling Pairs," *Sociology of Education* 57: 11-21.

Blake, Judith

1981 "Family Size and the Quality of Children," *Demography* 18: 421-442.
 1989 *Family Size and Achievement*. Berkeley: University of California Press.

Blau, P. M. and O. D. Duncan.

1967 *The American Occupational Structure*. New York: John Wiley.

Boudon, Raymond

1974 *Education, Opportunity and Social Inequality*. New York: John Wiley.

Bowles, Samuel

1972 "Schooling and Inequality from Generation to Generation," *Journal of Political Economy* 80: S219-251.

Chamberlain, Gary

1977 "An Instrumental Variable Interpretation of Identification in Variance -Components and MIMIC Models," in P. Taubman (ed.), *Kinometrics: The Determinants of Socioeconomic Success within and between Families*. Amsterdam :North-Holland.

Chamberlain, G. and Z. Griliches

1975 "Unobservables with a Variance Components Structures: Ability, Schooling and the Economic Success of Brothers," *International Economic Review* 16 (2): 422-449.

Duncan, Bevely

1968 "Trends in Output and Distribution of Schooling," In E.B. Sheldon and W. E. Moore (eds.), *Indicators of Social Change*. New York: Russell Sage Foundation.

Dunn, Judy and Robert Plomin

1991 "Why Are Siblings So Different? The Significance of Differences in Sibling Experiences Within the Family," *Family Process* 30: 271-283.

Griliches, Zvi

1977 "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica* 45 (1): 1-22.

1979 "Sibling Models and Data in Economics: Beginnings of a Survey,"
Journal of Political Economy 87: S37-S64.

Hauser, Robert M.

1988 "A Note on Two Models of Sibling Resemblance," *American Journal of Sociology* 6: 1401-1423.

Hauser, Robert M. and David L. Featherman

1976 "Equality of Schooling: Trends and Prospects," *Sociology of Education* 49: 99-120.

Hauser, R. M. and A. S. Goldberger

1971 "The Treatment of Unobservable Variables in Path Analysis," in H. L. Costner (ed.), *Sociological Methodology 1971*. Chapter 4. San Francisco: Jassay-Bass.

Hauser, R. M. and W. H. Sewell

1985 "Birth Order and Educational Attainment in Full Sibships," *American Educational Research Journal* 22 (1): 1-23.

1986 "Family Effects in Simple Methods of Education, Occupational Status and Earnings: Findings from the Wisconsin and Kalamazoo Studies," *Journal of Labor Economics* 4: S83-S115.

Hauser, R. M. and R. Wong

1989 "Sibling Resemblance and Inter-Sibling Effects in Educational Attainment," *Sociology of Education* 62: 149-71.

Hsueh, C. T. James

1992 "Sibling Resemblance in Educational Attainment: A Replication and Alternative," CDE Working Paper 92-11. University of Wisconsin—Madison.

Hurn, Christopher J.

1993 *The Limits and Possibilities of Schooling: An Introduction of Education*. Boston: Allyn and Bacon, Inc.

Jencks, C., M. Smith, H. Acland, M. J. Bane, D. Cohen, H. Ginits, B. Heyns, and S. Michelson

1972 *Inequality: A Reassessment of the Effects of Family and Schooling in America*. New York: Basic Books.

Jencks, C., S. Bartlett, M. Corcoran, J. Crouse, D. Eaglesfield, G. Jackson, K. McClelland, P. Mueser, M. Olneck, J. Schwartz, S. Ward and J. Williams

1979 *Who Gets Ahead?* New York: Basic Books.

Jensen, Arthur R.

1980 "Uses of Sibling Data in Educational and Psychological Research," *American Educational Research Journal* 17: 153-170.

Jöreskog, K. G. and A. S. Goldberger

1975 "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variables," *Journal of American Statistical Association* 70: 631-639.

Jöreskog, Karl G. and Dag Sörbom

1988 *LISREL 7: A Guide to the Program and Applications*. Chicago: SPSS.

Kuo, Hsiang-Hui D. and Robert M. Hauser

1995 "Trend in Family Effects on the Education of Black and White Brothers," *Sociology of Education* 68 (2): 136-160.

Lindert, P. H.

1977 "Sibling Position and Achievement," *Journal of Human Resources* 12: 198-219.

1978 *Fertility and Scarcity in America*. N.J.: Princeton University Press.

Mare, Robert D.

1993 "Educational Stratification on Observed and Unobserved Components of Family Background," in Yossi Shavit and Hans-Peter Blossfeld (eds.), *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, CO: Westview Press.

Mare, Robert D. and Meei-Shenn Tzeng

- 1989 "Fathers' Ages and the Social Stratification of Sons," *American Journal of Sociology* 95: 108-131.

Muthén, Bengt O.

- 1979 "A Structural Probit Model with Latent Variables," *Journal of American Statistical Association* 74 (368): 807-811.
- 1983 "Latent Variable Structural Equation Modeling with Categorical Data," *Journal of Econometrics* 22: 43-65.
- 1984 "A General Structural Equation Model with Dichotomous, Ordered Categorical, and Continuous Latent Variable Indicators," *Psychometrika* 49 (1): 115-132.
- 1987 *LISCOMP: Computer Program*. Scientific Software, Inc.

Olneck, Michael R.

- 1977 "On the Use of Sibling Data to Estimate the Effects of Family Background, Cognitive Skills, and Schooling: Results from the Kalamazoo Brothers Study," in Paul Taubman (ed.), *Kinometrics: Determinants Socioeconomic Success within and between Families*. Amsterdam: North-Holland.

Olneck, Michael R. and David B. Bills

- 1979 "Family Configuration and Achievement: Effects of Birth Order and Family Size in a Sample of Brothers," *Social Psychology Quarterly* 42: 135-148.

Raftery, Adrian E.

- 1986 "Choosing Models for Cross-Classifications (Comment on Grusky and Hauser, ASR, Feb., 1984)," *American Sociological Review* 51: 145-146.

Schwarz, Gideon

- 1978 "Estimating the Dimension of A Model," *The Annual of Statistics* 6 (2): 461-464.

Sewell, William H.

1971 "Inequality of Opportunity for Higher Education," *American Sociological Review* 36: 793-809.

Sewell, W. H. and R. M. Hauser

1977 "On the Effects of Families and Family Structure on Achievements,"
In P. Taubman (ed.) *Kinometrics: The Determinants of Socioeconomic Success within and between Families*. Amsterdam: North-Holland.

Taubman, Paul

1977 *Kinometrics: The Determinants of Socioeconomic Success within and between Families*. Amsterdam: North-Holland.

Winship, C. and R. D. Mare

1984 "Regression Models with Ordinal Variables," *American Sociological Review* 49: 512-525.

Zajonc, Robert B.

1976 "Family Configuration and Intelligence," *Science* 192: 227-236.

Zajonc, Robert B. and Gregory B. Markus

1975 "Birth Order and Intellectual Development," *Psychological Review* 82: 74-88.

Zajonc, R. B., H. Markus and G. B. Markus

1979 "The Birth Order Puzzle," *Journal of Personality and Social Psychological Review* 37: 1325-1341.

「對數線性模型」與「對數相乘模型」 的發展與運用：以社會流動的分析為例*

蔡瑞明**

壹、前言

本文主要是討論對數線性分析 (log-linear analysis) 的最新趨勢與運用。近年來，「對數線性模型」(log-linear models) 與「對數相乘模型」(log-multiplicative models) 有許多突破性的發展。這一類的研究方法利用預設的模型設計來分析行列表 (或稱列聯表 contingency tables)，檢驗理論假設，比傳統的對數線性分析的層級模型 (hierarchical models) 更為實用而有效率。對數線性分析不但可以用於社會流動，還可用於其他的類別資料 (categorical data) 的分析。最新

* 本文初稿曾發表於中央研究院中山人文社會科學研究所「社會科學計量方法之發展與應用」研討會，以及東海大學應用統計學研究所。作者感謝王德睦教授對本文初稿的評論，並且謝謝蔡淑鈴教授、瞿海源教授、研討會與會的學者和統計所的教授們、以及兩位匿名評審提供的許多寶貴意見。Yu Xie 與 Raymond Wong 兩位教授在 GLIM 的程式設計上不吝賜教，在此一併致謝。本文使用的分析資料是根據「台灣地區社會變遷基本調查」的問卷資料。作者感謝該調查研究的主持人瞿海源教授在資料取得與分析整理的協助與建議。本文寫作期間，國科會專題研究計劃 (NSC84-2412-H-029-004) 提供資助。當然，一切文責，仍由筆者自負。

** 東海大學社會學系副教授

的分析模型結合了對數線性與對數相乘的參數，不僅擁有 Goodman 的關連模型中不受類別項目排列影響的優越性，而且又可加入其他非垂直 (nonvertical) 的結構性參數。這一類的對數線性與對數相乘的模型具有清晰而內容豐富的優點。

社會科學的資料有許多是以名目尺度 (nominal scale) 測量的。對數線性模型常用來分析這些資料。由於許多類別資料都含有次序 (ordinal) 排列的特性，例如階級、政治取向，而此特性又有其理論及研究的意義；因此，使用可以分析此特質的方法是很重要的。有些對數線性及對數相乘模型，例如本文所要討論的關連模型，特別適用於內涵次序的名目尺度資料。其中，對數相乘模型的最主要優點是它提供次序尺度變項的類別間距離的資訊 (Clogg, 1982a: 114)。對數線性分析在社會科學的應用，愈來愈廣泛，不僅於社會流動的分析，¹ 研究者更將之運用於其他研究領域，諸如網絡分析 (例如 Marsden, 1981)、婚姻配對 (例如 Hout, 1982; Kalmijn, 1991)、政黨政治 (例如 Weakliem, 1991)、勞力市場 (例如 Clogg, Eliason and Wahl, 1990)、及組織關係 (例如 Liedka, 1991) 等方面。

以往，一般的分析方法是使用層級模型 (hierarchical models) 來分析類別資料，但是層級模型缺乏靈活性；特別是在行列的互動效應的分析上，配適整組分類項目的互動關連，常使用太多的自由度，分析過於粗略。飽和模型 (saturated model) 又因使用所有的自由度，不是一個精簡的 (parsimonious) 模型，驗證假設的能力不強。一般說來，在行列表內的互動效應有許多不同的組合，而各個不同的組合蘊含著不同的理論意義；對數線性模型的分析，可以根據理論假設對行列之間 (即表內不同格子) 的關連加上限制，以符合模型必須是精簡的，而且又具有統計檢定力的原則；換言之，在同一組合內，我們可以用同一參數來配適，節省自由度的數目，使之較具解釋力。就分析而言，這些是屬於「非層級模型」(non-hierarchical models)。

爲了有效地分析行列表，Goodman (1979) 提出一系列的關連模型 (association models) 用於分析內涵次序的名目尺度資料 (ordered categories)，爲行列表的分析開創出一片新的天地，也爲社會流動的研究在地位取得模型 (status attainment model) 典範之外，另闢新的研究路徑。Goodman (1979) 的關連模型分析方法 (analysis of association, 簡稱 ANOAS) 是以 odds-ratio 的設定來建構各種「簡單」的模型 (simple models)，主要可分爲虛無關連 (null association)、齊一關連 (uniform association)、行效應關連 (row-effect association)、列效應關連 (column-effect association)、行列效應第一模型 (row and column effects I) 及行列效應第二模型 (row and column effects II) 等模型。Goodman 以行列表中相鄰的 2×2 次表 (subtables) 的 odds-ratio 來估計行列的關連。² 假設 F_{ij} 爲行列表 ij 格子 (cell) 的期望值，則 odds-ratios $\theta_{ij} = (F_{ij}F_{i+1,j+1}) / (F_{i,j+1}F_{i+1,j})$ 。例如， $\theta_{ij} = 1$ ，則是虛無關連模型； $\theta_{ij} = \theta$ ，則爲齊一關連模型。上述的 ANOAS 模型皆可以 odds-ratios，清楚地設定表示。其中，尤其以 Goodman 的行列效應第二模型 (或簡稱 RC 模型) 在社會流動的研究中最具特色，是上述模型中唯一的對數相乘模型。透過 odds-ratio 的設定與估計，Goodman 所提出的模型不僅有效率，且具有統計上的簡潔特性。特別是因爲其行列效應的估計值不受行列次序的影響，近年來被廣泛地應用於分析社會流動的研究。³

Goodman 提出關連模型時正值地位取得模型研究如日中天，由於迴歸分析與地位取得在出身 (ascribed) 及成就 (achieved) 的職位與聲望的測量與模型建構上，相當容易配合，因而蔚爲潮流。此外，迴歸分析發展已久，應用上相當簡單，可根據理論上的需要，將各種相關的解釋變項加入分析模型。相較之下，Goodman 的方法，難以應用於地位取得的分析模型，因而未受到應有的重視。這種情況，一直到階級分析又受到重視之後，才逐漸改變。

就研究分析的角度看來，迴歸分析與對數線性分析兩者各有其優點。但是，地位取得模型是以職業聲望等連續性的變項為主要的指標，使用迴歸分析的技術來研究社會流動。基本上，迴歸分析將豐富而複雜的社會階級關係縮減成單一的迴歸係數，難以有效地分析社會流動範疇中，資源分配的相對關係。例如，就地位取得分析而言，一個人無法進入上層階級，只是因為她或他在某些出身或成就的變項上表現不佳。如此的說法，忽略了階級間的動態關係及階級藩籬的影響。⁴ 換言之，地位取得的分析在某種程度上忽略社會位置與社會流動的關係。而 Goodman 的流動表分析方法可說是將分析的面向擴大到階級(職業)的相對位置及結構關係。Clogg (1982a; 1982b) 及下面將提及的研究者將這些模型擴展成比較行列表的模型及其他行列關係的研究。迴歸分析在社會流動研究上的侷限，體現在地位取得模型 (Blau and Duncan, 1967) 的創始人 Otis D. Duncan 在1979的 *AJS* 發表「職業流動表內終點如何依賴起點」一文，提出並討論「齊一關連模型」在社會流動研究上的應用。Duncan 後來又與 Hout 及 Sobel 等人 (Sobel, Hout and Duncan, 1985; Hout, Duncan and Sobel, 1987) 在合力開發分析結構性流動的對數模型上，成就豐碩。⁵ Duncan 的研究轉向對數線性分析象徵著以往慣用的迴歸分析在研究社會流動上有其不足之處。

貳、關連模型的建構

在分析社會流動時，可以用流動表的邊際總數 (margins) 來直接測量其絕對流動 (absolute mobility) 的比率，計算外流 (outflow)、內流 (inflow) 及總流動 (total mobility) 比率；另一方面，也可以控制邊際總數 (margins) 來測量其相對流動 (relative mobility) 的程度。許多研究使用對數線性與對數相乘模型來分析相對流動，兩者皆可控制邊際分配的影響，並分析流動表內類別間的關係 (參見 Goodman, 1979; 1984; Clogg 1982a; 1982b; Agresti, 1990; Hout 1983; Hagenaaers, 1990;

McCullagh and Nelder, 1989: Chapter 6)。⁶有些職業流動與地位取得的社會學研究也使用類似的模型 (Hauser, 1984; Smith and Garnier, 1987; Wong, 1990; 1992; Xie, 1992)。這些模型結合對數線性與對數相乘的參數，它們不僅擁有一般常用的 Goodman 的關連模型 (Goodman, 1984) 中項目排列的優越性，測量垂直接流動 (vertical mobility)；而且又可加入其他非垂直接流動 (nonvertical mobility) 的結構性參數。與 Hauser (1978; 1979) 的層次模型 (levels model) 及 Hope (1982) 的鑽石模型 (diamond model) 相比較，這一類的對數線性與對數相乘的模型具有清晰與內容豐富的優點。Hauser 的模型被批評認為太具實驗性格；而 Hope 的模型則是太過狹隘 (Yamaguchi, 1987; Wong, 1992)。⁷

層級參數與交錯效應參數這兩組參數測量社會流動的起點 (origins) 與終點 (destinations) 的關連性 (association)，也就是 i 行與 j 列的期望次數 (F_{ij})。基本上，假定在流動表上，不同的區域有不同的流動密度，而有不同的關連性 (association)。此外，交錯效應參數 (crossing effects) 假設流動表的項目的排列次序是正確的。效應參數常用於分析對稱關係的流動表，如婚姻配對 (參見 Johnson, 1980; Pöntinen, 1982)。由於對稱的關係，模型參數通常比一般的社會流動表的分析來得精簡。

在測量非垂直接流動方面，Hauser (1978; 1979) 所提出的 topological model 或稱「層次模型」(levels model) 基本上是 Goodman 的準獨立模型 (quasi-independent model) 的擴大模型 (參見 Goodman, 1984)。⁸ Hauser 的模型雖被批評太具實驗性格，但是，後來有許多研究將之發展成較具理論驗證的模型。Goldthorpe 等人 (Goldthorpe, 1987; Erikson and Goldthorpe, 1992) 將流動表劃分成多個區隔，使用如繼承性 (inheritance)、親近性 (affinity) 等概念來檢驗不同的代間流動的形態。立基於層次模型，Stier and Grusky (1990) 提出「重疊持續模型」(overlapping persistence model, 簡稱 OP 模型)，利用較大的流動表，分析勞力市場的生涯流動 (career mobility) 與二元市場分隔的關係。他

們使用五個持續流動參數：次職業持續 (suboccupational)、職業持續 (occupational)、次層面持續 (substratum)、層面持續 (stratum) 及部門持續 (sector)，來分析生涯流動的結構限制。Wong (1990; 1992) 則比較國際間的變異，將非垂直的流動分成階級對角線繼承 (class-specific diagonal inheritance)、部門繼承 (sector inheritance) 及一般階級繼承 (general class inheritance)。此外，Breiger (1981; 1982; 亦參見 Snipp, 1985; Jacobs and Breiger, 1988) 依據韋伯的階級理論，應用 Hauser 及 Goodman 的模型，提出一種對數線性模型來檢定同質性結構 (homogeneous structure)，用以合併 (減少) 流動表中不具顯著性的階級或職業類別的數目，以提高分析的效率，並可用來建構階級分類架構 (亦參見 Goodman, 1981)。Breiger (1990) 並試圖利用相同的模型處理解決 Wright and Perrone (1977) (參見 Wright, 1985) 所謂的「矛盾位置」(contradictory class positions) 的分類問題。最近，馬克思主義的階級研究也開始運用各式的對數線性模型來分析階級互動關係及階級藩籬 (Wright and Cho, 1992; Western and Wright, 1994)。總之，這些研究靈活地運用層級模型發展出各類的對數模型來概念化及檢驗社會流動的理論假設。

一般的研究常用「齊一關連模型」(uniform association model) 來測量關連性 (Goodman, 1979; Duncan, 1979)。此模型使用一個參數來估計行列的關連，其優點在於它的參數與一般的迴歸係數很相似。但是，這個模型假設相鄰的項目的距離都相等，而且項目之排列次序是正確的。在許多情況下，它的假設是過度牽強了。估計齊一關連模型，一般都使用行列的次序 (如 1, 2, 3, ...) 來建構設計矩陣 (design matrix)。Hout (1984) 認為類別的垂直關係可以使用其他外在的指標來設定。Hout 採用 Haberman (1974) 的 linear-by-linear 互動模型，使用各類別內職業的平均 SEI 分數、職業的自主性 (Autonomy) 及職業的訓練 (Training in occupation) 三個指標來取代單一的社會經濟地位指標或沒有特殊意義的行列次序，所以名之為 SAT 模型。基本上，SAT

模型是 Duncan (1979) 的齊一關連模型的擴大模型。Hout 的 SAT 模型與 Goodman 的行列效應模型 (RC) 兩者都可用來分析垂直接流動，SAT 模型採用一種事前設定 (*a priori*) 的方式，利用一些社會經濟指標預設類別的關係；而 RC 模型則是使用事後設定 (*a posteriori*)，利用社會流動資料所呈現的結果，來估計階級或職業類別的相對關係 (Yamaguchi, 1983; 1987; Wong, 1990)。

社會流動的研究常用的許多模型是以 Goodman (1981;1984) 所發展出來的「行列效應第二模型」(row and column effects model II) 為基礎。⁹ 該模型不受分類項目次序的影響，而且是根據社會流動所形成的內在關係模式來決定項目之間的距離。在社會流動中，測量垂直接流動的形態時，即使行列的類別的排列有錯，其估計參數也不受影響。其他的對數線性模型的估計則會因類別的次序排列而變動，因此，此一模型對於無法確定次序排列的分類資源，特別適用。

本文討論的對數線性及對數相乘模型主要是用來分析兩種流動表：二維 (two-way) 與三維 (three-way) 流動表。雖然社會流動表皆為行列數目一樣的正方形行列表 (square tables)，本文討論的模型大多 (除了對稱模型) 可應用於行列的變項特質不一樣的行列表。分析二維流動表的關連模型可以設定如下：¹⁰

$$\log F_{ij} = \lambda + \lambda_i + \lambda_j + \Delta + \Psi \quad (1)$$

其中， λ 代表模型的總平均數； λ_i 與 λ_j 分別配適 (fit) 流動表的行邊際值與列邊際值；而 $\Psi = \phi\mu_i\nu_j$ ，配適行列效應， Δ 則通常是配適不同的層級參數 (levels parameters) 與交錯效應參數 (crossing effects) 所組成的參數群。¹¹ $\Psi = \phi\mu_i\nu_j$ 不是一個對數線性關係，而是對數相乘的關係。 μ_i 與 ν_j 分別是行與列類別的標準分數； ϕ 則測量行列間內在關連的程度。因此，加上類似此乘數的對數模型稱為「對數相乘模

型」。 $\Psi = \phi\mu_i\nu_j$ 並可由單一層面，擴大到多個層面，用以測量多層面的互動關連 (multidimensional RC)，則等式 1 可改寫成：

$$\log F_{ij} = \lambda + \lambda_i + \lambda_j + \sum_{m=1}^M \phi_m \mu_{im} \nu_{jm} \quad (2)$$

這個模型稱為 RC (M) 關連模型 (Goodman 1986; Becker and Clogg 1989)。Goodman 的 RC 模型是 RC (M) 的特殊模型 (即一個層面；M=1)。RC (0) 表示行列並無關連，即為獨立模型。有些研究則使用一個層面以上的模型 (例如 Gilula and Haberman, 1986; Smith and Garnier, 1987; Clogg, Eliason and Wahl, 1990; Weakliem, 1991)。此種模型的分析結果與解釋，與因素分析 (factor analysis) 類似。換言之，此模型將行列的關連分割成多個層面 (dimensions)，根據理論假設，可以得出各層面所代表的意義。但是，由於此模型仍未有如因素分析的轉軸 (rotations) 功能，層面之間僅能成垂直關係 (orthogonal)。RC (M) 關連模型除了可檢驗參數的顯著程度之外，又可分析關連本身含有多少層面及不同層面的關連程度有無差異。

除了分析二維的流動表之外，若將二維模型加上多層級的設定，即成爲一種限制性的三維對數線性模型。三維的流動表分析適用於跨國 (或社會、群體) 的比較研究。一般的分析模型仍然是以 Goodman 的關連模型爲主。根據 Goodman 的模型，Clogg (1982b) 提出一系列適用於比較分析流動表的模型。在社會流動的比較研究中 (例如 Hauser, 1984; Yamaguchi, 1987; Wong, 1990; 1992)，最有名的分析模型大概是由 Erikson 與 Goldthorpe (1987a; 1987b; 1992) 所提出的「共同流通模型」 (common fluidity model)。這個模型假設每個流動表 (代表一個國家或社會) 的起點 (i) 與終點 (j) 的每一個格子裡的互動程度是相同的。此模型可以用下面的方程式來表示：

$$\log F_{ijk} = \lambda + \lambda_i + \lambda_j + \lambda_k + \lambda_{ik} + \lambda_{jk} + \lambda_{ij} \quad (3)$$

其中， k 代表一個層級（通常指社會或國家）。簡單說來，這個模型將飽和模型的參數（即等式 3 加上一個參數群 λ_{ijk} ）設限成每個流動表的行列的互動形態設定為一致（ $\lambda_{ijk} = 0$ ）。Goldthorpe 與 Erikson 使用此模型的目的是為了驗證社會流動研究的 *FJH* 假設（Featherman, Jones, and Hauser, 1975）：各國的相對代間流動是相似的。這是一個相當簡單的模型。但是，它使用過多的自由度，不是一個精簡的（parsimonious）模型。在比較研究上，跨組（社會或國家）的差異其實是可以使用較少的參數來配適。就資料分析而言，共同流通模型很少能夠配適流動表；而且模型的設定忽略了職業或階級的層級性。在驗證 *FJH* 假設上，Goldthorpe and Erikson 模型的配適結果並不能有力地支持 *FJH* 的假設。總而言之，Goldthorpe 等人的研究基本上是以對數線性模型為主，很少使用 Goodman 提出的對數相乘模型來分析社會流動。若將上面所提的二維分析模型（等式 1）擴展，加上 K 層級的設定，則可形成下面的模型，用來分析跨組的比較：

$$\log F_{ijk} = \lambda + \lambda_i + \lambda_j + \lambda_k + \lambda_{ik} + \lambda_{jk} + \Delta + \phi_k \mu_{ik} \nu_{jk} \quad (4)$$

由於這個模型可以限制成每一層級內 μ 與 ν 是相等的，若與上述 Goldthorpe 和 Erikson 的模型相比，因為 μ 與 ν 測量階級關係的層級性，因而這個模型是相當精簡的。不僅節省多個自由度，又深具解釋力。因為對數相乘的特色，這個模型預設流動表的項目的次序在各國是相同的。總之，本模型不但可以限制行列的效應（ $\mu = \nu$ ）是一樣的；並可以限制一般關連效應（ ϕ ）為各行列表（國家或群體）相同或各不相同，並檢定比較各國的關連程度。一般關連效應（ ϕ ）參數測量行列之間總體的關連，也就是流動終點對起點的依賴程度（Duncan, 1979）。由此可見，這是一個強而有力又靈活的比較模型。

參、準對稱模型

每個社會在不同的時間點上，特別是在職業結構快速變動時，職業分配會隨之變動。職業結構的變動，提高兩代間職業的異質性，必然會降低代間傳承的比率。在比較研究方面，有些研究著重於分析結構性流動 (structural mobility) 與交換性流動 (exchange mobility) 分別的影響。許多研究者一直想分析一個社會的結構性流動狀況 (Rogoff, 1953; Yasuda, 1964; Featherman and Hauser, 1978; Hope 1982)。但是，如何估計結構性流動，則眾說紛紜。Hope (1982) 認為獨立模型與他所提的「半面模型」(halfway model) 的差異就是結構性流動。Featherman, Jones 與 Hauser (1975) 認為流動表起點與終點的邊際總和的參數即結構性流動；交換性流動 (循環性流動) 則是兩個參數的互動效應。這類的論點傾向於直接計算流動表的邊際差異可得到結構性流動；而結構性流動的殘差 (residual) 即為交換性流動。Sobel, Hout 與 Duncan (1985) 認為這類的說法缺乏一致性，因為結構性流動與交換性流動是並存而不可分開處理的。他們支持結構性流動可以由流動表的「起點」與「終點」的邊際異質性 (marginal heterogeneity) 來計算的論點。但是，他們證明流動表至少要符合「準對稱模型」(quasi-symmetry model) 的最低條件，才可以分離出結構性流動與交換性流動的各別程度。換言之，Sobel 等人將結構性流動定義為「流動表關連是對稱時的邊際差異值」(Sobel et al. 1985: 360-2)，這也就是「準對稱模型」成立的條件。準對稱模型表示一個行列表中沿著主要對角線 (main diagonal) 的互動效應是對稱的 (Bishop, Fienberg and Holland, 1975; Hout, 1983)。若將準對稱模型加以擴展，可用來分析比較多層的流動表；在比較研究方面，最通用的模型是「條件準對稱模型」(conditional quasi-symmetry) (Sobel, 1988)。這個模型設定每一比較組 (國家或社會) 內，符合準對稱模型。這個模型預測各組流動表內

起點 (例如父代的階級) 與終點 (例如子代的階級) 互動格子內的觀察值 (F_{ijk}) 為：

$$F_{ijk} = \mu_k \beta_{ik} \beta_{jk} \alpha_{jk} \delta_{ijk} \quad (5)$$

在每一比較組內， $\delta_{ij} = \delta_{ji}$ ； $\prod_j \alpha_j = 1$ ；而且 $\beta_i = \beta_j$ ；假若 $i = j$ 則 $\delta_{ij} = 1$ 。根據 Sobel et al. (1985; Hout, Duncan and Sobel, 1987)，在二維流動表內， α_j 測量結構性流動。 β_i 、 β_j 與 δ_{ij} 的相乘，代表交換性流動。在三維的模型裡，可以透過對特定參數的限制 (restrictions)，來檢驗比較組之間流動模式的差別。所以，此一模型不僅可以分析比較跨國或跨社會的結構性流動狀況，而且分析出那些參數 (流動的模式) 促成比較組結構性流動以及交換性流動的不同程度 (參見 Landale and Guest 1990)。有關臺灣社會結構的分析，Tsay (1996b) 應用此模型比較四個年齡層結構性流動對社會流動的影響，發現產業的轉型所帶動的階級流動是社會流動的主體。

肆、對數相乘層級效應模型

在分析和比較流動表方面，Yamaguchi (1987) 提出一個以單一參數來比較流動表的差異的模型，Xie (1992) 稱之為「齊一層級效應模型」(uniform layer effect model)。Yamaguchi 的模型比較兩個行列表的齊一差距，是一個精簡而又有實質意義的模型。但此模型亦有其弱點。首先，它如同「齊一關連模型」一樣，預設著行列的類別順序是正確的，而且行列表間的距離是一致的。但是，有許多的流動表的類別順序，與社會現實不盡符合，可能有錯，因而 Yamaguchi 的模型假設難免過於牽強。其次，此模型在分析的技術上並無法加入 Hauser 的等級模型 (levels model)；然而等級模型在分析流動表內的流動密度，非常有效。此外，Yamaguchi 的模型也不能用來與 Goodman 的第二行列效應關連模型 (RC model) (Xie, 1992: 380-1) 比較。為了解決上述的

問題及有效地比較行列表，Xie (1992: 380-4) 提出「對數相乘層級效應模型」(log-multiplicative layer effect model)，用於分析比較流動表。此模型的特色是將跨行列表的變異性限制成由一個「共同的關連」(a common association pattern) 及一個「行列表特有的參數」(a table-specific parameter) 所結合的對數相乘的乘積數。所以，此一模型假定各流動表間行列的關連有一個共同的模式。

$$\log F_{ijk} = \lambda + \lambda_i + \lambda_j + \lambda_k + \lambda_{ik} + \lambda_{jk} + \psi_{ij}\phi_k \quad (5)$$

其中， ψ_{ij} 測量行列表的起點與終點的關連性 (association)； ψ_{ij} 可設定成各種模型參數，如「齊一關連模型」、「行效應模型」(row effects model)、「列效應模型」(column effects model) 及 Goodman 「行列效應第一與第二模型」等。而 ϕ_k 代表各個流動表關連的差異特殊性 (deviation) (Xie, 1992)。此模型的特性之一是可將相同的限制 (restrictions) 加諸於各個行列表上，並同時測量比較各個行列表 (層級) 個別的關連性，從事一個類似「直接的」比較。此模型不僅可以分析各行列表的特性，也可比較行列表與行列表之間的差異，而且在模型設定上很靈活 (Xie, 1992)。

伍、模型估計與檢驗的方法

一個清楚明晰而解釋能力強的模型必須配適資料，並且不應有過度的探索性。研究者必須從多個競爭的模型 (competing models) 中，選擇出最佳的模型。模型的探尋可分成兩種，一是最簡單的模型為起始，當作一個基準模型 (baseline model)；然後，逐一加入相關的參數，一直到獲得最配適的模型為止。另一種方式是由最複雜的模型 (例如，飽和模型) 為起點，逐漸放鬆參數的限制，找尋最佳模型。

模型的選擇是利用配適的優劣 (goodness of fit) 來決定那一個模型是最適的。比較的準則，主要是使用 log-likelihood ratio chi-square

(L^2)，愈小的 L^2 ，表示模型愈配適。由於絕大部分的模型都是飽和模型 (saturated model) 內加上參數限制，所以可以使用 nested models 的方式來比較模型間的優劣。對於不是 nested 的模型，就無法直接比較 log-likelihood ratio chi-square 來選擇模型。傳統的檢定方式，容易受到樣本數目的影響。針對此問題，Raftery (1986) 提出的貝氏模型選擇條件 (Bayesian Information Criterion，簡稱 *BIC*)。晚近，*BIC* 已愈來愈被研究者所採用來衡量模型配適的優劣。Raftery 的方法是以 L^2 、自由度 (*df*) 及樣本數 (*N*) 三者來決定模型的配適程度。*BIC* 的公式如下：

$$BIC = L^2 - df (\log N) \quad (7)$$

BIC 不僅可以處理行列表樣本數目過大所造成的影響，而且更重要的是可以做非 nested models 間的比較。但有些研究 (例如 Weakliem, 1992) 已開始提出一些比 *BIC* 更有效的方法，不過計算上較困難，且仍有待標準化。在社會流動的研究方面，由於通常使用較小的流動表，自由度數目不大，因此使用並比較非 nested 模型，無論在理論的驗證上或經驗分析的需求上，都有其必要性。這也是近年來社會流動研究愈來愈倚賴 *BIC* 的原因之一。

對數線性模型的另一特色是可以利用兩個模型的比較，檢驗特定的參數是否顯著。利用 chi-square partition 的方法，比較兩個模型的 chi-square 差值。此一差值也會形成 chi-square 分配，因此可以用來檢驗參數是否顯著 (Bishop et al., 1975; Hout, 1983)。比較多維 (如跨國) 的流動表時，Yamaguchi (1987) 使用一種簡潔的方法來比較參數的差異，在數個二維流動表相同設定的條件下，選取並比較這些流動表一個參數的差異，稱之為「一個自由度檢驗」(one-degree-of-freedom test)。由於只有使用一個自由度，這種檢驗方法可以有效地比較多個二維流動表間的差異來源。不過，這種檢驗必須將其他的參數設限，模型的假定 (assumptions) 程度較強。

本文使用 GLIM (Numerical Algorithms Group 1993) 的程式 (macros) 來估計模型參數。GLIM 的程式不大，但在模型的估計上，能相當靈活地使用上一個計算過程的估計值再予以運算。有關使用 GLIM 配適對數線性模型的方法，請參見 Aitkin (1989)；在社會學的運用，可參考 Breen (1984)。¹² 一般的統計套裝軟體程式如 SAS 的 CATMOD 及 SPSS 的 LOGLINEAR 皆可估計一般的對數線性模型，但是要估計複雜的對數相乘模型 (如 Goodman 的 RC 模型) 則須利用上述套裝軟體的 Matrix 程式。¹³ 除了 GLIM 之外，APL (STSC, 1989)、GAUSS (Aptech Systems, 1991)、CDAS (Eliason, 1990)、LIMDEP (Greene, 1991) 等軟體程式也可用來估計對數相乘模型。

關於模型與參數的估計方法，一般的統計軟體大致上使用下列兩種方法之一：一是 Iterative Proportional Fitting (IPF) 程序，此方法運算程序簡單，容易估計模型，但只適用於估計層級模型 (hierarchical models)，且不能準確地計算參數估計值的標準誤；另一種方法是 Newton-Raphson 程序，此方法的估計步驟較複雜，須使用各類設計矩陣來估計參數，但可用於估計各類的關連模型，且能運算出參數估計值的標準誤 (Bishop et al., 1975; Haberman, 1974; Goodman, 1984)。由於一般的套裝軟體在分析對數線性模型，尤其是對數相乘模型，並無標準化的程式，因而分析上的困難度較高。

在分析配適上，對數線性及對數相乘模型最大的特色是分析過程中，設計矩陣 (design matrix) 的活用性。由於兩種對數模型的假設經常是複雜的，需要透過設計矩陣的規劃才能有效地配適。設計矩陣的主要功能是将流動表內依互動的同質性的程度，歸類成不同的幾組流動區，並假設在同一組的區內，流動的密度是一樣的。換言之，研究者可以根據理論假設，將之轉換成各類相異的流動形態，透過設計矩陣，納入對數線性模型的估計程式。模型配適的結果可以測量並比較代表不同區域參數的大小。表 1 是一個典型的設計矩陣，其中的數字只是代表不同的區域，並沒有等級之分。例如，其中的數字「2」假設

「高級非手工」與「一般非手工」兩個階級之間的關係在流動密度上具有同質性。

表 1：代間社會流動的設計矩陣

		子代階級					
		I	II	III	IV	V	VI
父親階級							
I	高級非手工 (High Non-Manual)	2	2	1	1	1	1
II	一般非手工 (Routine Non-Manual)	2	2	1	1	1	1
III	自營 (Self-Employed)	1	1	3	1	1	1
IV	技術性手工 (Skilled Manual)	1	1	1	4	4	1
V	非技術性手工 (Unskilled Manual)	1	1	1	4	4	1
VI	農人 (Farm)	1	1	1	1	1	5

		子代階級					
		I	II	III	IV	V	VI
父親階級							
I	高級非手工 (High Non-Manual)	7	7	6	6	6	6
II	一般非手工 (Routine Non-Manual)	7	7	6	6	6	6
III	自營 (Self-Employed)	6	6	8	6	6	6
IV	技術性手工 (Skilled Manual)	6	6	6	9	9	6
V	非技術性手工 (Unskilled Manual)	6	6	6	9	9	6
VI	農人 (Farm)	6	6	6	6	6	10

以上大致討論了對數線性模型與對數相乘模型的發展及其在社會流動上的運用，以下將使用上述的一些模型來分析臺灣的社會流動現象。

陸、臺灣兩性社會流動的比較

表 2 是由「臺灣地區社會變遷基本調查」(瞿海源 1991) 中，父子及父女由受訪者的第一個職業到父親的職業代間社會流動的資料。¹⁴ 本文的階級分類是以 Goldthorpe 等人的 EGP 分類架構為基礎 (Erikson-Goldthorpe-Portocarero Scheme, 簡稱 EGP) (Erikson, Goldthorpe and Portocarero, 1979; Goldthorpe, Llewellyn and Payne 1987; Erikson and Goldthorpe, 1992)。¹⁵ 由於社會變遷的調查資料是以「國際標準職業分類表」(International Standard Classification of Occupations, 簡稱 ISCO) (Treiman, 1977) 來登錄。為了將此資料轉成 EGP 分類架構，本研究利用由 Ganzeboom 等人所發展的轉換程序將臺灣的 ISCO 資料轉成 EGP 的分類 (參見 Ganzeboom, Luijkx and Treiman, 1989)。¹⁶ 由於資料筆數不大，無法使用 EGP 的整個分類架構，故而將流動表的階級分類縮減成六個階級：高級非手工階級 (high non-manual)，一般非手工階級 (routine non-manual)，自營階級 (self-employed)，技術性手工階級 (skilled manual)，非技術性手工階級 (unskilled manual) 及農人階級 (farm)。¹⁷ 臺灣的社會流動研究所使用的分類架構，大致可分成三種。早期的研究通常沿用政府部門 (如主計處) 所慣用的職業分類，將其集成成幾個大類。這種分類一方面相單簡單且容易取得，其次又可以直接與其他學門的研究及政府的報告資料比較。但是，將此分類架構所得的職業結構逕視為階級結構，則缺乏理論依據。以社會理論為基礎，建構的分類方式，大致有兩種：一種是應用 Wright 所提出的 Marxist 的階級分類 (Wright, 1985)；¹⁸ 另一種則是使用 Goldthorpe 及其同事以韋伯 (Weber, 1978) 的階級理

表 2：臺灣男性與女性代間流動表：1990資料

		男性						
		子代階級						
父親階級		I	II	III	IV	V	VI	總數
I	高級非手工	101	43	46	32	29	4	255
II	一般非手工	12	11	8	8	12	0	51
III	自營	51	45	76	35	49	6	262
IV	技術性手工	26	24	29	69	29	9	186
V	非技術性手工	28	18	21	30	47	3	147
VI	農人	93	47	112	159	150	325	886
總數		311	188	292	333	316	347	1787

		女性						
		子代階級						
父親階級		I	II	III	IV	V	VI	總數
I	高級非手工	82	95	47	18	21	5	268
II	一般非手工	11	29	19	10	10	0	79
III	自營	47	86	56	34	39	9	271
IV	技術性手工	9	31	25	55	53	4	177
V	非技術性手工	23	34	27	25	43	2	154
VI	農人	42	104	87	161	230	195	819
總數		214	379	261	303	396	215	1768

附註：總樣本數為 3555。

論所發展的 EGP 分類架構 (如 Wu, 1994; So and Hsiao, 1994)。¹⁹ 有些臺灣的研究指出 EGP 分類架構可能優於 Wright 的分類架構 (Wu, 1994)。但是，分類架構的優劣，涉及研究的理論觀點、社會的結構及發展程度、與分析的效度及信度，甚至是研究者的個人見解，因而難以遽下判斷，實有待進一步的探究。考量臺灣戰後的階級的形成與階級關係 (參見 Tsay, 1995)，及高度流動的社會階層結構 (Sheu, 1989; Tsay, 1996b)，本文決定使用 Goldthorpe 的 EGP 分類架構。

近年來，臺灣的社會學界在社會流動研究方面，成果豐碩，對臺灣社會的族群、性別與年齡層的比較研究，已有更進一步的瞭解 (許嘉猷, 1982; 1987; 1990; 1994; 瞿海源, 1983; 蔡淑鈴, 1986; 1987; 1994; 蔡淑鈴、文崇一, 1985; 張維安、王德睦, 1983; 謝雨生、余淑娟, 1990; 孫清山、黃毅志, 1994; 蔡淑鈴、瞿海源, 1992; 蔡瑞明, 1995; Sheu, 1989; Tsai and Chiu, 1993; Tsay, 1996a; 1996b)。跨國的社會流動比較研究，也日漸增加 (許嘉猷, 1990; 1994; Sheu, 1989; So and Hsiao, 1994)。除了少數如蔡淑鈴、文崇一 (1985) 及蔡淑鈴 (1986; 1994) 使用對數線性及對數相乘模型來分析社會流動表之外，以往的職業或階級流動研究偏重於以迴歸分析來分析資料，因而對流動的模式及流動的管道 (channels of mobility) 比較難以掌握。使用對數線性及對數相乘模型可分析階級 (類別) 間的相對關係，並檢驗階級間的流動性，有助於對階級關係的瞭解。

柒、對數線性模型分析

以下本文使用一系列的對數線性與對數相乘模型配適臺灣社會流動的資料。這是由兩個 6×6 的行列列表所組成的三維流動表。表 3 列出臺灣父子及父女代間社會流動模型分析的配適資料。對數線性分析與其他的統計方法相同，第一個步驟是設定模型 (model specifications)；然後比較資料與模型的配適程度；接著，檢定顯著水準。模型一「條

表 3 臺灣男性與女性代間流動配適的結果：1990 資料

模型說明	df	L^2	P	BIC
(1) 條件獨立模型 (Conditional independence)	50	1069.04	.000	660.2
(2) 同質對角線獨立模型 (Homogeneous independence with diagonals)	44	287.38	.000	-72.4
(3) 條件準獨立模型 (Conditional quasi-independence)	38	280.59	.000	-30.1
(4) 同質齊一關連對角線模型 (Homogeneous uniform association with diagonals)	43	109.27	.000	-242.3
(5) 相異齊一關連對角線模型 (Heterogeneous uniform association with diagonals)	42	97.02	.000	-246.4
(6) 同質列效應對角線模型 (Homogeneous row-effects with diagonals)	39	86.65	.000	-232.2
(7) 相異列效應對角線模型 (Heterogeneous row-effects with diagonals)	34	65.51	.000	-212.5
(8) 同質行效應對角線模型 (Homogeneous column-effects with diagonals)	39	77.42	.000	-241.4
(9) 相異行效應對角線模型 (Heterogeneous column-effects with diagonals)	34	56.06	.010	-221.9
(10) 同等行列效應對角線模型 II (Equal RC effect model with diagonals [$\phi_k = \phi_l$])	39	72.33	.001	-246.5
(11) 同等行列效應對角線模型 II (Equal RC effect model with diagonals)	38	63.83	.005	-246.9
(12) 同等行列效應模型 + 親近及繼承性 (Equal RC effect model with designs)	40	93.61	.000	-233.4
(13) 同等行列效應模型 + 相異 designs (Equal RC effect model with designs by sex)	36	77.75	.000	-216.6
(14) 行列效應對角線模型 (RC effect model with diagonals)	34	43.60	.125	-234.4
模型比較				
(1) Model 2 vs. Model 3	6	6.79	.341	
(2) Model 2 vs. Model 4	1	178.11	.000	
(3) Model 4 vs. Model 5	1	12.25	.005	
(4) Model 6 vs. Model 7	5	21.41	.000	
(5) Model 8 vs. Model 9	5	21.36	.000	

附註：模型加上對角線 (models with diagonals) 表示一組的對角線的效應加諸於流動表；而準對角線 (quasi) 模型表示每個對角線的效應都完全配適。總樣本數為 3555。

件獨立模型」(conditional independence model) 是一個基準模型(baseline model)，設定行列表的起點與終點之間無顯著關連。換言之，除了受到邊際分配的影響之外，樣本會隨機地分配到流動表的格子。本模型設定男女兩個流動表的「獨立模型」(independence model) 皆成立。由於這個模型意味著完全開放的社會流動，因而，這個模型通常不配適流動表，在此也不例外。對角線(main diagonals) 上的格子代表兩代地位的傳承；一般而言，由於代間地位的繼承性都不低，所以關連程度通常比其他非對角線的互動關連(父子地位流動部份) 來得強。社會流動的研究基本上是在比較流動與不流動的差異，因而對這兩者的差異程度，特別感到興趣。第二個模型——「同質對角線獨立模型」，即是將相同的對角線傳承模式加入模型一。此模型不僅假設對角線區與非對角線區有所差異，而且假設各個階級的傳承程度是不同的；但是，兩個流動表的傳承模式則是一致的。增加使用了六個自由度，log-likelihood ratio (L^2) 大幅的下降，而且 BIC 也變成負值。爲了驗證父子及父女代間社會流動的階級傳承是否有差異，模型三加入六個參數來估計階級傳承。換言之，每個流動表都使用六個參數來測量對角線上的流動密度。²⁰ 無論是 L^2 或是 BIC 值都略降，但模型仍未配適資料。模型二及模型三的差異並不顯著 ($\Delta L^2 = 6.79$; $df = 6$; $P = .341$)，顯示男女傳承模式並無差距。第四模型使用一個參數設定兩代間的關連。「齊一關連模型」假設行列間的相關是一致的；齊一關連參數與常用於地位取得的迴歸係數相似，代表父代地位對子代地位的線性關係的影響。本模型雖然比模型二只多一個估計參數，但模型的配適程度卻提高不少， L^2 及 BIC 統計值都下降很多。整體而言，顯示父親與子代的階級地位呈現某種程度的因果關連。模型五則進一步設定齊一關連參數在男女兩個流動表上是不一樣的。比較第二個及第四個模型的配適結果，兩者的差距是由加入的參數所造成的，顯示齊一關連的效應是顯著的 ($\Delta L^2 = 178.11$; $df = 1$; $P = .000$)。而且，模型五與模型四的比較也顯示出男女在兩代關連有顯著的差別(符號 $\Delta L^2 = 12.25$; $df = 1$; $P = .000$)。然而，整個模型的配適程度並沒有改善很

多，*BIC* 統計值由 -242.3 降為 -246.4，改變不大。模型六設定男女社會流動表的行效應是相同的，而模型七則放鬆模型六的限制，將行效應設定為不一樣的。模型六與模型七的比較顯示出男女在兩代流動的外流形態 (outflow patterns) 有顯著的差別 ($\Delta L^2 = 21.41$; $df = 5$; $P = .000$)。模型八設定男女社會流動表的列效應是相同的，而模型九則放鬆模型八的限制，將列效應設定為有差異的。模型八與模型九的比較發現男女在兩代流動的內流形態 (inflow patterns) 有顯著的差別 ($\Delta L^2 = 21.36$; $df = 5$; $P = .000$)。²¹

接著，模型十在男女流動表設定行效應與列效應是一樣的，並對行列的關連 (ϕ) 加上同質性的限制 (homogeneity restrictions)。這是 *RC* 模型的一種，其特色是不受行與列類別排列所影響，減少階級排列預設上的誤差。在此，我們使用 Xie (1992) 所提的層級模型來分析，在特定的條件下，比較男女兩代流動的關連程度。根據 *BIC* 統計值，模型十比上述的模型更配適流動表。模型十一放鬆模型十對關連同質性的限制。根據傳統的假設檢定的統計值，模型的改善非常顯著；但 *BIC* 統計值並無顯著的降低。模型十二及模型十三以表 1 的設計矩陣 (design matrix) 來代替對角線的效應。這個設計矩陣是一個典型的 levels model。正如本文前述 Hauser 的 levels model 深具實驗性，這個矩陣是在階級藩籬的區隔及階級親近性的指引下，搜尋數個矩陣模型之後所得的。其中，模型十二的設計矩陣假設男性與女性在白領階級 (以數字 2 為代表) 與藍領階級 (4) 的流動有某種相同的親近性；並凸顯自營業者 (3) 及農人階級 (5) 的自我傳承的特殊性。換言之，除了其他參數之外，本模型使用五個親近性及繼承性參數去配適男性與女性的流動表。²² 模型十三則使用不同的親近性及繼承性參數去配適兩性的流動表。模型十二雖然比模型十一少使用二個參數來估計，但從 *BIC* 統計值來判斷，模型十二並沒有比模型十一有較高的配適程度。同樣地，模型十三雖然增加四個參數來配適流動表，整個模型並不是

很恰當。這顯示出加入設計矩陣的參數並未增加模型的配適程度。因此，我們繼續探尋更佳模型。

模型十四是「行列效應對角線模型」，放鬆男女流動表行效應與列效應是一樣的限制，運用不同的參數(多用了四個自由度)來估計行列效應。雖然 P 的檢定值大幅改善，但 BIC 值卻不降反升。這種現象顯示模型估計可能已使用了過多的參數。由於 BIC 值同時考慮自由度及樣本數，意味著模型十四並不是一個精簡的模型。模型十四的結果顯示了模型已經足以解釋資料了。

經過了上述的模型蒐尋，模型十一「同等行列效應對角線模型」，應為比較男女代間傳承的最佳配適模型 ($L^2 = 63.84$; $df = 38$; $P = .005$; $BIC = -246.9$)。雖然選取模型十一為最佳配適模型，而且前述的幾個模型也顯示出模型十一的跡象，但必須注意的是在分析模型中有二、三個競爭模型(如模型五、九、十)，都相當配適資料，因此，在解釋最佳模型時應避免過度引申。這個模型假設男女的代間傳承除了受一般的行列邊際分配的影響之外，階級內流動沒有性別差異。但每個階級的自我流動有所不同，行效應與列效應是一樣的；此模型又假定兩代的關連效應因性別而異。

表 4：同等行列效應對角線模型配適男女代間流動表的參數估計值

性別/階級 各種效應	1	2	3	4	5	6
關連效應 (ϕ_k)	.565	.825				
行列效應 ($\mu_i = \nu_j$)	-.416	-.517	-.208	.289	.243	.611
對角線效應	.725 (.114)	-.204 (.206)	.265 (.115)	.664 (.122)	.267 (.135)	1.932 (.178)

附註：括弧內是參數估計標準誤 (standard errors)。

表 4 列出了最配適模型 (模型十一, 即「同等行列效應對角線模型」) 的各類參數估計值。從表 4 的行列效應的係數來看, 階級的排列除了「高級非手工階級」(-.416) 與「一般非手工階級」(-.517) 的順序及「技術性手工階級」(.289) 與「非技術性手工階級」(.243) 的順序不對之外, 整體來說, 與原先的階級間的層級假設是相近的。無論如何, 本模型是屬於 *RC* 模型, 因此, 模型估計值並不受此問題的影響。從行列效應看來, 白領階級 (非手工) 與藍領階級 (手工) 有個明顯的距離 (藩籬界線)。自營者與白領階級的距離比藍領階級接近; 而農人階級 (.611) 與藍領階級 (.243; .289) 的明顯差距, 表示農人在向上流動 (upward mobility) 的負擔沈重。從關連效應的係數來看, 男女的代間傳承有程度上的差別, 臺灣父子的階級關連 (.565) 低於父女的階級關連 (.825)。換言之, 在控制了上述的因數之後, 女性的階級受其父親的影響比男性來得高。此一研究結果與過去的研究 (如蔡淑鈴與文崇一, 1985) 有所不同。²³ 對角線的係數顯示, 不分男女性別, 農人階級的自我繼承比率 (1.932) 最高, 其次是高級專業人員 (.725) 及自營業者 (.664)。其他階級的自我繼承比率則不高, 且估計值並不很顯著。

捌、結 論

本文主要是討論近十幾年來對數線性模型與對數相乘模型的發展過程及其在社會科學上的運用。起初, Goodman 的關連模型帶動了社會流動分析的新路徑。後來的發展大多是建基於 Goodman 的模型; 本文的另一個重點在於介紹一些基於 Goodman 的 *RC* 模型所延伸的分析模型, 其估計值不受分類項目次序的影響, 並可使用社會流動的形態來分析階級間的相對垂直關係。這些模型又可加入等級模型參數及交錯效應參數來分析流動表, 檢驗流動表內的流動密度。本文除了討論分析單一的二維流動表的模型之外, 也討論一些研究如何將二維模型

加上多層級的設定，而成爲一種限制性的三維線性模型。由數個二維流動表所組成的三維流動表分析可適用於跨組(國)的比較分析。換言之，在相同設定的條件下，可直接精準地比較流動表參數的差異。另外，本文使用「臺灣地區社會變遷基本調查」的資料所得的男性與女性代間社會流動表，以最近發展出來的一系列對數線性與對數相乘模型來分析相對的社會流動形態。從資料的分析過程中，顯示對數線性模型在分析社會流動對結構性因素模型建構時，相當具有伸縮性。對數線性模型已逐漸爲臺灣的社會科學界所接受及應用。過去僅能從事簡單的關連分析的列聯表資料，對數線性模型已能從事更深入的分析，以探究更詳細的關連結構。本文所討論的面向，雖然以社會流動爲主，但是本文所討論的模型及方法都可應用於社會科學的其他類別資料。有些學門及研究取向所蒐集的資料含有特別多的類別資料，例如政黨選舉與政治，特別適用於以對數線性的分析，來探究一些深層的關係(參見 Weakliem, 1991利用 RC (M) 模型分析歐洲左右政黨的區隔化與選民職業的關係，或 De Graaf, Nieuwebeerta, and Heath, 1995的分析社會流動與政治取向的關係)。對數線性分析提供臺灣的政治學及社會學者另一個分析臺灣豐富的選舉資料的方法。

迴歸分析與對數線性分析是研究社會流動的兩大研究方法。社會流動的研究取向，受到新的統計技術引進的影響甚大。最近的發展，例如 RC (M) 模型及由 Xie (1992) 所提出的對數相乘層級效應模型，一方面以增加多層次的分析，一方面提高了層級間比較的靈活性。從以上的討論，顯示對數線性模型能有效地提供研究者一個分析多面向結構關係的工具。但是，就多變項的模型建構而言，對數線性分析方法卻無法如迴歸分析能納入多個獨立變項，對於那些社會因素造成流動表上的分配，難以提出一個有效的因果關係的解釋理論。使用對數線性的研究者，曾試圖突破這種限制，有不少社會流動的比較分析，如本文分析男女社會流動的例子，使用條件式的模型 (conditional models) (Clogg, 1982b 對此方法有詳細的說明) 來加入並探討其他相關

變項 (例如性別、時間及地區等因素) 的影響。還有些研究如 Hout (1984) 的 SAT 模型及 Stier and Grusky (1990) 的 OP 模型，則更進一步嘗試「間接地」將相關變項加入對數線性模型。DiPrete (1990) 可算是第一位正式直接將一般地位取得模型所常用的變項加入對數線性模型的研究者。雖然，DiPrete (1990) 所提出的模型在分析上，仍有其技術上的困難。但將迴歸分析及對數線性模型的優點相結合，是一個非常值得開發的研究路徑。

註 釋

- 1 作者僅就個人所涉獵的研究論文來舉例，大多限於社會學的研究。與對數線性模型相關的分析方法，也被社會學家所重視並廣泛的運用，例如 canonical correlation (其應用參見如 Wasserman and Faust, 1989) 及 correspondence analysis (參見如 Bourdieu, 1984: 226-83)。
- 2 有關這些模型的公式及設定，請參見 Goodman (1979)，此文也登印於 Goodman (1984) 的選集。
- 3 Goodman 所提出的模型非常之多，不僅止於本文所介紹的。在社會學的研究，Goodman 可說是對數線性分析的開山祖師，絕大部分的新創模型都是立基於 Goodman 所發展出來的模型。
- 4 例如，上層階級為維護其利益，可能抵擋並降低其他下層階級流入的機會；有些階級間可能因某些特殊的因素而使得它們之間的流動比較容易。
- 5 另一個例子是 Duncan 為 Goodman (1984) 的選集撰寫序言，讚賞 Leo Goodman (及 Clifford Clogg) 的成就，並提倡使用 Goodman 的模型來分析行列表。
- 6 本文主要是討論 log-linear 模型的最新發展，對於 log-linear 的一些基本概念，請參考 Knoke and Burke (1980), Agresti (1984; 1990)。

- 7 Hope (1982) 的「鑽石模型」將「獨立模型」分成行與列兩部分，即他所謂的「半面模型」(halfway model) 及「差異模型」(difference model)，是比「獨立模型」更小(更簡單)的分析模型，打破了一般對數線性分析以獨立模型為基準模型的原則。由於過度的精簡，Hope 的模型缺乏分析複雜模型的靈活性。近十多年來，對數線性分析的發展顯然是建立在 Goodman 的關連模型，包括他的同事 Goldthorpe (參見 Erikson and Goldthorpe, 1992) 在內的研究者，幾乎完全忽略了 Hope 在分析模型建構上的貢獻。
- 8 獨立模型又稱為完全流動模型(perfect mobility model)，因為模型假設行列表的行與列沒有顯著的關連，由某一特定的位置流動到任何位置的機率是一樣的。準獨立模型是一個對角線(diagonals)上的互動關連完全配適的獨立模型。
- 9 Goodman 的關連模型有助於決定排列的順序。在社會科學的資料分析上的運用，請參考 Clogg (1982a)。
- 10 本文依循慣例，以行列表(列聯表)的行(i)來代表「起點」(origins)；以列(j)來代表「終點」(destinations)。若以 R 代表行；以 C 代表列，則等式 1 應精確地寫成 $\log F_{ij} = \lambda + \lambda_i^R + \lambda_j^C + \Delta + \Psi$ 。為了簡單化，本文省略行與列的上標。
- 11 本文中，所有的列舉模型都必須在規格化設限(normalization)的條件下才能成立。有關的設限，請參考相關論文或 Goodman 1984, Agresti 1990。
- 12 GLIM 能將隨機(誤差)因數的分配設定成 Poisson, Normal, Binomial, Inverse-Gaussian 及 Gamma 分配，用處甚廣。
- 13 Agresti (1990) 的附錄 A 列有各式套裝軟體的樣本程式，用於分析對數線性模型，可供參考。
- 14 由於本文以討論對數線性分析技術為主，不擬討論兩性的社會流動的研究及文獻。有興趣的讀者，可參考 Abbot and Sapsford, 1987; Acker, 1973。有關臺灣的研究，請參考蔡淑鈴、文崇一，1985。

- 15 Goldthorpe 等研究者將歐美及日本十多個國家的就業資料，利用 EGP 的分類架構將原始資料重新分類。這個跨國的大型的研究計劃稱為 CASMIN，主要的研究目標是比較工業化國家的社會結構及社會流動。
- 16 感謝 Harry Ganzeboom 教授提供資料轉換的程式及建議。在轉換過程中，我曾將程式稍加修改，以符合臺灣社會的一些特殊性。但是，這些改變並不會影響分析結果。
- 17 Goldthorpe 等人在從事比較研究上（參見 Erikson and Goldthorpe, 1992），由於有些國家的分析樣本數不足，所以通常不使用整個分類架構（full version），而採用七個階級的縮減架構（collapsed version）。臺灣的階級研究之中，有些研究採用 Goldthorpe 分類架構分析階級關係，也由於資料的不足，必須使用較小的版本（吳乃德，1994；許嘉猷，1994；So and Hsiao, 1994）。
- 18 許嘉猷（1990; 1994）使用 Wright 的分類架構，並從事跨國的比較；最近又將 Wright 的「階級結構與階級意識」的比較研究，用於臺灣社會，做進一步的國際比較分析，詳見許嘉猷（1994）。
- 19 階級的研究中，分類架構的選定，一直是關鍵的課題；因為，分類的架構在某種程度裡已經決定了分析的結果（參見 Breiger, 1981; Goodman, 1981）。雖然，社會學界應致力於發展出一套最佳的分類架構（例如 Breiger, 1981; 1990）；但是過度批判分類架構又無法提出一套完美的分類原則，並無助於階級研究的進展。就目前而言，基於科學強調分析的可重複性的原則，研究者應清楚交代選用分類架構的理由及架構設定的步驟，使得將來的研究者能夠加以檢證批評，並比較不同分類架構的優劣及其影響。
- 20 若以前述 Hauser（1978）的 levels model 來解釋，模型一將列聯表分成流動部份與非流動部份兩個區域，假設各有其不同的關連密度。模型二則將非流動部份再劃分成六個區域，也是各有其關連密度。
- 21 這幾個模型通常都不配適資料，列出來只是供做參考。一般的學者，特別是美國的研究者習慣將相關的模型一一列出；有些則只列出少數較顯著的模型（如 Erikson and Goldthorpe, 1992）。

- 22 由於其中有一個參數是多餘的 (redundant)，因此，僅需要運用四個自由度。
- 23 本文以研究方法為主，對於社會流動的分析結果，並不擬深入討論。作者希望能針對此問題，另文討論。有關分析結果的差異，有可能是因時間相隔長久 (兩個調查 — 1978與1990 — 相差十餘年)；本文所使用的階級分類與他們的行業分類也不易直接比較；其次，本文所使用的資料較大且模型較複雜，也限制了兩者比較的可行性。

參考資料

吳乃德

- 1994 〈階級認知與階級認同：比較瑞典、美國、臺灣，和兩個階級架構〉，許嘉猷 (主編)，《階級結構與階級意識比較研究論文集》，頁109-150。臺北：中央研究院歐美研究所。

孫清山、黃毅志

- 1994 〈社會資源、文化資本與地位取得〉，《東海學報》35: 127-150。

許嘉猷

- 1982 〈出身與成就：臺灣地區的實證研究〉，見陳昭南、江玉龍、陳寬政 (主編)，《社會科學整合論文集》，頁265-299。臺北：中央研究院三民主義研究所。
- 1987 〈臺灣的階級結構〉，《中國社會學刊》11: 25-60。
- 1990 〈臺灣的階級流動及其與美國的一些比較〉，《中國社會學刊》14: 1-30。
- 1994 《階級結構與階級意識比較研究論文集》(主編)。臺北：中央研究院歐美研究所。

張維安、王德睦

- 1983 〈社會流動與選擇性婚姻〉，《中國社會學刊》7: 191-214。

蔡淑鈴

1986 〈職業地位結構——臺灣地區的變遷研究〉，見瞿海源、章英華(主編)，《臺灣社會與文化變遷(上冊)》，頁299-351。臺北：中央研究院民族學研究所。

1987 〈社會地位取得：山地、外省及閩客之比較〉，見楊國樞、瞿海源(主編)，《變遷中的臺灣社會》，頁1-44。臺北：中央研究院民族學研究所。

1994 〈臺灣之婚姻配對模式〉，《人文及社會科學集刊》6: 335-371。

蔡淑鈴、文崇一

1985 〈性別與社會流動〉，《中央研究院民族研究所集刊》10: 121-147。

蔡淑鈴、瞿海源

1992 〈臺灣教育階層化變遷〉，《人文及社會科學研究彙刊》10: 98-118。

蔡瑞明

1995 《臺灣、美國、日本社會流動的比較研究》，國科會研究計劃報告。

瞿海源

1983 〈勞力市場與出身對成就之影響〉，《中央研究院民族研究所集刊》53: 133-152。

1991 《臺灣地區社會變遷基本調查計畫：第二期第一、二次調查計畫執行報告》(主編)。臺北：中央研究院民族學研究所。

謝雨生、余淑嬪

1990 〈臺灣的社會階級及其流動〉，《中國社會學刊》14: 31-63。

Abbot, Pamela and Roger Sapsford

1987 *Women and Social Class*. London: Tavistock.

Acker, Joan

1973 "Women and Social Stratification: A Case of Intellectual Sexism," *American Journal of Sociology* 78: 936-45.

Agresti, Alan

1984 *Analysis of Ordinal Categorical Data*. New York: Wiley.

1990 *Categorical Data Analysis*. New York: Wiley.

Aitkin, Murray

1989 *Statistical Modelling in GLIM*. London: Falmer Press.

Aptech Systems

1991 *GAUSS 2.1 System and Applications Manuals*. Kent, WA: Aptech System Inc.

Becker, Mark P. and Clifford C. Clogg

1989 "Analysis of Sets of Two-Way Contingency Tables Using Association Models," *Journal of the American Statistical Association* 84: 142-51.

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul M. Holland

1975 *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass.: MIT Press.

Blau, Peter M. and Otis Dudley Duncan

1967 *The American Occupational Structure*. New York: Wiley.

Bourdieu, Pierre

1984 *Distinction: A Social Critique of the Judgement of Taste*, translated by Richard Nice. Cambridge, Mass.: Harvard University Press.

Breen, Richard

1984 "Fitting non-hierarchical and Association Log-linear Models using GLIM," *Sociological Methods and Research* 13: 77-107.

Breiger, Ronald L.

1981 "The Social Class Structure of Occupational Mobility," *American Journal of Sociology* 87: 578-611.

1982 "A Structural Analysis of Occupational Mobility," in Peter V. Marsden and Nan Lin (eds.), *Social Structure and Network Analysis*, pp.17-32. Beverly Hills, CA: Sage Publications.

- 1990 "Intermediate Classes and Social Structure," in Ronald L. Breiger (ed), *Social Mobility and Social Structure*, pp.225-246, Cambridge: Cambridge University Press.

Clogg, Clifford. C.

- 1982a "Using Association Models in Sociological Research: Some Examples," *American Journal of Sociology* 88: 114-34.
- 1982b "Some Models for the Analysis of Association in Multiway Cross-Classification Having Ordered Categories," *Journal of the American Statistical Association* 77: 803-15.

Clogg, Clifford C., Scott R. Eliason and Robert J. Wahl

- 1990 "Labor -Market Experiences and Labor-Force Outcomes," *American Journal of Sociology* 95: 1536-76.

De Graaf, Nan Dirk, Paul Nieuwbeerta and Anthony Heath

- 1995 "Class Mobility and Political Preferences: Individual and Contextual Effects," *American Journal of Sociology* 100: 997-1027.

DiPrete, Thomas A.

- 1990 "Adding Covariates to Loglinear Models for the Study of Social Mobility," *American Sociological Review* 55: 757-773.

Duncan, Otis D.

- 1979 "How Destination Depends on Origin in the Occupational Mobility Table," *American Journal of Sociology* 84: 793-803.

Eliason, Scott R.

- 1990 *The Categorical Data Analysis System Version 3.50 User's Manual*. Department of Sociology, University of Iowa.

Erikson, Robert and John H. Goldthorpe

- 1987a "Communality and Variation in Social Fluidity in Industrial Nations, Part I," *European Sociological Review* 3: 54-77.
- 1987b "Communality and Variation in Social Fluidity in Industrial Nations, Part II," *European Sociological Review* 3: 145-66.

- 1992 *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Clarendon Press.
- Erikson, Robert John H. Goldthorpe, and Lucienne Portocarero
 1979 "Intergenerational Class Mobility in Three Western European Societies," *British Journal of Sociology* 30: 415-441.
- Featherman, David L. and Robert M. Hauser
 1978 *Opportunity and Change*. New York: Academic Press.
- Featherman, David L., F. Lancaster Jones, and Robert M. Hauser
 1975 "Assumptions of Social Mobility Research in the U.S.: The Case of Occupational Status," *Social Science Research* 4: 329-60.
- Ganzeboom, Harry B.G., Ruud Luijkx, and Donald J. Treiman
 1989 "Intergenerational Class Mobility in Comparative Perspective," *Research in Social Stratification and Mobility* 8: 3-84.
- Gilula, Zvi and Shelby J. Haberman
 1986 "Canonical Analysis of Contingency Tables by Maximum Likelihood," *Journal of the American Statistical Association* 88: 780-88.
- Goldthorpe, John H., C. Llewellyn and C. Payne
 1987 *Social Mobility and Class Structure in Modern Britain* (2nd Edition). Oxford: Clarendon Press.
- Goodman, Leo A.
 1979 "Some Multiplicative Models for the Analysis of Occupational Mobility Tables and Other Kinds of Cross-Classification Tables," *American Journal of Sociology* 84: 804-19.
 1981 "Criteria for Determining Whether Certain Categories in a Cross-Classification Table Should Be Combined," *American Journal of Sociology* 87: 612-650.
 1984 *The Analysis of Cross-Classified Data Having Ordered Categories*. Cambridge, Mass.: Harvard University Press.

- 1986 "Some Useful Extensions of the Usual Correspondence Analysis Approach and the Usual Log-Linear Models Approach in the Analysis of Contingency Tables," *International Statistical Review* 54: 243-309.

Greene, William H.

- 1991 *LIMDEP Version 6 Manual*. New York: Econometric Software Inc.

Hagenaars, Jacques A.

- 1990 *Categorical Longitudinal Data*. Newbury Park, CA: Sage Publications.

Haberman, Shelby J.

- 1974 *The Analysis of Frequency Data*. Chicago: University of Chicago.

Hauser, Robert M.

- 1978 "A Structural Model of the Mobility Table," *Social Forces* 56: 919-53.
- 1979 "Some Exploratory Methods for Modeling Mobility Tables and Other Cross-Classified Data," in David R. Heise (ed.), *Sociological Methodology*, 1980, pp.141-158. San Francisco: Jossey-Bass.
- 1984 "Vertical Class Mobility in England, France, and Sweden," *Acta Sociologica* 27: 87-110.

Hout, Michael

- 1982 "The Association Between Husbands' and Wives' Occupations in Two-Earner Families," *American Journal of Sociology* 88: 397-409.
- 1983 *Mobility Tables*. Beverly Hills, CA: Sage Publications.
- 1984 "Status, Autonomy and Training in Intergenerational Mobility," *American Journal of Sociology* 89: 1379-1409.

Hout, Machael, Otis D. Duncan, and Michael E. Sobel

- 1987 "Association and Heterogeneity: Structural Models of Similarities and Differences," in Clifford C. Clogg (ed.), *Sociological*

Methodology 1987, pp.145-84. Washington, DC: American Sociological Association.

Hope, Keith

1982 "Vertical and Nonvertical Class Mobility in Three Countries," *American Sociological Review* 47: 99-113.

Jacobs, Jerry A. and Ronald L. Breiger

1988 "Careers, Industries, and Occupations," in George Farkas and Paula England (eds.), *Industries, Firms, and Jobs*, pp.43-63. New York: Plenum Press.

Johnson, Robert A.

1980 *Religion Assortative Mating in the United States*. New York: Academic Press.

Kalmijn, Matthijs

1991 "Status Homogamy in the United States," *American Journal of Sociology* 97: 496-523.

Knoke, David and P. J. Burke

1980 *Log-Linear Models*. Beverly Hills, CA: Sage.

Landale, Nancy S. and Avery M. Guest

1990 "Generation, Ethnicity, and Occupational Opportunity in Late 19th Century America," *American Sociological Review* 55: 280-296.

Liedka, Raymond V.

1991 "Who Do You Know in the Group? Location of Organizations in Interpersonal Networks," *Social Forces* 70: 455-474.

Marsden, Peter V.

1981 "Models and Methods for Characterizing the Structural Parameters of Groups," *Social Networks* 3: 7-27.

McCullagh, P. and J.A. Nelder

1989 *Generalized Linear Models* (2nd Edition). London: Chapman and Hall.

Numerical Algorithms Group

1993 *The GLIM System Release 4 Manual*. Oxford: Oxford University Press.

Pöntinen, Seppo

1982 "Models and Social Mobility Research: A Comparison of Some Log-Linear Models of a Social Mobility Matrix," *Quality and Quantity* 16: 406-21.

Raftery, Adrian E.

1986 "Choosing Models for Cross-Classifications," *American Sociological Review* 51: 145-46.

Rogoff, Natalie

1953 *Recent Trends in Occupational Mobility*. Glencoe, Ill.: Free Press

Sheu, Jia-You

1989 "The Class Structure in Taiwan and Its Changes," in Hsin-Huang M. Hsiao, Wei-Yuan Cheng and Hou-Sheng Chan (eds.), *Taiwan: A Newly Industrialized State*, pp.117-150. Taipei: Department of Sociology, National Taiwan University.

Smith, Herbert L. and Maurice A. Garnier

1987 "Scaling via Models for the Analysis of Association: Social Background and Educational Careers in France," in Clifford C. Clogg (ed.), *Sociological Methodology 1987*, pp.205-45. Washington, D.C.: American Sociological Association.

Snipp, C. Matthew

1985 "Occupational Mobility and Social Class: Insight from Men's Career Mobility," *American Sociological Review* 50: 475-493.

So, Alvin and H. H. Michael Hsiao

1994 "The Making of the Middle Classes in East Asia: Some Tentative Hypotheses," Paper Presented in International Conference on East Asian Middle Classes and National Development in Comparative

Perspective in Institute of Ethnology, Academia Sinica, Taipei, Taiwan, December 19-21.

Sobel, Michael

1988 "Some Models for the Multiway Contingency Table with a One-to-one Correspondence Among Categories,," in Clifford C. Clogg (ed.), *Sociological Methodology 1988*, pp.165-192. Washington, D.C.: American Sociological Association.

Sobel, Michael E., Michael Hout, and Otis D. Duncan

1985 "Exchange, Structure, and Symmetry in Occupational Mobility," *American Journal of Sociology* 91: 359-72.

Stier, Haya and David B. Grusky

1990 "An Overlapping Persistence Model of Career Mobility," *American Sociological Review* 55: 736-56.

STSC

1989 *APL*PLUS System for the PC User's Manual*. Rockville, MD: STSC Inc. (現已改名為 Manugistics, Inc)

Treiman, Donald J.

1977 *Occupational Prestige in Comparative Perspective*. New York: Academic Press.

Tsai, Shu-Ling and Hei-Yuan Chiu

1993 "Changes in Educational Stratification in Taiwan," in Yossi Shavit and Hans-Peter Blossfeld (eds.), *Persistent Inequality*, pp.193-228. Boulder, CO: Westview Press.

Tsay, Ruey-Ming

1995 "Economic Development and Political Attitudes: The Formation of Political Authoritarianism in Taiwan," *Tunghai Journal of Social Science* 36: 205-236.

1996a "Who Marries Whom? The Associations Between Wives' and Husbands' Educational Attainment and Class in Taiwan,"

Proceedings of National Science Council, ROC. Part C: Humanities and Social Sciences 6: 258-277. (國家科學委員會研究彙刊：人文及社會科學)

- 1996b “Leaving Farmland: Class Structure Transformation and Social Mobility in Taiwan,” 將刊登於張荳雲 (主編), 《九〇年代的臺灣社會》。臺北：中央研究院社會學研究所。

Wasserman, Stanley and Katherine Faust

- 1989 “Canonical Analysis of the Composition and Structure of Social Networks,” in Clifford C. Clogg (ed.), *Sociological Methodology 1989*, Pp.1-42. Washington, D.C.: American Sociological Association.

Weakliem, David L.

- 1991 “The Two Lefts? Occupation and Party Choice in France, Italy, and the Netherlands,” *American Journal of Sociology* 96: 1327-61.
- 1992 “Comparing Non-Nested Models for Contingency Tables,” in Peter V. Marsden (ed.), *Sociological Methodology 1992*, pp.147-178. Washington, D.C.: American Sociological Association.

Weber, Max

- 1978 “Status Groups and Classes, in Economy and Society,” in Guenther Roth and Claus Wittich (eds.), *Economy and Society*, pp.302-307. Berkeley: University of California Press.

Western, Mark and Erik Olin Wright

- 1994 “The Permeability of Class Boundaries to Intergenerational Mobility Among Men in the United States, Canada, Norway and Sweden,” *American Sociological Review* 59: 606-629.

Wong, Raymond Sin-Kwok

- 1990 “Understanding Cross-National Variation in Occupational Mobility,” *American Sociological Review* 55: 560-73.

- 1992 "Vertical and Nonvertical Effects in Class Mobility: Cross-National Variations," *American Sociological Review* 57: 396-410.
- Wright, Erik Olin
- 1985 *Classes*. New York: Verso.
- Wright, Erik Olin and Donmoon Cho
- 1992 "The Relative Permeability of Class Boundaries to Cross-Class Friendships: A Comparative Study of the United States, Canada, Sweden and Norway," *American Sociological Review* 57: 85-102.
- Wright, Erik Olin and L. Perrone
- 1977 "Marxist Class Categories and Income Inequality," *American Journal of Sociology* 42: 32-55.
- Wu, Naiteh
- 1994 "Class Awareness and Class Identity," in Jia-You Sheu (ed.), *Comparative Studies of Class Structure and Class Consciousness*, pp.109-150. Taipei: Institute of European and American Studies, Academia Sinica. (in Chinese)
- Xie, Yu
- 1992 "The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables," *American Sociological Review* 57: 380-395.
- Yamaguchi, Kazuo
- 1983 "The Structure of Intergenerational Occupational Mobility: Generality and Specificity in Resources, Channels, and Barriers," *American Journal of Sociology* 88: 718-745.
- 1987 "Model for Mobility Tables: Toward Parsimony and Substance," *American Sociological Review* 52: 482-94.
- Yasuda, S.
- 1964 "A Methodological Inquiry Into Social Mobility," *American Sociological Review* 29: 16-23.

附 錄

GLIM配適同等行列效應對角線模型的指令範例：

```
$C This is a GLIM program to estimate intergenerational class
$C Mobility: Men vs. Women, Taiwan 1990.
$C Equal Row and Column Effect Model with Diagonals
$C Sector factor is added.
$C Date: 2-19-1995
$echo
$units 72 $data Freq $read
```

[DATA FILE]

```
$data sect $read
```

[DESIGN MATRIX]

```
$calc %O=6:%D=6:%T=2
$calc d=%gl(6,1):o=%gl(6,6):c=%gl(%t,6*6)
$calc inh = %if(%eq(o,d),2,1)
$calc diag= %if(%eq(o,d),o,0):diag=diag+1
$FACTOR O %O D %D DIAG 7 INH 2 C 2
$CALC %N=%CU(FREQ)
$Macro model O*C+D*C+diag.C
$ENDMac
$MAC STAT
$CALC %B=%DV-%DF*%LOG(%N)
```

```

$CALC %U=%CU(%SQRT((%FV-FREQ)**2))
$CALC %X=%CU(%FV):%U=50*(%U/%X)
$PRINT 'LOG LIKELIHOOD RATIO' %DV
$PRINT 'PEARSON CHI-SQUARE' %X2
$PRINT 'DEGREE OF FREEDOM' %DF
$PRINT 'SAMPLE SIZE' %N
$PRINT 'INDEX OF DISSIMILARITY' %U
$PRINT 'BIC STATISTIC' %B
$ENDMAC
$MAC ST2
$CALC %B=%DV-(%DF-4)*%LOG(%N):%Q=%DF-4
$CALC %U=%CU(%SQRT((%FV-FREQ)**2))
$CALC %X=%CU(%FV):%U=50*(%U/%X)
$PRINT 'LOG LIKELIHOOD RATIO' %DV
$PRINT 'PEARSON CHI-SQUARE' %X2
$PRINT 'DEGREE OF FREEDOM (corrected)' %Q
$PRINT 'SAMPLE SIZE' %N
$PRINT 'INDEX OF DISSIMILARITY' %U
$PRINT 'BIC STATISTIC (corrected)' %B
$ENDMAC
$MAC K2
$CAL %I=%I+1 $
$PRINT 'Iteration' *2 %I
$CALC V1=0
$CALC V2=((%EQ(O,2))*H(D)+(%EQ(D,2))*H(O))/2
$CALC V3=((%EQ(O,3))*H(D)+(%EQ(D,3))*H(O))/2
$CALC V4=((%EQ(O,4))*H(D)+(%EQ(D,4))*H(O))/2
$CALC V5=((%EQ(O,5))*H(D)+(%EQ(D,5))*H(O))/2
$CALC V6=((%EQ(O,6))*H(D)+(%EQ(D,6))*H(O))/2

```

```

$FIT V1+V2+V3+V4+V5+V6+#model $
$EXTRACT %PE
$CALC H1(O)=%PE(1+O) $CALC H=(H+H1)/2
$CALC F=(PE-%PE)**2:%F=%CU(F) $CALC PE=%PE
$CALC %S=%GE(%F,10**%C)
$ENDMAC
$ACCU 5
$YVAR FREQ $ERROR P $LINK L
$MAC INIT
$CALC %C=-4:%P=40 $CALC %S=1:%I=0 $VAR %P PE
$CALC PE=100
$CALC H(O)=O $WHILE %S K2
$PRINT 'Program converged after ' *2 %I ' Iteration'
$USE STAT$
$VAR %D SE$
$CALC SE(O)=H1(O) $CALC %M=%CU(SE)/%D
$CALC SM=(SE-%M)
$CALC %V=%SQRT(%CU(SM**2)) $CALC S1=SM/%V
$CALC %H=%CU(SM**2)
$PRINT : 'Fully Normalized Scale Values (Mean=0; Sum of Square=1):'
$PRINT S1 $
$PRINT : 'Partially Normalized Scale Values (Mean=0):'
$PRINT SM $
$PRINT : 'Phi Coefficient:'
$PRINT %H $
$C The following part will estimate the standard error of the Phi
$C Coefficients using Row and Column Effect scores
$CALC RS=S1(D):CS=S1(O):PH1=RS*CS
$FIT #model + PH1

```

```
$USE ST2 $DISP M E $  
$ENDMAC  
$USE INIT $  
$STOP
```

團體比較研究中結構差異的考慮與處理： 論迴歸標準化及對數線性消除法的應用*

謝雨生** 鄭宜仲***

壹、前言

社會科學研究中，研究者常會需要進行團體間的比較。這裡所說的團體，可能是不同國家或社會，也可能是同一個社會的不同地區或不同群體、族群，亦有可能是同一群體的不同時間點，因此，不同時間點也可視為比較研究中的不同群體。此外，依研究性質的不同，團體間的比較研究，大致可分為三類：

- (一) 比較某一解釋變項對依變項影響效果的團體差異。譬如，比較男女性的教育投資報酬率 (return to schooling) 的差異。
- (二) 比較某一依變項平均數的團體差異及其成因。比較男女性平均薪資的差異及造成差異的原因等研究即屬於這一類的比較研究。

* 本論文所採用的部分資料來自於國家科學委員會所支助，瞿海源教授主持之「臺灣地區社會變遷基本調查第二期第三次」調查結果。感謝研討會評論人謝靜琪教授及兩位匿名審查人所提出的寶貴意見，特此誌謝。本論文如有任何疏漏之處，由作者負責。

** 臺灣大學農業推廣學系教授。

*** 臺灣大學農業推廣學系博士候選人。

(三) 探討某一社會現象 (以現象率表示之) 的團體差異。如比較男女性離婚率、失業率或人力適當運用率等社會現象的差異，或分析臺灣離婚率、失業率或人力適當運用率等社會現象的變動等，則屬於此類之比較研究。

上述第一類的比較研究，研究者可以直接檢驗個別團體迴歸方程式中自變項 (如教育) 對依變項 (如所得) 的迴歸係數母數是否有差異；或是使用互動項 (性別×教育) 於團體資料合併後的迴歸分析中，再行檢驗互動項的母數是否為零，就可以回答解釋某一自變項對依變項影響效果的團體差異問題。

除了第一類比較研究的分析較為單純和直接外，其餘兩類的比較研究皆牽涉到團體間可能存在結構性差異的問題。就上述第二類的研究而言，當研究者在進行依變項平均數之團體差異比較時，其前提條件是團體間同一自變項的分配狀態必須相同或者相似。因為即使某一自變項對依變項的影響效果是相同的，而不同團體之該自變項的平均數不同時，也會自然地產生不同的依變項平均數。這種因為自變項的團體差異而自然誘導出的依變項平均數差異即為「團體結構性差異效果」。

若研究者所進行的比較研究屬於上述的第三類研究的話，通常研究者是在比較各團體之現象率 (rate of event) 或現象粗率 (crude rate) 的差異。但是粗率同時決定於兩個重要成份的變化，即團體的組成分配 (distribution of composition) 與組成特殊率 (composition-specific rate)。也就是說，即使兩個團體有相同的組成特殊率，一旦他們具有不同的組成分配，則他們還是會有不同的現象粗率產生。

因此，研究者如果沒有將上述的這種團體間的結構差異予以考慮，並在團體比較時予以必要且適當的處理，則常會使得整個研究得到偏離事實真相的研究結論。這種團體結構性差異沒有給予正確處理的情形在社會科學研究中卻經常可以發現。本論文的目的是希望藉由

研究實例的說明，討論並提供兩種團體比較研究中團體結構差異的考慮與處理的方法，以分別滿足上述第二類及第三類團體比較研究的資料分析需求。其中一種方法適用於上述第二類的比較研究，稱之為迴歸標準化 (regression standardization) 或迴歸調整法 (regression adjustment method)。另一種方法則是屬於非連續變項的處理方法，適用於上述第三類之比較研究，稱之為對數線性消除法 (log-linear purging method) (Clogg, 1978; Clogg and Shockey, 1985; Clogg and Eliason, 1988; Xie, 1989)。前者以多元迴歸分析為基礎的調整法，而後者則以對數線性模式 (log-linear model) 為主的調整法。

貳、團體結構差異及處理

(一) 迴歸標準化或迴歸調整法

一般而言，研究者在探討團體間某一依變項平均數的差異時，往往會進一步解析差異的原因。因為依變項平均數的團體差異是可觀察到的差異 (observed difference) 或是粗差異 (crude difference)，不見得是代表團體間的真正差異 (true difference)。所以，研究者為了能掌握團體間依變項的真正差異，往往必須有效地控制或消除團體間的結構性差異才能奏效，也才不致於下了偏差的研究結論。換言之，研究者在進行前述的第二類之比較研究時，常需要將團體間依變項的粗差異分解為團體結構差異所造成的差異部分和非團體結構差異所形成的差異部分 (詳細情形容後再進一步討論)。

基本上，當依變項受到一群自變項的影響，且研究者在進行依變項平均數差異的比較時，必須基於團體間的同自變項需具有相同或相似的分配狀態時為之，才能顯現出團體間依變項的真正差異。否則，如果團體間自變項的立基點不同 (即分配狀態不同) 之情形下，所進行的依變項平均數差異之比較結果，容易造成研究結論的偏差。因此，研究者必須對團體的各自變項之分配狀態予以必要的調整，使團

體間的同一直變項分別具有相同或相似的分配狀態，以為具有基準點的有效比較。

一個團體的任一自變項之分配狀態可以由其平均數、標準差、偏態及峰度四個母數(或參數)予以呈現出來。因此，理想上研究者應該使同一個解釋變項在團體間均具有相同的分配狀態後，再進行調整後依變項平均數的比較。但是，同時考慮上述四個母數的調整方式之團體比較並不容易，且在經驗研究上也未發現。然而，不論是在經濟學或社會學的經驗研究上，只調整自變項平均數的團體比較方式卻是相當常見，尤其是在不平等(inequality)的經驗研究中更是常見。這些研究的內容則涵蓋男女性平均薪資或所得的差異研究(Semyonov, 1988; Wellington, 1994)、不同勞動力市場平均薪資差異及性別或種族歧視的研究(Rosenfeld and Kalleberg, 1990; Semyonov, 1988; Semyonov and Cohen, 1990)或貧窮的性別差異及其跨國比較(Casper and McLanahan, 1994)等。

在團體比較研究中，即使某一自變項對依變項的影響效果是相同的，但是不同團體在該自變項的平均數不同，也會自然地產生不同的依變項平均數。這種解釋變項就具有團體差異，而自然誘導出的依變項平均數差異，稱之為「團體結構性差異效果」或「結構差異效果」。因此，在經驗研究中，研究者通常將團體間依變項平均數之粗差異加以調整，進而將粗差異分解為三個或四個不同性質的成份，其中一個成份就是團體結構差異效果。而「調整」的基本精神是假定被比較的團體具有和所選擇的標準團體(或參考團體)完全一樣的自變項平均數時，再計算出被比較團體之依變項平均數，即為調整後平均數。原來有結構差異的團體，其調整後平均數將會有所改變(標準團體者則不會改變)。基於這種調整假定，再經過適當的數學運算，團體間的粗差異可以被分解為三個或四個不同性質的成份，如式 1 至式 3 所示。原則上，粗差異的分解結果取決於標準團體的選擇和交互作用效

果的考慮與否。再者，標準團體的選擇和交互作用效果的考慮與否又受到研究者之研究目的或研究企圖的影響。

為方便說明粗差異的分解結果，本文將只討論兩個團體（以男、女性為例）比較的情形。底下將只列出分解後的結果，詳細的分解過程則請參見附錄一。依變項（如所得）平均數的性別差異可以分解為四個部份，分別為（1）截距的性別差異，（2）自變項的影響量或報酬率之性別差異所造成的差異，（3）自變項平均數的性別差異所造成的差異及（4）平均數差異和報酬率差異之交互作用所造成的差異。其數學式如式 1 所示（Rosenfeld and Kalleberg, 1990; Iams and Thornton, 1975; Jones and Kelley, 1984）。

$$\begin{aligned} \text{Ln}\bar{y}_m - \text{Ln}\bar{y}_f &= \frac{(a_m - a_f)}{(1)} + \frac{\sum \bar{X}_{if}(b_{im} - b_{if})}{(2)} \\ &+ \frac{\sum b_{if}(\bar{X}_{im} - \bar{X}_{if})}{(3)} \\ &+ \frac{\sum (b_{im} - b_{if})(\bar{X}_{im} - \bar{X}_{if})}{(4)} \end{aligned} \quad (1)$$

其中 \bar{y} 為依變項（所得）的平均數。m 和 f 分別表示男性和女性（團體）。a 是迴歸方程式中的常數項， \bar{X} 和 b 分別為自變項平均數和迴歸母數估計值（或未標準化迴歸係數）。另外，i 表示不同的自變項。

若是研究者只選擇其中一團體為標準團體（先假設為男性），則兩個團體依變項平均數粗差異可以分解如式 2。此時，原式 1 中的第（2）項和第（4）項在式 2 中結合成一項。

$$\begin{aligned}
 \bar{Lny}_m - \bar{Lny}_f &= \frac{(a_m - a_f)}{(1)} + \frac{\sum \bar{X}_{im}(b_{im} - b_{if})}{(2) + (4)} \\
 &+ \frac{\sum b_{if}(\bar{X}_{im} - \bar{X}_{if})}{(3)} \quad (2)
 \end{aligned}$$

但若選擇另一團體為標準團體 (此例為女性)，則兩個團體依變項平均數粗差異可以分解為式 3 中的三部份。同樣的，此式中的第三部分為式 1 中第 (3) 項和第 (4) 項的結合。可見，以多元迴歸分析為基礎的調整法 (或標準化)，其分解的方式與研究者所選擇的標準組之不同而略有差別。但是由於式 1 中的第 (4) 項 (交互作用項) 常被視為殘差項 (residual)，且其數值常常都不大。因此，採用式 1 到式 3 的迴歸標準化或調整方法所得到的研究結論常會相近，不致於差別太大。本論文的實例討論將採用式 1，將團體依變項平均數粗差異分解為四個成份。

$$\begin{aligned}
 \bar{Lny}_m - \bar{Lny}_f &= \frac{(a_m - a_f)}{(1)} + \frac{\sum \bar{X}_{if}(b_{im} - b_{if})}{(2)} \\
 &+ \frac{\sum b_{im}(\bar{X}_{im} - \bar{X}_{if})}{(3) + (4)} \quad (3)
 \end{aligned}$$

在式 1 中四個成份隨研究主題的不同，在概念意義的詮釋及使用上也有些微不同。Semyonov (1988) 將式 1 中的第一部分及第二部分解釋為團體平均數差異中「無法被解釋」的差異部分，它分別由團體會員 (group membership) 和人力資本報酬 (returns on human resources) 所造成。並且把它們視為勞動力市場歧視 (market discrimination) 的差異

部分。因為不知道造成人力資本報酬差異及團體會員差異的確切原因，所以將它們視為無法被解釋的部分。而第三個部分是由於團體間人力資本或人力組成或才能 (composition or endowment) 差異所形成的。本文特別把這個部分稱之為團體結構性差異效果 (structural difference effect or component)，其主要的意義是表示這部分之差異是由於團體間自變項的分配差異所自然誘導出來的。結構性差異效果是粗差異中可以被解釋的部分，亦即它是由已知的原因 (自變項平均數差異) 所產生的差異。結構性差異的解釋會隨變項的概念意義或特性而有所不同。基本上，就一個自變項而言，它有可能是團體歧視所造成的，但卻是有別於前述的勞動力市場歧視。¹ 另外，也有可能只是抽樣過程所造成的結果。最後式 1 中的第四部分是組成或才能差異和報酬率差異的聯合交互作用所產生的差異；此部分可以視為殘差項，它的絕對數值通常是不大。Semyonov and Cohen (1990) 把第三和第四部分合稱為團體間「正當的或合法的差異」(legitimate gap)。本論文的實例討論將特別著重於第三部分，所謂結構性差異的討論。

在實證研究上，Rosenfeld 和 Kalleberg 以本文中式 1 分析美國、加拿大、挪威和瑞典四國的男女性平均所得差異並進行分解，結果顯示包括人力資本、工時、職業等九個自變項組的結構性差異效果分別為 28%、43%、35% 和 35% (Rosenfeld and Kalleberg, 1990: Table 6)。另外，Semyonov 探討以色列境內阿拉伯人與猶太人所組成的雙種族勞動力市場 (bi-ethnic labor market) 和單種族勞動力市場 (mono-ethnic labor market) 之社會經濟不平等，結果發現其結構性差異效果 (或組成資源差異) 從 13% 至 120% 不等 (Semyonov, 1988: Table 4)，視比較的團體不同而有所差異。再者 Semyonov 和 Cohen (1990) 的以色列 33 個地方勞動力市場的所得研究結果指出真正的市場歧視效果僅佔 40%；換言之，結構差異效果將近 60%。可見，在依變項為連續性變項的團體比較研究中，結構性差異效果如果沒有給予適當的處理的話，那麼研究結果不正確的可能性很高。

(二) 對數線性消除法

當研究者所進行的比較研究，其性質若屬於在前言中所說明的第三類的話，通常研究者必須比較各團體之現象率的差異。基本上，這裡所說的現象率就是現象粗率。而現象粗率是決定於兩個重要成份，一是團體的組成分配，另一是組成特殊率(參見文中式 6)。相似於迴歸標準化的理由，即使兩個團體有相同的組成特殊率，一旦團體具有不同的組成分配，則團體還是會有不同的現象粗率產生。因此，直接將現象粗率加以比較的時候，往往很容易得到不恰當的研究結論(Clogg and Eliason, 1988；謝雨生，1990, 1991)，尤其是當團體間的組成分配差異很大時。所以研究者常需要控制或調整擬比較的團體之組成分配，進而計算出調整率(adjusted rate)。這程序就是人口學上所謂的標準化(standardization)。這種「標準化」工作的基本目的就是擬排除團體組成差異所造成的差異，也可更進一步將現象粗率加以分解。如此，研究者不僅可以掌握團體間欲探討之現象的真正差異，同時也可以瞭解團體組成差異所造成的結構差異效果。一般來說，研究者最想問的問題當然是「當團體間的組成差異予以控制或消除後，依變項有無團體間的差異存在？」換句話說，在消除團體間的組成結構差異效果後，依變項的現象(或現象率)是否還有團體間的差異存在？同時，組成結構效果到底有多少？

這種目的之研究，其研究資料可說都是組成(composition)、團體(group)和依變項(dependent variable)三變項的 $I \times J \times K$ 之交叉表形式。這裡的組成、團體和依變項分別以C、G和D符號表示，且每一個變項都是類別變項或非連續變項。另外，I、J和K分別代表C、G和D變項的類別數目或層次數目(the number of levels)。組成和團體變項是可以調換使用的，亦即組成變項和團體變項可以根據研究目的而加以調換使用，但不論是組成變項或團體變項都可以是其他變項的聯合變項。譬如，組成變項可以是年齡和性別的聯合變項。因此，本論文底下所討論的方法之使用將不限於三個變項的分析，它可以用於

多變項的分析。同時每一個變項的層次也不限於二個層次，可以是多個層次。

由於這種研究所分析的資料都是屬於組成、團體和依變項之類別變項交叉而得的次數分配表。在傳統上，標準化(包括直接標準化或間接標準化)的調整方式，在文獻上已有相當豐富的討論，尤其是在人口學上，標準化方法有相當長的發展歷史(Kitagawa, 1955, 1964; Keyfitz, 1968; Shryock and Siegel, 1980; Spiegelman, 1968; Wunsch and Termote, 1978)。Clogg and Eliason (1988) 更指出了直接標準化和對數線性消除法的異同之處，也指出對數線性消除法的優點。本文的目的之一也就是要說明對數線性消除法的理論基礎及使用情形。

讓 f_{ijk} 表示三個變項，組成(C)、團體(G)和依變項(D)，交叉表中(i, j, k)空格的觀察次數，而 F_{ijk} 為某一對數線性模式下相對應空格的期望次數。組成特殊率則定義為：

$$r_{ij(k)} = F_{ijk} / F_{ij+} \quad (4)$$

其中「+」表示根據該下標之累加。 $F_{ij+} = N_{ij}$ 表示第 j 個團體的第 i 類組成的總次數。同時 $r_{ij(k)}$ 為第 j 個團體，第 i 個組成中的第 k 類依變項的發生率 (the rate of prevalence)。因此，第 j 個團體的第 k 類依變項之粗率則為：

$$r_{j(k)} = F_{+jk} / F_{+j+} \quad (5)$$

F_{+j+} 是第 j 群的所有觀察次數。也可以說粗率是由團體(G)和依變項(D)兩變項交叉表的邊際分配導得。

若使用式 4，則式 5 可以改寫為：

$$r_{j(k)} = \sum C_{ij} r_{ij(k)} \quad (6)$$

$$C_{ij} = F_{ij+} / F_{tj+} \quad (7)$$

C_{ij} 為第 j 團體中第 i 類組成所佔的比例。因此每個團體的 $\sum C_{ij} = 1$ 。我們已經可以發現式 6 中現象粗率是組成特殊率 ($r_{ij(k)}$) 的加權平均數，其權重為 C_{ij} 。換言之，現象粗率是由組成分配和組成特殊率所共同決定的。傳統上直接標準化是經由選擇一組所謂的標準組成權重 (C_i^s)，以控制團體間的組成差異。如此，調整率就可以定義為：

$$r_{j(k)}^s = \sum C_i^s r_{ij(k)} \quad (8)$$

很清楚地，在式 8 中團體間的組成差異已經被加以控制，因為每一個團體皆已經使用了相同的組成權重。但是標準率 (或調整率) 又與 C_i^s 的選擇有密切的關係，因而，文獻上常會提到選擇適當的組成權重問題 (Clogg and Eliason, 1988; Keyfitz, 1968; Freeman and Holford, 1980)。Clogg and Eliason (1988) 更指出假如存在組成 \times 團體 \times 依變項 (C-G-D) 三因子交互作用時，則上述的一組標準組成權重的選擇不見得適當；而對數線性消除法卻能克服這種問題。

如以乘法模式 (multiplicative model) 來表示組成、團體和依變項三個變項交叉表所呈現的關係，則可以表示如下：

$$F_{ijk}^{CGD} = \tau \tau_i^C \tau_j^G \tau_k^D \tau_{ij}^{CG} \tau_{ik}^{CD} \tau_{jk}^{CD} \tau_{ijk}^{CGD} \quad (9)$$

三因子交互作用是由一組 IJK 個 τ_{ijk}^{CGD} 的母數來代表。當有一些空格之 $\tau_{ijk}^{CGD} \neq 1$ 時則表示團體與依變項間的關係會因組成變項的層次不

同而不同。當所有 τ_{ijk}^{CGD} 值為 1 時，則式 9 中三個兩因子的母數則分別測量了當第三個變項被加以控制後的兩個變項間的淨關係 (partial association)。因此， τ^{CG} 測量了組成與團體的淨關係， τ^{GD} 測量了團體與依變項的淨關係，是所謂淨團體效果 (partial group effect)；而 τ^{CD} 則是測量了組成與依變項的淨關係，亦即為淨組成效果 (partial composition effect)。

如把式 9 代入式 5，就可以看出粗率決定於淨組成效果、淨團體效果和淨組成一團體交互作用效果。但是母數 τ^{CG} 和 τ^{CGD} 會干擾團體粗率的直接比較，這也是需要考慮使用對數線性消除法的基本理由所在。

以對數線性消除法來進行標準化，在對數線性模式上的意義即是同時消除了 τ_{ij}^{CG} 和 τ_{ijk}^{CGD} 的影響效果，再利用去除干擾效果後的次數 (F_{ijk}^*) 來計算出調整率。這兩個步驟可以分別由式 10 和式 11 來加以表示：

$$F_{ijk}^* = F_{ijk} / (\tau_{ij}^{CG} \tau_{ijk}^{CGD}) \tag{10}$$

$$r_{j(k)}^* = F_{+jk}^* / F_{+j+}^* \tag{11}$$

一般而言，研究者常常會選擇一個特定的團體，以作為其他團體的比較對象。假如第 s 個團體被選為參考團體或所謂的標準組，則其意義為：

1. 第 s 個團體的粗率和調整率是完全相同的。
2. 干擾之交互作用項將以第 s 團體者為參考團體地被加以控制。

3. 標準團體以外的每一個團體之調整率皆將與第 s 團體加以比較。

上述的有標準團體之調整後次數如式 12 所列。

$$F_{ijk}^{**} = F_{ijk} (\tau_{is}^{CG} \tau_{isk}^{CGD}) / (\tau_{ij}^{CG} \tau_{ijk}^{CGD}) \quad (12)$$

其中 s 代表標準團體之代號。同樣的，有標準團體的各團體調整率 $r_{j(k)}^{**}$ 就可以用 F_{ijk}^{**} 去計算出來。再將式 12 所運算出來的調整率 ($r_{j(k)}^{**}$) 和標準團體調整率 (也就是粗率) 分別加以比較，不僅可以瞭解團體間的真正差異，而研究者也可以掌握團體間組成差異所造成的效果，亦即結構差異的效果。

不過，如果組成變項是一個聯合變項，則上面所談的結構差異效果是一種總體的效果；換言之，它是聯合變項的總體效應。假如研究者需要進一步瞭解聯合組成變項中每一個組成變項的單獨效果及其交互作用的效果，則需要再進行結構差異效果的分解工作。Liao (1989) 所提出的現象率差異分解方法則適合這個工作需要。他設立了一組線性聯立方程來說明粗率差異分解的情形，如式 13 至式 16 所示。

$$RE + AE + BE + ABE = CRD \quad (13)$$

$$RE + \quad \quad BE + ABE = RD (A) \quad (14)$$

$$RE + AE + \quad \quad ABE = RD (B) \quad (15)$$

$$RE \quad \quad \quad = RD (AB) \quad (16)$$

其中 RE 代表真正的差異，也就是率效果 (rate effect)，它是由前述 $r_{j(k)}^{**}$ 與標準團體之現象率的差異而來的。AE 代表聯合組成變項 (如

年齡一性別) 中第一個組成變項 (如年齡) 的團體結構差異效果, BE 表示第二個組成變項 (如性別) 的團體結構差異效果, 而 ABE 則代表兩個組成變項之團體結構差異的交互作用效果。而 AE、BE 和 ABE 的總和就是團體結構差異總效果。換言之, 如果聯合組成變項是兩個變項交叉結合而成, 則其團體結構差異總效果可以分解為上述的三個部分。另外, CRD (crude rate difference) 為團體間粗率差異, RD (A)、RD (B) 和 RD (AB) 則分別表示消除第一個組成變項的團體結構差異效果後之調整率的差異 (與標準團體之差異)、消除第二個組成變項的團體結構差異效果後之調整率的差異及消除聯合組成變項的團體結構差異效果後之調整率的差異。因此, 只要 RD (A)、RD (B) 和 RD (AB) 從對數線性消除法中取得, 則 RE、AE、BE 和 ABE 就可以從解式 13 至式 16 的聯立方程式求得。

至於對數線性消除法與傳統人口學所用的標準化方法的差異所在請詳細參閱 Liao (1989) 的研究。而這部分的經驗研究上, Sweet (1984) 曾研究指出 1970 年至 1980 年的十年間美國所增加的家戶數量, 有三分之二是由於人口結構改變所造成的。另外, 同期間美國犯罪率的增加有 43% 是源於人口結構改變的結果 (Weeks, 1989: 242)。Clogg 和 Shockey (1985) 以對數線性消除法探討美國 1969 年至 1980 年的勞力未適當運用現象, 發現不論是就絕對量或相對量來看, 美國的年齡和性別組成變遷並沒有對美國勞力未適當運用產生影響。Liao (1989) 比較美國 1970 年和 1985 年的死亡率差異時, 粗差異為 $-0.683(\%)$, 但在消除年齡與種族組成的結構差異效果後, 死亡率的真正差異為 $-3.464(\%)$ (Liao, 1989: Table 2)。另外, 臺灣的經驗研究也發現 1980 年至 1988 年的勞動力市場之未適當運用率的粗差異為 6.51 個百分點, 但是排除年齡與性別組成的結構差異效果後, 人力未適當運用率的真正差異為 12.32 個百分點 (謝雨生, 1990: 表 6)。從以上的經驗研究之結果來看, 可以發現在群體比較或跨時間的變遷研究中, 研究者應該謹

慎地處理群體結構差異或變遷的效應，否則很容易產生偏差的研究結論。

以上的討論，我們可以瞭解不論使用迴歸標準化(調整法)或使用對數線性模式的對數線性消除法，研究者皆可以掌握到團體比較研究中必須特別處理的團體結構差異的部分。如此，研究者可以明確地說明團體間的真正差異所在。本論文即以實例說明上述兩種比較研究中團體結構差異的考慮和處理方法之實際使用情形與結果，以為參考。

參、實例的資料來源與分析

在不平等的經驗研究中，大部分的作者發現女性的平均薪資總是遠低於男性的平均薪資 (Jones, 1983; Kalleberg, 1988; Kalleberg and Lincoln, 1988; Rosenfeld and Kalleberg, 1990; Semyonov, 1988; Semyonov and Cohen, 1990; Kelley and Evans, 1993)。不僅是已開發的國家如此，新興工業化國家也是如此，臺灣自然也不例外。雖然這種男女性平均薪資差距在這些年有些微縮小，但是性別差異仍然相當大。以臺灣為例，1989年時，女性平均薪資是男性平均薪資的63.9%；至1993年時這個比例略為提高至66.0%。承如前面的討論所述，這種男女性平均薪資的差異是粗差異，並不能全部視為真正的差異。換言之，在分析男女性平均薪資差異時，必須有效地去除前述的結構性差異後，才可能對真正差異有一正確的瞭解。不論是經濟學領域，或是社會學領域皆有不少理論和經驗研究探究男女性所得差異的原因。本論文的焦點並不是所得性別差異理論的討論，²而是以此為例說明粗差異的分解並說明結構性差異的處理情形，尤其強調比較團體間結構差異所造成的差異部分之討論。

為完成這部分迴歸標準化或迴歸調整法的實例討論，本部分的實證資料取自中央研究院民族學研究所執行完成之「臺灣地區社會變遷

表 1：變項測量說明

變項名稱	說 明
月工作所得	現在工作的個人月收入，單位：元
Ln(月工作所得)	現在工作的個人月收入之自然對數值
年齡	以81減去出生年
婚姻地位	虛擬變項，未婚者為0，其它為1
教育年數	教育程度轉換成教育年數
場所規模	現在工作之公司或機關所有員工數，以下列兩個 虛擬變項表示：
11—50人	虛擬變項，場所規模11—50人者為1，其它為0
51人以上	虛擬變項，場所規模51人以上者為1，其它為0
行業	以下列兩個虛擬變項表示：
製造業	虛擬變項，製造業為1，其它行業為0
服務業	虛擬變項，服務業為1，其它行業為0
職業聲望	引自 Tsai and Chiu (1991) 臺灣職業聲望建構

基本調查」第二期第三次調查的原始資料。分析資料同時過程自問卷 I 和問卷 II 的適用對象。分析對象包括563個男性及489個女性，他們都是在「私人事業」單位(但非農林漁牧、礦業者)的有酬工作者。實例中的各變項之定義與測量方式如表 1 所列。自變項包括了人力資本變項(年齡、婚姻地位和教育年數)和工作特性(場所規模、行業和職業聲望)。資料分析程序約略如下：

- (一) 分別取得影響男性及女性工作所得的多元迴歸方程式和各變項之平均數。
- (二) 以式 1 為基礎，進行平均數粗差異的分解工作，而將粗差異分四個成份。

(三) 確認團體結構差異效果之成份，並進行必要的解讀工作。

另外，本論文的第二個實例的主題是想探討臺灣地區人力運用現象的變遷情形，其分析資料取自行政院主計處和行政院經建會所發表之「臺灣地區人力運用調查報告」，時間從1980年至1993年，共計14年時間。分析資料是由組成變項、時間(當作團體)和勞力運用狀況(當作依變項)所形成的三向度交叉表。其中組成變項是由年齡和性別組合的聯合變項。男、女性的勞動力年齡組分15—19、20—24、25—29、…、65⁺等十一組，所以年齡—性別聯合組成則有二十二類，以C表示之；所比較的時間(團體)由1980年至1993年，共計14年，亦即有十四個團體，以G表示；而依變項是勞力運用狀況，共包括勞力適當運用、工作時數不足、所得偏低、教育與職業不相稱和失業等五類，以D表示之。

資料分析時，先以對數線性消除法(如式10至式12所列)去除勞動力聯合組成變項(年齡—性別)和時間(團體)及聯合組成變項×時間×勞力運用三因素之交互作用，並選擇1980年為標準團體，進而取得式13至式16所列的RD(A)、RD(B)、RD(AB)及CRD。再依Liao(1989)所提的粗率差異分解法，將各年的各類勞力運用率與1980年之粗率差異分為四個成份，即RE、AE、BE和ABE。

由於資料處理複雜，所以再詳細說明如後。首先以 $22 \times 14 \times 5$ 的聯合組成×時間×勞力運用之交叉表，以飽合對數線性模型的處理(如式9)，選擇1980年的聯合組成為標準組成，再經式9至式12的步驟調整，可以求得各年消除聯合組成之變遷(即團體差異)後的勞力運用調整率。將每一年之調整率與1980年之調整率(等於其粗率)比較，即可求得RD(AB)。再者，RD(A)和RD(B)的獲得比較複雜，所以也詳細說明如後：

- (一) 分別將男性的 $11 \times 14 \times 5$ 和女性的 $11 \times 14 \times 5$ 的年齡組成 \times 時間 \times 勞力運用之交叉表，也以式 9 至式 12 的過程，分別求得各年男性或女性勞力運用的調整率，及 $r_{j(k)}^{**}$ 。
- (二) 再以各年男女性勞力之比例為權重，將 (一) 中的男女性調整率乘上權重，而求得各年在消除年齡組成變遷影響效果後的調整率 ($r_{j(k)}^{***}$)。如式 17 所示：

$$r_{j(k)}^{***} = w_m \times {}_m r_{j(k)}^{**} + w_f \times {}_f r_{j(k)}^{**} \quad (17)$$

其中 w 為男女性之勞動力的比例， m 和 f 分別表示男女性。

- (三) 最後將各年的 $r_{j(k)}^{***}$ 與 1980 年標準團體之調整率 (等於其粗率) 比較，即可獲得 RD (A)。

同樣地，RD (B) 也可以上述相同的程序而求得。只是所使用的資料改為 11 個年齡組的 $2 \times 14 \times 5$ 的性別組成 \times 時間 \times 勞力運用之交叉表，而且其權重為各年勞動力之年齡組成的比例。

一旦 RD (AB)、RD (A)、RD (B) 及 CRD 皆求得之後，就可以利用式 13 至式 16 的聯立方程組解得 AE、BE、ABE 和 RE。至此，團體比較之粗率差異的分解才算完成。研究者就可以據以討論團體結構差異效果和現象真正的團體差異。底下將呈現和討論實例運作後的結果。

肆、團體結構差異效果：實例說明

在第一個實例中，我們想要探討臺灣男女性平均工資的差異，同時瞭解團體結構差異(包括人力資本及工作特性)所造成的差異之效果。在分析樣本中，1992年男性的平均月工作所得為31,412元，女性則為18,599元。女性的月平均工作所得僅為男性者的59%，絕對數相差12,813元(見表2)。可見，男女性的平均月工作所得仍有相當大的差異。

我們假定工作者之年齡、婚姻地位和教育年數等人力資本變項及場所規模、行業和職業聲望等工作特性變項可用來解釋工作所得。但是從表2中，我們可以發現樣本中男女性的平均年齡、平均教育年數、工作於服務業的比例及平均職業聲望分數皆有顯著的性別差異。所以男女性的平均月工作所得粗差異(12,813元)將會有一部分是由於年齡、教育、工作於服務業及職業聲望等四個自變項的性別結構差異所造成的團體結構差異效果，而不能將粗差異全部歸因於勞動力市場上的性別歧視。

爲了瞭解這裡團體(性別)結構差異效果有多少，我們進行了粗差異的分解工作。利用表3所列的男、女性的月工作所得(取log)之多元迴歸方程式和表2所列的各自變項的平均數及前述的式1進行粗差異的實際分解，其結果如表4所列。由表4，我們可以發現，男女性的平均月工作所得差異中有13%是性別結構差異效果，而87%是屬於勞動力市場之性別歧視所造成的。值得注意的是由於這個說明實例並沒有周全地處理所有可能解釋所得的影響因素(因爲受到現成資料的限制)，因此，上述的研究結論都是暫時性的。以這個實例來說，研究者在進行男女性月工作所得差異之討論時，起碼應該先消除或考慮這13%的性別結構差異效果，再進行其他的相關討論。我們真正的企圖是藉此實例，提出團體結構差異效果在團體比較研究中考慮之重要性，並提供實際的處理過程。

表 2：平均月工作所得與各影響變項平均數 (標準差) 及其差異檢定

變項名稱	平均數		差異檢定 t 值
	男 性 (N=563)	女 性 (N=489)	
月工作所得	31412.0782 (16722.159)	18599.1820 (10832.692)	12812.896 ^{***}
Ln(月工作所得)	10.2498 (.447)	9.7125 (.489)	.537 ^{***}
年齡	35.7052 (10.159)	32.9714 (9.590)	2.734 ^{***}
婚姻地位 (未婚 = 0)	.6696 (.471)	.6196 (.486)	.050
教育年數	11.1030 (3.748)	10.0511 (4.656)	1.052 ^{***}
場所規模 (10人以下 = 0)			
11 — 50人	.3002 (.459)	.3252 (.469)	-.025
51人以上	.4227 (.494)	.3967 (.490)	.026
行業 (營造業 = 0)			
製造業	.5453 (.498)	.5153 (.500)	.030
服務業	.3073 (.462)	.4397 (.497)	-.132 ^{***}
職業聲望	54.6191 (8.905)	53.2817 (7.300)	1.337 ^{**}

** P < .01, *** P < .001。

表 3：影響男女性工作所得之迴歸分析

變 項	男 性		女 性	
	b	s.e.	b	s.e.
年齡	-.000228	.0002	-.003380	.0029
婚姻地位 (未婚 = 0)	.176732***	.0411	.134283**	.0496
教育年數	.031845***	.0062	.029567***	.0069
場所規模 (10人以下 = 0)				
11 - 50人	.025466	.0438	.094425	.0502
51人以上	.058163	.0443	.194253***	.0495
行業 (營造業 = 0)				
製造業	-.203469***	.0505	-.329147***	.0947
服務業	-.093795	.0548	-.230472*	.0952
職業聲望	.012848***	.0022	.012997***	.0036
常數項	9.191842***	.1385	8.914274***	.2220
R ²	.2562		.2697	
N	563		489	

* P < .05, ** P < .01, *** P < .001。

表 4：男女性平均月工作所得差異之分解結果

變 項	$\bar{X}_f(B_m - B_f)$	$B_f(\bar{X}_m - \bar{X}_f)$	$(B_m - B_f)(\bar{X}_m - \bar{X}_f)$
年齡	.10393	-.00924	.00862
婚姻地位 (未婚 = 0)	.02630	.00671	.00212
教育年數	.02290	.03110	.00240
場所規模 (10人以下 = 0)			
11 - 50人	-.02243	-.00236	.00172
51人以上	-.05399	.00505	-.00354
行業 (營造業 = 0)			
製造業	.06476	-.00987	.00377
服務業	.06010	.03051	-.01810
職業聲望	-.00794	.01738	-.00020
Σ	.19363 -- (2) (36.0)*	.06929 -- (3) (12.9)	-.00320 -- (4) (-0.6)
常數項差異	.27757 -- (1) (51.7)		
Ln(月工作所得) 差異	.53730 (100)		

* 括號內數字是以 Ln(月工作所得) 差異值為100之比例值。

表 5：臺灣地區勞動力人力運用情形，1980~1993年

年度	適當運用 (1)	未適當運用 (2)= (3)+(4)+(5)+(6)	工作時數不足 (3)	低所得 (4)	教育與職業 不相稱(5)	失業 (6)
1980	84.98	15.02	1.10	7.54	5.18	1.19
1981	86.10	13.90	1.52	5.78	5.58	1.01
1982	82.97	17.03	2.68	6.36	6.02	1.98
1983	76.65	23.35	3.68	10.82	6.43	2.42
1984	77.68	22.32	3.23	9.74	7.23	2.11
1985	75.34	24.66	2.69	12.20	7.20	2.57
1986	75.57	24.43	3.78	10.84	7.47	2.34
1987	77.47	22.53	2.02	10.18	8.38	1.94
1988	78.41	21.59	1.71	8.82	9.32	1.74
1989	77.26	22.74	1.63	10.56	9.04	1.50
1990	79.15	20.85	1.61	9.05	8.71	1.48
1991	79.50	20.50	1.35	8.63	9.08	1.43
1992	78.59	21.41	1.82	8.97	9.22	1.40
1993	82.53	17.47	1.55	9.10	5.54	1.29

資料來源：臺灣地區人力運用調查報告，1980年至1993年。

在第二部分的實例中，我們說明了對數線性消除法用於掌握比較研究中團體結構差異效果的實際使用情形。表 5 列出了從 1980 年至 1993 年的臺灣地區人力運用狀況。常常有人直接使用表中所列的現象率加以比較並下結論。譬如，有人說，1980 年的人力未適當運用率為 15.02%，到 1992 年增加為 21.42%，成長了 43%，並藉以指出臺灣勞動力就業市場環境的種種不利變遷。基本上，這樣的資料使用及其說明可能會是不正確的。誠如前面已經討論的，任何現象粗率都受到組

成特殊率和組成分配(或結構)兩因素的影響。尤其是跨團體(或跨時間)比較時,團體間(或不同時間點)的組成分配差異很大(或已經發生很大的變遷)時,直接使用現象粗率來討論現象的變遷都是不恰當的作法。

表 6：臺灣地區勞動力年齡－性別組成及變化，1980年及1993年

年齡	男 性			女 性		
	1980年	1993年	差異	1980年	1993年	差異
15—19	8.27	3.02	-5.25	17.53	5.07	-12.46
20—24	9.78	7.69	-2.09	23.86	16.31	-7.55
25—29	18.37	16.03	-2.34	15.41	16.84	1.43
30—34	11.99	17.09	5.10	9.58	15.66	6.08
35—39	11.23	15.79	4.56	9.08	15.06	5.98
40—44	9.90	12.96	3.06	8.59	12.41	3.82
45—49	9.44	7.93	-1.51	7.28	6.74	-0.54
50—54	9.53	7.30	-2.23	4.34	5.46	1.12
55—59	6.84	5.80	-1.04	2.94	3.64	0.70
60—64	3.45	4.39	0.94	1.13	1.97	0.84
65 ⁺	1.24	2.00	0.76	0.27	0.84	0.57
△ [*]			14.44			20.55
合計(千人)	4354	5463		2213	3353	

資料來源：臺灣地區人力運用調查報告，1980年及1993年。

* △為差異指數(Index of Dissimilarity)。 $\Delta = (1/2) \sum |p_i - q_i|$ ， p_i 及 q_i 分別為1980年及1993年各年齡層之比例。

表 7：臺灣地區勞動力運用類型年齡—性別分佈特殊率，1980年及1993年

年齡組	男 性					女 性				
	工作時數 不 足	所得偏低	教育與職 業不相稱	失業	適當運用	工作時數 不 足	所得偏低	教育與職 業不相稱	失業	適當運用
1980 年										
15-19	0.64%	21.15%	3.51%	2.52%	72.18%	0.25%	9.79%	3.19%	2.71%	84.06%
20-24	0.86	3.73	11.52	3.30	80.59	0.29	4.58	7.91	2.65	84.56
25-29	0.86	2.56	10.99	1.33	84.27	1.39	6.51	9.30	1.77	81.02
30-34	0.81	1.63	6.49	0.62	90.45	1.09	9.56	3.26	0.35	85.74
35-39	1.56	2.79	3.12	0.29	92.24	0.87	10.26	0.94	0.13	87.80
40-44	1.42	5.42	2.40	0.19	90.58	1.74	11.29	0.42	-	86.56
45-49	1.44	7.34	1.41	0.58	89.23	2.34	11.07	0.29	-	86.30
50-54	1.50	8.19	2.97	0.75	86.59	1.43	12.05	0.52	-	85.99
55-59	1.25	11.89	4.13	0.33	82.41	1.02	16.85	0.34	-	81.79
60-64	2.42	20.25	2.17	0.75	74.41	2.65	20.22	2.32	-	74.82
65 ⁺	2.08	25.26	0.68	-	71.98	-	32.16	-	-	67.84
全部	1.18	6.92	5.59	1.07	85.24	0.95	8.77	4.39	1.42	84.46
1993 年										
15-19	0.99	30.94	5.45	3.17	59.45	0.20	27.15	10.30	4.17	58.18
20-24	0.74	4.51	9.12	4.84	80.80	0.58	9.07	8.63	3.83	77.89
25-29	0.68	2.38	8.94	1.95	86.05	0.62	6.06	7.47	1.41	84.44
30-34	1.39	2.31	7.75	0.76	87.79	1.37	9.23	6.34	0.85	82.21
35-39	1.47	3.35	5.98	0.98	88.22	2.01	10.37	4.52	0.25	82.86
40-44	1.55	3.78	4.67	0.53	89.46	1.79	11.78	3.18	0.49	82.75
45-49	2.19	5.76	3.09	0.60	88.36	3.23	13.49	1.32	-	81.95
50-54	2.37	10.34	1.49	0.60	85.20	3.55	15.40	0.24	-	80.81
55-59	2.81	19.61	0.99	0.56	76.02	3.71	22.39	-	-	73.90
60-64	2.56	25.59	1.16	0.46	70.22	2.61	25.44	0.26	-	71.70
65 ⁺	2.25	47.01	0.64	0.21	49.90	3.43	35.85	-	-	60.71
全部	1.53	7.49	5.64	1.28	84.04	1.58	11.71	5.36	1.30	80.05

資料來源：台灣地區人力運用調查報告，1980年及1993年。

另外，我們從表 6 資料發現，男女性的年齡組成從 1980 年至 1993 年間已經產生了相當大的變化，亦即是比較團體間其組成分配差異很大。再者，年齡別人力運用狀況也有差異。換言之，組成特殊率也存在團體差異（見表 7）。這兩種狀況綜合顯示「直接」使用人力運用各種現象率之團體（及時間）比較，極容易產生錯誤結論。因此，對數線性消除法的使用不僅可以去除團體結構差異的干擾，同時可以確認團體結構的差異效果。

經過對數線性消除法的使用，去除了勞動力組成變遷對人力運用現象粗率的干擾效果後，調整率列於表 8 第 3 欄。從調整率和粗率的比較或其指數的比較，我們可以很清楚地看到兩者間之差異，亦即組成變遷效果或結構差異效果相當明顯。同時，我們更可發現，如果沒有組成分配的變遷或差異，則 1992 年的人力未適當運用率應該較 1980 年者成長 74%，而不是使用粗率比較時的 43%。一方面顯示人力未適當運用的實際狀況要比粗率表面上的變化嚴重得多；另一方面也顯示勞動力組成變遷的結果對臺灣人力適當運用是有利的發展。這個說明實例的結構差異效果則列於表 8 的第 5 和第 6 欄。這些數據在在顯示團體研究中，團體結構差異考慮的重要性。

爲了進一步說明粗率差異的分解情形，我們進行了人力未適當運用及其各細類的粗率差異分解，以瞭解各人力運用現象的真正差異、年齡結構差異效果、性別結構差異效果，及後兩者的交互作用效果。分析結果如表 9 所示。從表 9 中，我們可以將粗差異 (CRD) 分解爲真正差異效果 (RE) 和結構差異效果兩大部分；同時，結構差異效果又分爲年齡結構差異效果 (AE)、性別結構差異效果 (BE) 和後兩者之交互作用 (ABE)。可見，在消除團體結構差異效果後，人力未適當運用的真正差異才得以呈現出來。另外，在結構差異效果中，年齡結構差異和性別結構差異對不同類型的人力未適當運用（如所得偏低和失業）的影響也不同。這顯示藉由粗率差異分解的工作，研究者可以更細緻和有效地掌握聯合組成變項中每一個個別組成變項的差異影響量或其交互

作用的影響量，有助於研究者釐清現象本身和團體結構差異的本質。

表 8：臺灣地區勞力未適當運用的結構差異效果

年度	粗率	指數 ^a	調整數	指數 ^b	結構差異效果	
					(5)=(1)-(3)	(6) ^c =(2)-(4)
1980	15.02	100	15.02	100	—	—
1981	13.90	93	13.58	90	0.32	3
1982	17.03	113	19.87	132	-2.84	-19
1983	23.35	155	25.92	173	-2.57	-18
1984	22.32	149	21.78	145	0.54	4
1985	24.66	164	25.89	172	-1.23	-8
1986	24.43	163	26.98	180	-2.55	-17
1987	22.53	150	27.45	183	-4.92	-33
1988	21.59	144	27.22	181	-5.63	-37
1989	22.74	151	25.38	169	-2.64	-10
1990	20.85	139	24.10	160	-3.25	-21
1991	20.50	136	22.95	153	-2.45	-17
1992	21.41	143	26.06	174	-4.65	-31
1993	17.47	116	18.34	122	-0.87	-6

a 以1980年的粗率為100。

b 以1980年的調整率為100。

c 為指數差異，亦即結構差異效果以粗率和調整率之指數差異來表示。

表 9：臺灣地區勞力未適當運用之結構差異效果及其分解效果

比較的年份	粗差異 (CRD)	結構差異效果			
		真正 差異效果 (RE)	年齡結構 差異效果 (AE)	性別結構 差異效果 (BE)	年齡與性別結構 差異交互效果 (ABE)
勞力未適當運用					
1981-1980	-1.13	-1.44	-0.66	-1.89	2.87
1982-1980	2.01	4.84	-3.03	-0.77	0.97
1983-1980	8.33	10.89	-1.79	-1.93	1.16
1984-1980	7.29	6.76	0.26	-1.53	1.80
1985-1980	9.64	10.87	-1.70	-2.87	3.34
1986-1980	9.41	11.96	-2.08	-1.27	0.79
1987-1980	7.51	12.43	-4.07	-2.35	1.49
1988-1980	6.57	12.19	-4.91	-2.73	2.03
1989-1980	7.72	10.36	-2.24	-2.06	1.66
1990-1980	5.82	9.08	-4.17	-3.17	4.09
1991-1980	5.48	7.92	-3.22	-2.71	3.48
1992-1980	6.39	11.04	-3.97	-2.71	2.03
1993-1980	2.45	3.32	-1.11	-1.50	1.74
工作時數不足					
1981-1980	0.41	0.32	-0.01	0.14	-0.03
1982-1980	1.57	1.21	0.19	0.22	-0.05
1983-1980	2.57	3.35	-0.50	-0.17	-0.10
1984-1980	2.13	1.63	0.29	0.21	0.00
1985-1980	1.58	2.21	-0.37	-0.29	0.03
1986-1980	2.67	3.30	-0.41	-0.17	-0.05
1987-1980	0.92	1.23	-0.19	-0.26	0.13
1988-1980	0.61	1.01	-0.24	-0.32	0.16
1989-1980	0.53	0.90	-0.24	-0.31	0.19
1990-1980	0.51	0.62	-0.02	-0.44	0.35
1991-1980	0.25	0.60	-0.21	-0.29	0.14
1992-1980	0.71	1.08	-0.18	-0.52	0.34
1993-1980	0.45	0.85	-0.27	-0.45	0.32

表 9：(續)

比較的年份	粗差異 (CRD)	真正 差異效果 (RE)	結構差異效果		
			年齡結構 差異效果 (AE)	性別結構 差異效果 (BE)	年齡與性別結構 差異交互效果 (ABE)
所得偏低					
1981-1980	-1.76	-1.92	-0.13	0.33	-0.05
1982-1980	-1.18	-2.04	-0.09	0.72	0.23
1983-1980	3.28	3.13	0.30	0.23	-0.38
1984-1980	2.20	2.71	-0.56	0.18	-0.13
1985-1980	4.66	4.69	0.04	0.80	-0.88
1986-1980	3.30	3.26	-0.01	0.66	-0.61
1987-1980	2.64	2.37	0.22	0.85	-0.80
1988-1980	1.28	1.07	-0.05	0.65	-0.39
1989-1980	3.01	3.16	-0.02	0.85	-0.97
1990-1980	1.50	1.59	-0.11	0.78	-0.76
1991-1980	1.09	1.29	-0.28	0.74	-0.66
1992-1980	1.42	1.66	-0.43	0.73	-0.54
1993-1980	1.55	2.17	-0.65	0.72	-0.69
教育與職業不相稱					
1981-1980	0.40	-0.57	0.46	-0.28	0.80
1982-1980	0.84	3.77	-2.05	-0.61	-0.27
1983-1980	1.25	1.17	0.51	-0.22	-0.22
1984-1980	2.05	0.36	1.73	-0.22	0.17
1985-1980	2.02	0.00	1.54	-0.32	0.79
1986-1980	2.29	1.81	0.84	-0.60	0.24
1987-1980	3.20	4.75	-0.76	-0.84	0.05
1988-1980	4.14	6.52	-1.42	-0.91	-0.05
1989-1980	3.86	4.47	0.00	-1.06	0.45
1990-1980	3.53	3.62	0.46	-1.01	0.46
1991-1980	3.90	1.70	1.93	-1.23	1.51
1992-1980	4.04	6.82	-1.73	-1.48	0.43
1993-1980	0.36	-0.88	0.97	-0.78	1.04

表 9：(續)

比較的年份	粗差異 (CRD)	真正 差異效果 (RE)	結構差異效果		
			年齡結構 差異效果 (AE)	性別結構 差異效果 (BE)	年齡與性別結構 差異交互效果 (ABE)
失業					
1981-1980	-0.18	0.73	-0.98	-2.08	2.15
1982-1980	0.78	1.90	-1.08	-1.10	1.06
1983-1980	1.23	3.24	-2.09	-1.78	1.86
1984-1980	0.92	2.05	-1.20	-1.68	1.76
1985-1980	1.38	3.96	-2.91	-3.07	3.40
1986-1980	1.14	3.59	-2.50	-1.16	1.21
1987-1980	0.75	4.07	-3.34	-2.11	2.12
1988-1980	0.55	3.59	-3.20	-2.15	2.31
1989-1980	0.31	1.84	-1.98	-1.54	1.98
1990-1980	0.28	3.24	-4.49	-2.50	4.04
1991-1980	0.24	4.33	-4.66	-1.93	2.50
1992-1980	0.21	1.48	-1.62	-1.44	1.80
1993-1980	0.10	1.17	-1.16	-0.99	1.07

伍、討 論

本文所談論的迴歸標準化(或迴歸調整法)及對數線性消除法是團體比較研究中相當一般性的處理程序。它們可以分別應用於許多不同主題的團體比較研究中，但是值得強調的是這並不代表這兩種方法可以完全取代其他常規性的分析方法。此外，迴歸標準化使用之前，主題性的理論邏輯發展，研究者更需講究，以使得迴歸標準化之結果解釋更具理論概念意義。同時，對數線性消除法更需與原來的對數線性模式的分析配合使用。可以說，對數線性消除法把原來研究者所選擇

的最佳、最精簡模式中與結構差異無關的母數轉換到一種比較容易詮釋的調整現象率上，來加以比較。當然，研究者並不是非用飽合模式不可。因為最佳、最精簡的模式並不一定是飽合模式。基本上，如果研究者所選擇的最佳、最精簡的模式中有牽涉到「組成變遷」的母數存在時，研究者就得針對此結構性差異效果加以有效的處理。

雖然在迴歸標準化的實例說明上，我們只討論了兩個團體的比較，但是這個方法是可以應用到多團體的比較研究。此外，對數線性消除法也可以推演至多變項的分析上。假如依變項是 y ，且有 X_1, X_2, \dots, X_q 為其解釋變項 ($q \geq 2$)。所有變項都是非連續變項。若研究者想探討當 $X_2 \sim X_q$ 受到控制時， X_1 對 y 的影響，則研究者可以將 X_1 當作團體變項 (G)，將 $X_2 \sim X_q$ 的交叉組合處理成聯合組成變項 (C)。同樣的，這種程序可以每次輪流提出一個解釋變項當作團體變項，其餘的解釋變項再交叉成一個聯合組成變項。當然，這樣的處理有一個很重要的基本要求，就是樣本大小 (sample size) 必須夠大。不論是迴歸標準化或對數線性消除法皆適用於多團體比較，也適用於多變項模式分析。前者更可應用在邏輯迴歸分析或多類別邏輯迴歸分析 (ordered or multinomial logit regression) 時。後者也可應用到多類別邏輯迴歸分析裡，只是當自變項若為連續性變項時，研究者需要將它加以分組或分類。

本文以實例說明了團體結構差異對團體比較研究結果的影響，研究者應根據其研究目的和資料特性選擇適當的團體結構差異處理方式，以期得到正確的研究結論。

註 釋

- 1 性別歧視與勞動力市場歧視兩種概念並不完全相同。研究者在使用上述的分解法來討論歧視的問題時，應予以必要的釐清和分開討論。換言

之，一般所謂的性別歧視包含了勞動力市場的性別歧視和非勞動力市場的性別歧視。若結構性差異的原因是團體歧視所造成的，那麼它應該屬於非勞動力市場的性別歧視。至於如何處理性別歧視端視研究者的研究企圖而定。

- 2 因為受到研究資料的限制，本研究中所使用的所得解釋變項在理論上並不完備。理論上，應該還有本研究所用的六個解釋變項以外的其他變項可以詮釋所得的差異。所幸本研究的真正分析重點並不在此。

參考資料

謝雨生

1990 〈臺灣勞動力人口組成變遷對勞力未適當運用的影響〉，《人口學刊》13: 59-82。

1991 〈勞動力人口結構變遷對人力運用的影響〉，《人口學刊》14: 1-28。

Casper, Lynne M. and Sara S. McLanahan

1994 "The Gender-Poverty Gap: What We Can Learn from Other Countries," *American Sociological Review* 59: 594-605.

Clogg, C. C.

1978 "Adjustment of Rates Using Multiplicative Models," *Demography* 15: 523-539.

Clogg, C. C. and J. W. Shockey

1985 "The Effect of Changing Demographic Composition on Recent Trends in Underemployment," *Demography* 22: 395-414.

Clogg, C. C. and Scott R. Eliason

1988 "A Flexible Procedure for Adjusting Rates and Proportions, Including Statistical Methods for Group Comparisons," *American Sociological Review* 53: 267-283.

Freeman, Daniel H. and Theodore R. Holford

1980 "Summary Rates," *Biometrics* 36: 341-352.

Jams, H. H. and A. Thornton

1975 "Decomposition of Differences: A Cautionary Note," *Sociological Methods and Research* 3: 341-352.

Jones, Frank L.

1983 "Sources of Gender Inequality in Income: What the Australian Census Says," *Social Forces* 62: 134-152.

Jones, Frank L. and Jonathan Kelley

1984 "Decomposing Difference Between Groups : A Cautionary Note on Measuring Discrimination," *Sociological Methods and Research* 12: 323-343.

Kalleberg, Arne L.

1988 "Comparative Perspectives on Work Structures and Inequality," *Annual Review of Sociology* 14: 203-225.

Kalleberg, Arne L. and James R. Lincoln

1988 "The Structure of Earnings Inequality in the United States and Japan," *American Journal of Sociology* 94: S121-S153.

Kelley, Jonathan and M. D. R. Evans

1993 "The Legitimation of Inequality: Occupational Earnings in Nine Nations," *American Journal of Sociology* 99: 75-125.

Keyfitz

1968 *Introduction to the Mathematics of Population*. Reading, MA: Addison-Wesley.

Kitagawa, Evelyn M.

1955 "Components of a Difference between Two Rates," *Journal of the American Statistical Association* 50: 1168-1194.

1964 "Standardized Comparisons in Demographic Research," *Demography* 1: 296-315.

Liao, T. F.

- 1989 "A Flexible Approach for the Decomposition of Rate Differences,"
Demography 26: 717-726.

Rosenfeld, Rachel A. and Arne L. Kalleberg

- 1990 "A Cross-National Composition of the Gender Gap in Income,"
American Journal of Sociology 96: 69-106.

Semyonov, Moshe

- 1988 "Bi-Ethnic Labor Market, Mono-Ethnic Labor Markets, and
Socioeconomic Inequality," *American Sociological Review* 53: 256-
266.

Semyonov, Moshe and Yinon Cohen

- 1990 "Ethnic Discrimination and the Income of Majority-Group
Workers," *American Sociological Review* 55: 107-114.

Shryock, Henry S. and Jacob S. Siegel

- 1980 *The Methods and Materials of Demography*. Washington, DC: U.S.
Government Printing Office.

Spiegelman, Mortimer

- 1968 *Introduction to Demography*. Cambridge, MA: Harvard University
Press.

Sweet, J.

- 1984 "Components of Change in Number of Households: 1970-1980,"
Demography 21: 129-140.

Tsai, Shu-Ling and Hei-Yuan Chiu

- 1991 "Constructing Occupational Scales for Taiwan," *Research in Social
Stratification and Mobility* 10: 229-253.

Weeks, J. R.

- 1989 *Population: An Introduction to Concepts and Issues*. 4th ed.
California: Wadsworth Publishing Company.

Wellington, Alison J.

1994 "Accounting for the Male/Female Gap Among Whites: 1976 and 1985," *American Sociological Review* 58: 839-848.

Wunsch, Guillame J. and Marc G. Termote

1978 *Introduction to Demographic Analysis: Principles and Methods*.
New York: Plenum Press.

Xie, Y.

1989 "An Alternative Purging Method: Controlling the Composition-Dependent Interaction in an Analysis of Rates," *Demography* 26: 711-716.

附錄一：迴歸標準化或調整法中團體平均數差異分解詳細過程

因爲男女性原迴歸方程式分別爲式A和式B，

$$\text{Ln}\hat{y}_m = a_m + \sum b_{im} X_{im} \quad (\text{A})$$

$$\text{和 } \text{Ln}\hat{y}_f = a_f + \sum b_{if} X_{if} \quad (\text{B})$$

其中 \hat{y} 爲依變項(所得)的估計值， m 和 f 分別表示男性和女性(團體)， a 是迴歸方程式中的常數項， X 和 b 分別爲自變項和迴歸母數估計值(或未標準化迴歸係數)，另外， i 表示不同的自變項。

所以男女性依變項平均數可以表示如式C和式D，

$$\text{Ln}\bar{y}_m = a_m + \sum b_{im} \bar{X}_{im} \quad (\text{C})$$

$$\text{和 } \text{Ln}\bar{y}_f = a_f + \sum b_{if} \bar{X}_{if} \quad (\text{D})$$

其中 \bar{y} 及 \bar{X} 分別代表依變項與自變項的平均數。

另外，假如男性具有和女性相同的自變項平均數，則可以式 E 表示其調整平均數，以 $\text{Ln}\bar{y}_m^s$ 表示。同樣的，如果女性擁有和男性相同的自變項平均數，則可以式 F 表示其調整平均數，以 $\text{Ln}\bar{y}_f^s$ 表之。

$$\text{Ln}\bar{y}_m^s = a_m + \sum b_{im} \bar{X}_{if} \quad (\text{E})$$

$$\text{和 } \text{Ln}\bar{y}_f^s = a_f + \sum b_{if} \bar{X}_{im} \quad (\text{F})$$

其中 s 代表調整後的意思。因而，文中式 1 的取得過程如下所示，亦即將式 C、式 D、式 E 和式 F 分別代入下列式子，並加以簡化而得。

$$\begin{aligned} \text{Ln}\bar{y}_m - \text{Ln}\bar{y}_f &= (\text{Ln}\bar{y}_m - \text{Ln}\bar{y}_m^s) + (\text{Ln}\bar{y}_f - \text{Ln}\bar{y}_f^s) \\ &\quad + (\text{Ln}\bar{y}_m^s - \text{Ln}\bar{y}_f) + (\text{Ln}\bar{y}_f^s - \text{Ln}\bar{y}_f) \\ &= \frac{(a_m - a_f)}{(1)} + \frac{\sum \bar{X}_{if} (b_{im} - b_{if})}{(2)} \\ &\quad + \frac{\sum b_{if} (\bar{X}_{im} - \bar{X}_{if})}{(3)} + \frac{\sum (\bar{X}_{im} - \bar{X}_{if}) (b_{im} - b_{if})}{(4)} \end{aligned}$$

另外，文中式 2 和式 3 的詳細過程，可以如下所示：

$$\text{Ln}\bar{y}_m - \text{Ln}\bar{y}_f = (\text{Ln}\bar{y}_m - \text{Ln}\bar{y}_f^s) + (\text{Ln}\bar{y}_f^s - \text{Ln}\bar{y}_f)$$

$$\begin{aligned}
 &= \frac{(a_m - a_f) + \sum \bar{X}_{im}(b_{im} - b_{if})}{(1) \quad (2) + (4)} \\
 &+ \frac{\sum b_{if}(\bar{X}_{im} - \bar{X}_{if})}{(3)}
 \end{aligned}$$

$$\begin{aligned}
 \text{Ln}\bar{y}_m - \text{Ln}\bar{y}_f &= (\text{Ln}\bar{y}_m - \text{Ln}\bar{y}_m^s) + (\text{Ln}\bar{y}_m^s - \text{Ln}\bar{y}_f) \\
 &= \frac{(a_m - a_f) + \sum \bar{X}_{if}(b_{im} - b_{if})}{(1) \quad (2)} \\
 &+ \frac{\sum b_{im}(\bar{X}_{im} - \bar{X}_{if})}{(3) + (4)}
 \end{aligned}$$

Covariance Structure Analysis with Intraclass Dependent Observations*

Li-Jen Weng** P. M. Bentler***

I. Introduction

Independence among observations is one of the basic assumptions in covariance structure analysis. However, as was noted, for example, by Freedman (1985), some of the individuals in a conventional cluster sample may have known and interacted with each other. Under such circumstances, independence is unlikely. Violation of the independence assumption may bias the estimates. Moreover, no matter what distributional assumptions are made for the variables, all current statistical theory developed for covariance structure analysis requires the assumption of independence among observations. Once independence is questionable, the associated tests may become invalid.

* The authors would like to thank two anonymous reviewers for their helpful comments on an earlier version of this paper. The research was supported in part by USPHS grants DA 01070 and DA 00017.

** Department of Psychology, National Taiwan University

*** University of California, Los Angeles

In the present study, an attempt is made to approach the problem of dependent observations in covariance structure analysis. A dependence structure among all the observations is assumed. To simplify the problem, the same dependence structure is assumed for all the observed variables. Under the assumption of normality, the matrix normal distribution (Arnold, 1981: 313; de Waal, 1982; Nel, 1977) gives the joint distribution of the elements of a data matrix where both the variables and the observations can be correlated. It is assumed throughout this research that the data matrix follows a matrix normal distribution. The crucial problem for covariance structure analysis is to find a consistent estimator of the covariance matrix among the variables with the dependent effects among observations filtered.

The present research focuses on one case. The dependent structure studied has practical implications, especially for family, genetic, or classroom research. This dependence structure is block-diagonal intraclass, implying independence among groups. Members in each group are assumed to correlate equally with one another. Data collected from twins, couples, or siblings are examples. Many researchers have worked on the asymptotic properties of maximum likelihood estimators (MLE's) with dependent observations (e.g., Amemiya, 1985; Bar-Shalom, 1971; Bhat, 1974; Crowder, 1976; Heijmans & Magnus, 1986a, 1986b, 1986c; Weiss, 1971, 1973). However, the study of special cases is necessary because the published theorems are usually very general and are not guaranteed to be easily applicable to all cases.

Amemiya (1985) presented theorems on consistency and asymptotic normality of extremum estimators. By extremum estimators he means "estimators obtained by either maximizing or minimizing a certain function defined over the parameter space." Observations are not

required to be independent or identically distributed. These theorems are applied to establish the consistency and the asymptotic normality of the MLE's obtained in this study by verifying the assumed conditions. Special attention is given to the dependence structure among observations because not all dependence structures satisfy the sufficient conditions set forth.

This paper is organized as follows. The notation and the matrix normal distribution are first introduced. The asymptotic properties of the maximum likelihood estimators are presented. A two-stage procedure is proposed to use the MLE of the covariance matrix obtained from dependent observations in covariance structure analysis. Finally, examples from simulated and real data are presented. The simulation studies investigate the effects of ignoring dependence among observations while the data are dependent. The real data example analyzes the factor structure among six personality scales based on 77 couples. The implications and the limitations of the model are discussed. Directions for future work are suggested.

II. Notation and the Matrix Normal Distribution

The following notation, unless indicated elsewhere, is used throughout the paper. The $n \times p$ data matrix X represents the observed values from a random sample of n observations on p variables. The vec operator stacks rows of a matrix into a long column. The symbol $\lambda(X)$ represents an eigenvalue of matrix X . The $n \times p$ matrix μ represents the expected values of data matrix X . In other words, $E(x_{ij}) = \mu_{ij}$, $X = [x_{ij}]$ and $\mu = [\mu_{ij}]$. The symmetric positive definite $p \times p$ matrix Σ represents the covariance matrix between the columns of the data matrix

X (i.e., the variables). The symmetric positive definite matrix R ($n \times n$) with all the diagonal elements equal 1.0 represents the correlations between rows of the data matrix X (i.e., observations). Matrix R specifies the dependence structure among the observations.

A matrix X with a moment generating function

$$M_X(t) = \exp[\text{tr}(\mu't) + 2^{-1}\text{tr}(t'Rt\Sigma)],$$

where t is of the same order as X , is said to follow a matrix normal distribution with parameters μ , R , and Σ (Arnold, 1981). $X \sim N_{n,p}(\mu, R, \Sigma)$. The joint density function is

$$f(X) = (2\pi)^{-np/2} |R|^{-p/2} |\Sigma|^{-n/2} \exp\{-2^{-1}\text{tr}[R^{-1}(X-\mu)\Sigma^{-1}(X-\mu)']\},$$

$-\infty < x_{ij} < \infty$, $X = [x_{ij}]$. The usual assumption of independent observations leads to a special case of the above density function with $R = I$, the identity matrix.

The matrix normal distribution implies that the covariance between any two data points depends not only on the covariation between the associated variables but also on the correlation between the observation units. In mathematical form,

$$\text{cov}(x_{ij}, x_{kl}) = \rho_{ik} \sigma_{jl},$$

Where ρ_{ik} is an element of R and σ_{jl} is an element of Σ (Arnold, 1981: 311). Therefore, when observations are independent as assumed in most statistical methods, $\rho_{ik} = 0$ for all $i \neq k$, and data points between any two

observation units yield zero covariance, i.e. $\text{cov}(x_{ij}, x_{kl}) = 0$ for all $i \neq k$. All the observations are identically distributed. The ordinary formula for computing the sample covariance works. However, if observations are dependent, i.e., $\rho_{ik} \neq 0$ for some $i \neq k$, computation of the sample covariance matrix using the ordinary formula and ignoring the dependence structure may yield an estimator which confounds covariation between variables with correlation between observations.

III. Asymptotic Properties of the Maximum Likelihood Estimators $\hat{\rho}$ and $\hat{\Sigma}$

The dependence structure studied is block-diagonal intraclass, implying independence among groups. Each group may consist of different number of observations. Members in each group are assumed to correlate equally with one another. Data collected from classrooms or siblings are examples. Donner and Koval (1980) estimated the intraclass correlation among siblings in univariate cases. They did not prove any asymptotic properties of the maximum likelihood estimators where the observations were dependent.

Let n_g be the number of observations in group g , G be the total number of groups, and n be the sum of all n_g 's. The group size n_g is assumed to be fixed. Further, assume $\lim_{n \rightarrow \infty} G/n = \delta$, $0 < \delta < 1$. Let m_k be the number of groups of size k ; in other words, $m_k = \{\text{number of } g : n_g = k\}$. Assume $\lim_{G \rightarrow \infty} \frac{m_k}{G} = f_k$, $0 \leq f_k \leq 1$ and $\sum_{k=1}^K f_k = 1$ for some finite positive integer K , $K = \max_g n_g < \infty$. Suppose an industrial

psychologist is interested in studying the working attitudes of employees. The employees are organized into work groups of two or three. There are an equal number of these two types of work groups. In the study, there will be 100 samples of group size 2, and 100 samples of group size 3. Thus $n = 500$, $G = 200$, $m_2 = 100$, $m_3 = 100$, $k = 2$ or 3 , $K = 3$, and in the limit $\delta = 0.4$, and $f_2 = f_3 = 0.5$.

For each positive integer n , $X(n)$ is an $n \times p$ random matrix to be observed. Without loss of generality, $X(n)$ is assumed to follow a matrix normal distribution with zero means, row correlation matrix $R(\rho)$, and column covariance matrix Σ ; $X(n) \sim N_{n,p}(0, R(\rho), \Sigma)$. Here $R(\rho)$, representing the dependence structure, is block-diagonal. Each $n_g \times n_g$ diagonal block in $R(\rho)$ has 1.0 on the diagonal and ρ elsewhere. Denote the vector containing the p^* distinct elements in the lower triangle of Σ_0 as σ_0 . Let $\theta_0 = (\sigma_0, \rho_0)$ be the true parameter vector of order q ; $q = p^* + 1$, with $p^* = p(p + 1)/2$. The total number of parameters q is independent of the sample size n . The domain of $\theta = (\sigma, \rho)$, or the parameter space Θ , consists of all possible values that θ can take. σ is the vector with the p^* distinct elements in the lower triangle of Σ . The parameter space Θ is a subset of q -dimensional Euclidean space; $\Theta = \{(\sigma, \rho): \lambda_2 I_p < \Sigma < \lambda_1 I_p, -\varepsilon < \rho < \gamma, \text{ for } 0 < \lambda_2 < \lambda_1 \text{ and some } \varepsilon < K^{-1} \text{ and } \gamma < 1\}$. For matrices A and B of the same order, $A > B$ in the Loewner sense of inequality, if and only if $A - B$ is positive definite (see, e.g., Browne, 1974: 10; Browne & Shapiro, 1988: 207; Kano & Shapiro, 1987, Theorem 1). The definition of the parameter space implies that the eigenvalues of Σ are bounded between λ_2 and λ_1 .

Theorem. The maximum likelihood estimators of Σ and ρ , $\hat{\theta} = (\hat{\sigma}, \hat{\rho})$, under the conditions specified above are consistent and

asymptotically normal with $\text{plim } \hat{\theta} = \theta_0$, and $\sqrt{n}(\hat{\theta} - \theta_0) \rightarrow N[0, A(\theta_0)^{-1}]$, where $A(\theta_0) = \lim E n^{-1}(\partial^2 L_n / \partial \theta \partial \theta')$ at θ_0

$$= \frac{1}{2} \begin{bmatrix} K_p^{-1} (\Sigma_0^{-1} \otimes \Sigma_0^{-1}) K_p^{-1} & \left[\delta \sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} - \frac{1-\delta}{1-\rho_0} \right] K_p^{-1} \text{vec} \Sigma_0^{-1} \\ \text{Sym.} & \left[\delta \sum_{k>1}^K f_k \left(\frac{k-1}{1+(k-1)\rho_0} \right)^2 + \frac{1-\delta}{(1-\rho_0)^2} \right] \mathbf{p} \end{bmatrix}$$

The matrix K_p is of order $p^2 \times p^2$ with typical element $[K_p]_{ij,gh} = 2^{-1}(\delta_{ig}\delta_{jh} + \delta_{ih}\delta_{jg})$, $i \leq p, j \leq p, g \leq h \leq p$ and δ_{ij} represents Kronecker's delta. And $K_p^{-1} = (K_p' K_p)^{-1} K_p'$ is a left inverse of K_p of order $p(p+1)/2 \times p^2$. See, for example, Browne (1974).

The asymptotic variance of $\sqrt{n}(\hat{\rho} - \rho_0)$, denoted as $\psi(\rho_0)$, is

$$2p^{-1}\delta^{-1} \left\{ \sum_{k>1}^K f_k \left[\frac{k-1}{1+(k-1)\rho_0} \right]^2 - \delta \left[\sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} \right]^2 + 2 \frac{1-\delta}{1-\rho_0} \sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} + \frac{1-\delta}{(1-\rho_0)^2} \right\}^{-1}$$

So, an estimate of the asymptotic variance of $\hat{\rho}$ is given by $\psi(\rho_0)/n$.

Amemiya's (1985) theorem 4.1.6 on extremum estimators is applied to prove the consistency and asymptotic normality of the MLE's of ρ and Σ . The verification of the required conditions is given in the Appendix for the sake of completeness.

The maximum likelihood estimator $\hat{\theta}$ is asymptotically efficient because its asymptotic variance-covariance matrix reaches the Cramer-Rao lower bound (see, e.g., Amemiya, 1985, Definition 4.2.1). The likelihood ratio test of $\rho_0 = 0$ is defined as $2(F_0 - F_1)$; where F_0 is the function value of the negative log-likelihood under the null hypothesis H_0 of no dependence and F_1 is the function value under the alternative hypothesis H_1 of non-zero ρ_0 . This statistic is asymptotically distributed as a chi-square variate with 1 degree of freedom (see, e.g., Amemiya, 1985, Section 4.5.1). This test can be used to evaluate the appropriateness of the dependence assumption among observations.

IV. A Two-Stage Procedure for Covariance Structure Analysis

A Consistent estimator of the covariance matrix is one basic component in covariance structure analysis. $\hat{\Sigma}$ and $\hat{\rho}$ in the model have been shown to be consistent and asymptotically normal. Weng and Bentler (1987) discussed the use of $\hat{\Sigma}$ in covariance structure analysis when the data follow a matrix normal distribution and the dependence structure R is known. In the present research with an unknown parameter in the dependence structure, if the sample size is sufficiently large and one takes ρ at its estimated value $\hat{\rho}$, $\hat{\Sigma}$ can be used in any computer package for covariance structure analysis. Using $\hat{\Sigma}$ instead of S in the

analysis yields appropriate estimates. S represents the usual sample covariance matrix estimate under independence.

A two-stage procedure is proposed for cases with sufficient sample sizes. $\hat{\Sigma}$ and $\hat{\rho}$ are estimated and evaluated at the first stage. At the second stage, covariance structure analysis is performed with an appropriately chosen input matrix. If the result of the likelihood ratio test indicates that ρ_0 is not significantly different from zero, observations can be treated as independent, and the usual analysis procedures using S follow. If ρ_0 is significantly different from zero, one can use $\hat{\Sigma}$ in analysis, while fixing ρ at $\hat{\rho}$. With ρ fixed at its estimate, the observed data can be transformed to independently and identically distributed vectors, while still preserving the original covariance matrix among the variables. Standard asymptotic statistical theories for covariance structure analysis are applicable.

Suppose X follows a matrix normal distribution, $N_{n,p}(0, R, \Sigma)$, with $R = R(\rho)$. Decompose R as AA' , where A is a square matrix of order n . The transformed matrix, $Y = A^{-1}X$, is distributed as $N_{n,p}(0, I, \Sigma)$. With $\hat{\rho}$ being consistent, we have $\hat{R} \xrightarrow{P} R$ and $\hat{A} \xrightarrow{P} A$, where $\hat{R} = R(\hat{\rho})$ and $\hat{A} = A(\hat{\rho})$. Asymptotically, $Y^* = \hat{A}^{-1}X$ has the same distribution as Y . Note that each row in Y is independently, identically distributed as $N(0, \Sigma)$. Since the transformed data have the same covariance matrix as the original data, analyses can be carried out by using the transformed independent data instead of the original dependent data. Existing programs for covariance structure analysis can be used. Statistical theories based on the i.i.d. assumption are applicable.

The two-stage procedure has several advantages. First, one can evaluate the degree of dependence among observations prior to any

analysis. If the degree of dependence is negligible, the data can be treated as independent and the usual analysis procedure follows. Second, when dependence exists among observations, with ρ fixed at its estimate and with sufficient sample size, the transformation approach discussed above is legitimate. Calculation of the sample covariance matrix among observed variables is straightforward without the need for developing new estimator for it. Third, as long as $\hat{\Sigma}$ is obtained, any standard package program for covariance structure analysis can be used to get model parameter estimates and other statistics. However, when sample size is not large enough, the transformed data may not be independently, identically distributed because of the sampling error of $\hat{\rho}$. The sampling error of $\hat{\rho}$ may affect parameter estimates and associated standard errors, as well as the chi-square statistics.

Alternatively, one may simultaneously estimate the intraclass correlation and the model parameters. However, if the model is rejected with the one-stage estimation procedure, it is difficult to detect which part of the model breaks down. Moreover, the statistical properties of estimators under the one-stage procedure have to be proved otherwise. Statistical theory in covariance structure analysis under the independence assumption is not applicable. The two-stage procedure has the advantage of decomposing the problem into two components and examining them separately. As long as the transformation approach is appropriate, existing computer programs for covariance structure analysis can be used to obtain parameter estimates and available statistical theory is applicable.

Note that $\hat{\rho}$, the MLE of ρ , is not the only estimator that can be used in the two-stage procedure. Any consistent estimator is appropriate. One may be able to obtain estimators simpler than MLE.

V. Simulations and Example

This section presents the results from two simulation studies and one real data analysis. The purpose of the first simulation is to compare S and $\hat{\Sigma}$, where S represents the usual sample covariance matrix under independence and $\hat{\Sigma}$ represents the MLE of the population covariance matrix under the assumption of a matrix normal distribution with an intraclass dependence structure. The second simulation investigates the effect of using S and $\hat{\Sigma}$ in a tow-factor analytic model. The real data example looks into the factor structure of six personality measures on 77 couples.

(I) Simulation 1

Two cases of six variables are studied. The total sample size is fixed at 200, which is considered acceptable for six variables. In Case I, $m=2$ and $G=100$. In Case II, $m=10$ and $G=20$. One hundred replications were performed for each case. The population intraclass correlation (ρ_0) ranges from 0.0 to 0.9. The population covariance matrix (Σ_0) is:

$$\Sigma_0 = \begin{bmatrix} 9.000 & 4.410 & 4.410 & 1.323 & 1.323 & 1.323 \\ 4.410 & 9.000 & 4.410 & 1.323 & 1.323 & 1.323 \\ 4.410 & 4.410 & 9.000 & 1.323 & 1.323 & 1.323 \\ 1.323 & 1.323 & 1.323 & 9.000 & 4.410 & 4.410 \\ 1.323 & 1.323 & 1.323 & 4.410 & 9.000 & 4.410 \\ 1.323 & 1.323 & 1.323 & 4.410 & 4.410 & 9.000 \end{bmatrix}$$

Two estimators of the covariance matrix, S and $\hat{\Sigma}$, are obtained. $S = X'X/(n - 1)$. $\hat{\Sigma}$ is obtained from the Newton-Raphson optimization procedure (see, e.g., Fox, 1971, Section 2.10). The starting values are set as $\hat{\Sigma} = S$ and $\hat{\rho} = 0$. The iteration stops if the sum of the absolute values of the parameter changes is no greater than 0.001.

Table 1 RMSE^a of s and $\hat{\sigma}$

ρ_0		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
m	I										
2	S	0.743	0.736	0.728	0.734	0.794	0.835	0.858	0.890	0.950	0.970
2	$\hat{\Sigma}$	0.740	0.731	0.715	0.706	0.748	0.736	0.739	0.749	0.747	0.749
10	S	0.713	0.745	0.853	0.971	1.127	1.268	1.525	1.686	1.845	2.103
10	$\hat{\Sigma}$	0.710	0.722	0.702	0.767	0.772	0.775	0.820	0.844	0.910	0.933

$$^a \text{Average RMSE}(s) = \frac{1}{100} \sum_{i=1}^{100} \sqrt{\sum_{j=1}^p \sum_{k=1}^j (s_{ijk} - \alpha_{0ijk})^2 / p^*}, \text{ and}$$

$$\text{Average RMSE}(\hat{\sigma}) = \frac{1}{100} \sum_{i=1}^{100} \sqrt{\sum_{j=1}^p \sum_{k=1}^j (\hat{\sigma}_{ijk} - \sigma_{0ijk})^2 / p^*}.$$

m = Number of observations in each group

I = Input matrix used in the analysis

Table 1 presents the average root-mean-squared-error (RMSE) over 100 replications for S and $\hat{\Sigma}$. The average RMSE's for $\hat{\Sigma}$ are small and the average RMSE of S is, in general, greater than that of $\hat{\Sigma}$ for any given intraclass correlation. All the average RMSE's fall between 0.7 and 1.0 except for those of S in Case II. In Case I, the Difference between the average RMSE's for S and $\hat{\Sigma}$ is less than 0.1 for $0.0 < \rho_0 < 0.4$, and the

difference gradually increases as ρ_0 increases. In Case II, where m equals 10, the average RMSE's of S are much greater than those of $\hat{\Sigma}$. The results suggest that, when the total sample size is fixed, the difference between S and $\hat{\Sigma}$ increases (1) as the degree of dependence represented by the intraclass correlation increases, and (2) as the number of observations in each group increases or equivalently as the number of groups decreases.

The RMSE and bias of $\hat{\rho}$ are presented in Table 2. The MLE of the intraclass correlation under a matrix normal distribution performs quite well with small RMSE and bias.

Table 2 RMSE^a and Bias^b of $\hat{\rho}$

$m \backslash \rho_0$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
RMSE										
2	0.038	0.044	0.040	0.038	0.036	0.028	0.029	0.023	0.014	0.009
10	0.014	0.023	0.032	0.034	0.033	0.034	0.040	0.034	0.024	0.013
BIAS										
2	0.000	0.002	0.000	-0.004	-0.005	-0.002	0.004	0.000	0.003	-0.001
10	0.001	-0.004	0.008	-0.003	-0.002	-0.003	-0.003	-0.005	-0.005	0.000

$${}^a\text{RMSE}(\hat{\rho}) = \sqrt{\frac{100}{\sum_{i=1}^{100} (\hat{\rho}_i - \rho_0)^2} / 100}.$$

$${}^b\text{Bias}(\hat{\rho}) = E(\hat{\rho}) - \rho_0 = \frac{100}{\sum_{i=1}^{100} \hat{\rho}_i} / 100 - \rho_0.$$

m = Number of observations in each group

The likelihood ratio test (LRT) of the hypothesis: $H_0: \rho_0 = 0$ versus $H_1: \rho_0 \neq 0$ is performed in each sample. The likelihood ratio test when

H_0 is true has an asymptotic chi-square distribution with 1 degree of freedom. The probability for Type I error is chosen at the $\alpha = 0.5$ level. The average test values and the empirical power or probability of Type I error (for $\rho_0 = 0.0$) of the test during the 100 replications are presented in Table 3. The empirical power of the test is the frequency with which H_0 was correctly rejected over the total number of replications. The mean of the chi-square statistic is close to 1.0 at $\rho_0 = 0$, but far from 1.0 for $\rho_0 > 0$. The test appears to be very powerful. In Case I with $m = 2$, the empirical power of the test reaches 1.00 for $\rho_0 \geq 0.2$. In Case II with $m = 10$, the empirical power of the LRT reaches 1.00 for $\rho_0 \geq 0.1$. A power analysis is conducted to study the theoretical power of the test. The series of hypotheses tested are $H_0: \rho_0 = 0.0$ versus $H_1: \rho_0 = 0.1$, up to 0.9. For the sake of simplicity, the power analysis is based on the asymptotic normal distribution of $\hat{\rho}$. Power = Prob ($\rho^2 / (\psi(\rho_0) / n) > 3.841 \mid H_1$ is true). $\rho^2 / (\psi(\rho_0) / n)$ under H_1 is asymptotically distributed as a non-central chi-square variate with 1 degree of freedom and non-centrality parameter $\rho_0^2 / (\psi(\rho_0) / n)$, $\rho_0 = 0.1$ to 0.9 (see, e.g., Hogg & Craig, 1978, Section 8.4). The middle of Table 3 gives the theoretical power of the test for population intraclass correlation ranging from 0.1 to 0.9. The results of the power analysis support the high frequency of rejection in the simulation.

Smaller sample sizes should decrease the power of the test. Another simulation is conducted to look into the empirical power of the test with smaller samples in comparison with theoretical power. The same setup is used except for the sample size. In this case, the sample size reduces to 100 with $m = 2$ and $G = 50$. The results are given in the last two rows of Table 3. The power is reduced as a result of a decrease in sample size. But, the power of the test is still very high and reaches 1.00 for $\rho_0 \geq 0.3$.

Table 3 Mean of the Chi-Square Statistic and Power^a of the Test

$m \backslash \rho_0$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
(n = 200)										
<i>Mean of the Test Statistic</i>										
2	0.86	7.91	24.83	54.49	99.99	167.59	266.35	396.56	609.82	977.42
10	0.96	33.64	118.35	211.47	338.92	497.64	701.55	961.46	1347.85	2074.14
<i>Empirical Power of the Test</i>										
2	0.02	0.69	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
10	0.04	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
<i>Theoretical Power of the Test</i>										
2	0.05	0.70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
10	0.05	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(n = 100)										
<i>Empirical Power of the Test</i>										
2	0.07	0.46	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00
<i>Theoretical Power of the Test</i>										
2	0.05	0.42	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00

^aFor $\rho_0=0.0$ the number represents the probability for Type I error rather than the power of the test.

m = Number of observations in each group

n = Total sample size

(II) Simulation 2

The purpose of the second simulation is to compare the results of a confirmatory factor analysis using either S or $\hat{\Sigma}$ as the input matrix, while the data are matrix normally distributed with intraclass dependence structure. The analysis based on $\hat{\Sigma}$ corresponds to the proposed two-

stage methodology. In this model $\Sigma_0 = \Lambda_0 \Phi_0 \Lambda_0' + \Psi_0$. The two-factor model includes six variables with three indicators for each factor. The population factor loading matrix (Λ_0), the factor covariance matrix (Φ_0), and the residual matrix (Ψ_0) are as follows.

$$\Lambda_0 = \begin{bmatrix} 2.1 & 0 \\ 2.1 & 0 \\ 2.1 & 0 \\ 0 & 2.1 \\ 0 & 2.1 \\ 0 & 2.1 \end{bmatrix}$$

$$\Phi_0 = \begin{bmatrix} 1.0 & 0.3 \\ 0.3 & 1.0 \end{bmatrix}$$

$$\Psi_0 = \begin{bmatrix} 4.59 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4.59 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4.59 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4.59 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4.59 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4.59 \end{bmatrix}$$

The two cases studied are the same as in Simulation 1. The sample size of 200 is considered sufficient for the two-factor model studied. Therefore, the transformation approach and the proposed two-stage methodology may be applied. One hundred replications were performed. The parameters of the factor model were estimated using the MLE option in EQS (Bentler, 1989).

Table 4 (continued)

ρ_0		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9		
P	m	l	S	2.1174	2.1109	2.0988	2.1102	2.1098	2.1314	2.0979	2.1218	2.0499	2.1340
				(.202)	(.226)	(.254)	(.324)	(.382)	(.412)	(.416)	(.527)	(.622)	(.705)
10	\hat{S}	2.1124	2.1062	2.0984	2.1079	2.0880	2.1320	2.0930	2.1199	2.1236	2.0897		
		(.201)	(.234)	(.208)	(.232)	(.235)	(.222)	(.258)	(.255)	(.278)	(.270)		
λ_{42}	2	S	2.1056	2.1110	2.1182	2.0959	2.1151	2.1337	2.1184	2.0941	2.0939	2.1040	
			(.236)	(.240)	(.232)	(.222)	(.241)	(.231)	(.255)	(.284)	(.296)	(.323)	
10	\hat{S}	2.1009	2.1044	2.1102	2.0921	2.1143	2.1280	2.1052	2.1084	2.0984	2.1072		
		(.236)	(.237)	(.229)	(.211)	(.219)	(.223)	(.221)	(.240)	(.238)	(.232)		
2	S	2.1127	2.1259	2.0940	2.0684	2.0974	2.0808	2.1247	2.0797	2.0745	2.1511		
		(.239)	(.238)	(.248)	(.260)	(.346)	(.377)	(.491)	(.549)	(.581)	(.764)		
10	\hat{S}	2.1077	2.1184	2.1046	2.0803	2.0966	2.0652	2.0829	2.1090	2.1567	2.0861		
		(.238)	(.235)	(.206)	(.207)	(.216)	(.226)	(.286)	(.263)	(.241)	(.272)		
λ_{52}	2	S	2.0726	2.1153	2.1014	2.0556	2.1043	2.1668	2.0893	2.0912	2.0815	2.1110	
			(.208)	(.225)	(.233)	(.227)	(.241)	(.242)	(.276)	(.288)	(.254)	(.352)	
10	\hat{S}	2.0667	2.1114	2.0874	2.0501	2.1145	2.1624	2.0855	2.0970	2.0945	2.1290		
		(.209)	(.223)	(.238)	(.223)	(.239)	(.227)	(.216)	(.208)	(.232)	(.256)		
2	S	2.1406	2.1123	2.0715	2.1383	2.1253	2.1640	2.1106	2.0342	2.1245	2.0540		
		(.209)	(.251)	(.296)	(.307)	(.383)	(.482)	(.481)	(.477)	(.635)	(.680)		
10	\hat{S}	2.1343	2.1084	2.0678	2.1182	2.1211	2.1236	2.1091	2.1043	2.1169	2.0942		
		(.210)	(.230)	(.230)	(.234)	(.229)	(.228)	(.260)	(.210)	(.273)	(.269)		

Table 4 (continued)

ρ_0		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
P	m										
	I										
λ_{62}	S	2.0918 (.247)	2.1136 (.228)	2.0935 (.232)	2.1138 (.259)	2.0901 (.249)	2.1150 (.254)	2.0592 (.237)	2.0408 (.269)	2.1189 (.339)	2.1925 (.316)
	$\hat{\Sigma}$	2.0863 (.246)	2.1044 (.227)	2.0904 (.225)	2.1175 (.244)	2.0962 (.230)	2.0910 (.234)	2.0798 (.202)	2.0524 (.231)	2.1150 (.246)	2.1543 (.255)
.10	S	2.1088 (.234)	2.1148 (.237)	2.0922 (.265)	2.1241 (.295)	2.1424 (.343)	2.1256 (.410)	2.1066 (.513)	2.1020 (.484)	2.0848 (.636)	2.0693 (.761)
	$\hat{\Sigma}$	2.1017 (.235)	2.1032 (.219)	2.1041 (.224)	2.1354 (.237)	2.1345 (.203)	2.0925 (.238)	2.0703 (.227)	2.0688 (.220)	2.1226 (.263)	2.0916 (.280)
2	S	0.3062 (.083)	0.2847 (.099)	0.3067 (.095)	0.3027 (.094)	0.3026 (.093)	0.2912 (.098)	0.2977 (.119)	0.2765 (.108)	0.2969 (.116)	0.3203 (.113)
	$\hat{\Sigma}$	0.3072 (.083)	0.2861 (.097)	0.3076 (.094)	0.3033 (.092)	0.3004 (.086)	0.2919 (.085)	0.2965 (.106)	0.2720 (.096)	0.2938 (.083)	0.3118 (.090)
ϕ_{21}	S	0.2925 (.083)	0.3139 (.088)	0.2882 (.103)	0.2979 (.120)	0.3145 (.138)	0.2864 (.187)	0.2929 (.187)	0.2828 (.229)	0.3195 (.259)	0.3124 (.330)
	$\hat{\Sigma}$	0.2926 (.083)	0.3137 (.086)	0.2905 (.088)	0.3092 (.092)	0.3037 (.083)	0.3073 (.096)	0.2983 (.097)	0.3031 (.096)	0.3078 (.096)	0.3217 (.080)

^aStandard deviation is in parentheses.

P = Parameter

m = Number of observations in each group

I = Input Matrix used in the analysis

Means and standard deviations of the six factor loading estimates and the estimated factor covariance over the 100 replications are given in Table 4. All the means of the parameter estimates are very close to the true values regardless of the population intraclass correlation, the number of observations in each group, or the input covariance matrix used in the analysis. As $\rho_0 = 0$, the factor loading estimates from S are always greater than those from $\hat{\Sigma}$ by approximately 0.005. This results from different denominators used for S and $\hat{\Sigma}$: (n-1) for S and n for $\hat{\Sigma}$. Therefore, S equals approximately $(200/99)\hat{\Sigma}$. Factor loadings estimates from S are greater than factor loadings from $\hat{\Sigma}$ approximately by a multiplier of $(200/199)^{1/2}$.

When we look at standard deviations of the parameter estimates, discrepancies among cases appear. The standard deviations of all factor loading estimates show similar patterns. The standard deviations of parameter estimates when S is used are usually larger than those with $\hat{\Sigma}$ being used, except for very small ρ_0 , say under 0.4 for Case I and under 0.2 for Case II. This result seems to indicate that the parameter estimates using S have less empirical efficiency than using $\hat{\Sigma}$. When $\hat{\Sigma}$ is used, the standard deviations of the parameter estimates have small differences between the two cases. When S is used, the standard deviations in Case II are larger than those in Case I, and the difference increases as ρ_0 increases. The results seem to indicate that although the mean parameter estimates are close regardless of group size or the input matrix used, the use of S leads to estimates with greater dispersion, especially for data with large groups and high degrees of dependence, when the total sample size is held constant.

Table 5 Frequency of Model Rejection

ρ_0	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
m I										
2 S	3	6	6	7	10	9	13	23	32	35
2 $\hat{\Sigma}$	3	6	6	4	5	2	3	7	7	3
10 S	6	12	16	43	57	75	86	94	98	98
10 $\hat{\Sigma}$	6	9	7	7	7	5	4	4	8	6

m = Number of observations in each group

I = Input matrix used in the analysis

Table 6 Mean and Standard Deviation^a of the Chi-Square Statistic: $\chi^2(8)$

ρ_0	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
m I										
2 S	8.526 (4.2)	8.757 (4.2)	8.224 (4.3)	8.267 (4.1)	9.938 (4.5)	8.928 (4.4)	10.741 (4.6)	12.268 (5.6)	13.080 (6.4)	13.743 (6.5)
2 $\hat{\Sigma}$	8.520 (4.2)	8.648 (4.0)	7.939 (4.3)	7.687 (3.7)	8.311 (4.1)	7.314 (3.4)	7.550 (3.2)	7.914 (3.9)	8.260 (4.5)	7.683 (3.3)
10 S	7.695 (3.8)	9.239 (5.2)	10.708 (4.7)	14.812 (7.3)	18.339 (9.1)	25.201 (11.4)	34.415 (17.7)	42.673 (20.0)	55.675 (25.7)	71.865 (31.8)
10 $\hat{\Sigma}$	7.734 (3.8)	8.660 (4.6)	8.154 (3.7)	8.659 (4.2)	8.002 (4.4)	7.841 (3.9)	7.938 (4.1)	7.861 (4.5)	8.375 (4.2)	8.225 (4.6)

^aStandard deviation is in parentheses.

m = Number of observations in each group

I = Input matrix used in the analysis

The frequency with which the two-factor model was rejected and the mean and standard deviation of the chi-square test statistic are summarized in Table 5 and Table 6. The expected frequency of rejection is 5 with $\alpha = .05$. The test associated with the model has 8 degrees of freedom. The mean and standard deviation of the statistic are expected to be 8 and 4, respectively.

The model tends to be rejected too often when S is used as the input matrix, especially in Case II. When $\hat{\Sigma}$ is used, the frequency of model rejection ranges from 2 to 9, and the chi-square test statistic behaves very well with a mean close to 8 and a standard deviation close to 4 in both cases. In Case I with S as the input matrix, the test performs fine for $\rho_0 < 0.4$, while its mean and standard deviation begin to increase as $\rho_0 \geq 0.4$. In Case II using S, the test did not perform well even with $\rho_0 = 0.1$, while the mean and standard deviation of the test statistic increases dramatically as ρ_0 increases. The results indicate that the test statistic is asymptotically distributed as a chi-square variate with 8 degrees of freedom if $\hat{\Sigma}$ is used as the input matrix, but it does not have an asymptotic chi-square distribution if S is used in the analysis.

(III) An Example

The sample consisted of 77 couples. The Bentler Psychological Inventory (BPI) (Comrey, Backer, & Glaser, 1973), which assesses 28 personality traits, was administered to the couples. An interested reader may refer to Bentler and Newcomb (1978) for detailed description of the sample, the data collection procedures, and the BPI. For demonstration purpose, six personality traits were selected based on husband-wife correlations on the scales (Bentler & Newcomb, 1978) and the factor structure among the personality traits (see, e.g. Stein, Newcomb, &

Bentler, 1987). Deliberateness, Diligence, Orderliness, Law Abidance, Liberalism, and Religious Commitment were chosen. The first three reflect the factor of Conscientiousness and the rest reflect the factor of Social Conformity.

A confirmatory two-factor analytic model with three indicators on each factor is employed. Separate analyses were first conducted on husband and wife samples to ensure the appropriateness of pooling together the data. The fits in both samples are acceptable. For husbands, $\chi^2(8) = 10.247$, $p = .248$, the factor loading of liberalism on Social Conformity was not significant. For wives, $\chi^2(8) = 13.953$, $p = .083$, and all the parameter estimates were significant. The data on husbands and wives were pooled together to form the couple sample of size 154. The husband and wife in each couple are assumed to be dependent. The sample size of 154 is considered at least moderate for the transformation approach and the two-stage methodology to be used.

The model using S as the input matrix for the couple data cannot be rejected; $\chi^2(8) = 12.079$, $p = .148$ and all the parameter estimates are significant. Means were subtracted from the original data prior to the estimation of $\hat{\Sigma}$. The Newton-Raphson iteration method was used to obtain $\hat{\Sigma}$ and $\hat{\rho}$. $\hat{\rho} = 0.193$ and the LRT, asymptotically distributed as a chi-square variate with 1 degree of freedom, yields 22.056, rejecting the hypothesis of no dependence among husbands and wives on the six personality traits. The two-factor model, using $\hat{\Sigma}$ as the input matrix, cannot be rejected, either; $\chi^2(8) = 12.039$, $p = .149$, with all the parameter estimates being significant. The parameter estimates and their associated standard errors are summarized in Table 7. The high similarity in the results coincides with the simulation results with $\hat{\rho}$ being close to

0.2, since when $\rho_0 = 0.2$, the results for S and $\hat{\Sigma}$ in both simulation studies are essentially the same for Case I with $m = 2$.

Table 7 Parameter Estimates and Associated Standard Errors for the Couples Example

Input Matrix Used Parameter (θ)	S		$\hat{\Sigma}$	
	$\hat{\theta}$	S.E.	$\hat{\theta}$	S.E.
λ_{11} (DB)	2.577	(.656)	2.875	(.676)
λ_{21} (DG)	4.322	(.791)	4.201	(.769)
λ_{31} (OD)	4.029	(.772)	4.185	(.767)
λ_{42} (LA)	4.087	(.678)	4.177	(.696)
λ_{52} (LB)	-2.041	(.632)	-1.799	(.610)
λ_{62} (RC)	3.878	(.785)	3.647	(.755)
ϕ_{21}	0.511	(.122)	0.528	(.122)
ψ_{11} (DB)	33.463	(4.436)	34.936	(4.756)
ψ_{22} (DG)	30.806	(6.293)	32.893	(6.053)
ψ_{33} (OD)	33.245	(5.967)	32.849	(6.023)
ψ_{44} (LA)	11.453	(4.873)	10.682	(5.141)
ψ_{55} (LB)	36.945	(4.500)	33.589	(4.269)
ψ_{66} (RC)	38.627	(6.198)	36.466	(5.718)

DB = Deliberateness

DG = Diligence

OD = Orderliness

LA = Law Abidance

LB = Liberalism

RC = Religious Commitment

VI. Discussion

The presented model has practical applications in many areas of research. For example, an industrial psychologist may observe similarity in working attitudes among employees under the same supervisor. This similarity can be either the cause or the result of being assigned to the same supervisor. Suppose the psychologist is interested in the relationships among various working attitudes and intends to conduct a factor analysis on these attitude scales. One immediate difficulty encountered is the dependence existing among the employees under the same supervisors. The current research offers one solution to this problem.

The model is limited in some aspects. The limitations provide directions for future research. First, the matrix normal distribution assumes one dependence structure for all the variables. This constraint is unacceptable in certain research. Second, the degree of dependence among observations within groups are assumed to be identical. This restriction can be relaxed to allow for different intraclass correlations. In addition, other dependence structures can be studied for different research designs. Stadge (1984) pointed out that not every dependence structure yields consistent estimators. One has to prove the properties of the estimators under different dependence structures.

Transformation is a useful approach for certain classes of dependence structures. Suppose the dependence structure can be decomposed as follows: $R = KLK'$, with K known and L unknown diagonal. The intraclass dependence studied is a special case of this general model. If the eigenvectors of the dependence structure do not depend on any parameters, the data can be transformed to independent,

but not identically distributed vectors. The distributions of the vectors depend on L and Σ . The proofs involving independent, but not identically distributed observations, are usually simpler than those involving dependent observations. Extra conditions are necessary for the parameter estimates to have desirable statistical properties.

Another line of research involves dependent but non-normally distributed data. The matrix elliptical distribution is one possibility. The transformation approach should give uncorrelated and not identically distributed vectors. Whether the transformed vectors are independent depends on the nature of the matrix-valued distributions. More work is needed.

References

Amemiya, T.

1985 *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

Arnold, S. F.

1981 *The Theory of Linear Models and Multivariate Analysis*. New York: Wiley.

Bar-Shalom, Y.

1971 "On the Asymptotic Properties of the Maximum-likelihood Estimate Obtained from Dependent Observations," *Journal of the Royal Statistical Society Series B*, 33: 72-77.

Bentler, P. M.

1989 *EQS Structural Equations Program Manual*. Los Angeles: BMDP Statistical Software, Inc.

Bentler, P. M., & M. D. Newcomb

- 1978 "Longitudinal Study of Marital Success and Failure," *Journal of Consulting and Clinical Psychology* 46: 1053-1070.

Bhat, B. R.

- 1974 "On the Method of Maximum-Likelihood for Dependent Observations," *Journal of the Royal Statistical Society Series B*, 36: 48-53.

Bradley, R. A., & J. J. Gart

- 1962 "The Asymptotic Properties of ML Estimators When Sampling from Associated Populations," *Biometrika* 49: 205-214.

Browne, M. W.

- 1974 "Generalized Least Squares Estimators in the Analysis of Covariance Structures," *South African Statistical Journal* 8: 1-24.

Comrey, A. L., T. E. Backer & E. M. Glaser

- 1973 *A Sourcebook for Mental Health Measures*. Los Angeles: Human Interaction Research Institute.

Crowder, M.J.

- 1976 "Maximum Likelihood Estimation for Dependent Observations," *Journal of the Royal Statistical Society Series B*, 38: 45-53.

de Waal, D. J.

- 1982 "Matrix-Valued Distributions," in S. Kotz & N. Johnson (eds.), *Encyclopedia of Statistical Sciences*, 5: 326-333. New York: Wiley.

Donner, A. & J. J. Koval

- 1980 "The Estimation of Intraclass Correlation in the Analysis of Family Data," *Biometrics* 36: 19-25.

Fox, R. L.

- 1971 *Optimization Methods for Engineering Design*. Reading, MA: Addison-Wesley.

Freedman, D. A.

- 1985 "Statistics and the Scientific Method," in W. M. Mason & S. E.

Fienberg (eds.), *Cohort Analysis in Social Research*, pp.343-366.
New York: Springer-Verlag.

Graybill, F. A.

1983 *Matrices with Applications in Statistics, 2nd Ed.* Belmont, CA:
Wadsworth.

Heijmans, R. D. H., & J. R. Magnus

1986a "On the First-order Efficiency and Asymptotic Normality of
Maximum Likelihood Estimators Obtained from Dependent
Observations," *Statistic Neerlandica* 40: 169-188.

1986b "Consistent Maximum-Likelihood Estimation with Dependent
Observations: The General (Non-normal) Case and the Normal
Case," *Journal of Econometrics* 32: 253-285. Corrigenda in:
Journal of Econometrics 35: 395.

1986c "Asymptotic Normality of Maximum Likelihood Estimators
Obtained from Normally Distributed but Dependent
Observations," *Econometric Theory* 2: 374-412.

Hoadley, B.

1971 "Asymptotic Properties of Maximum Likelihood Estimators for the
Independent Not Identically Distributed Case," *The Annals of
Mathematical Statistics* 42: 1977-1991.

Hogg, R. V., & A. T. Craig

1978 *Introduction to Mathematical Statistics, 4th Ed.* New York:
Macmillan.

Nel, H. M.

1977 "On Distributions and Moments Associated with Matrix Normal
Distribution," *Technical Report No.24*, Department of Mathematical
Statistics, University of the Orange Free State, Bloemfontein, South
Africa.

Philippou, A. N., & G. G. Roussas

- 1973 "Asymptotic Normality of the Maximum Likelihood Estimate in the Independent Not Identically Distributed Case," *Annals of the Institute of Statistical Mathematics* 27: 45-55.

Stadje, W.

- 1984 "A Note on Sample Means and Variances of Dependent Normal Variables," *Mathematische Operationsforschung und Statistik, series statistics* 15: 205-218.

Stein, J. A., M. D. Newcomb & P. M. Bentler

- 1987 "Personality and Drug Use: Reciprocal Effects across Four Years," *Personality and Individual Differences* 8: 419-430.

Weiss, L.

- 1971 "Asymptotic Properties of Maximum Likelihood Estimators in Some Nonstandard Cases," *Journal of the American Statistical Association* 66: 345-350.
- 1973 "Asymptotic Properties of Maximum Likelihood Estimators in Some Nonstandard Cases, II," *Journal of the American Statistical Association* 68: 428-430.

Weng, J. L.

- 1990 "Aspects of Covariance Structure Analysis with Dependent Observations," Ph. D. Dissertation, University of California, Los Angeles (Unpublished).

Weng, L.-J., & P. M. Bentler

- 1987 "Linear Structural Equation Modeling with Dependent Observations," *Proceedings of the Social Statistics Section, American Statistical Association*: 498-500.

Appendix

Theorem 4.1.6 of Amemiya (1985) is used to prove the consistency and the asymptotic normality of the maximum likelihood estimators $\hat{\theta}$. Let $L_n(\theta)$ denote the function to be minimized over the parameter space Θ . The following conditions have to be met.

(A) $\partial^2 L_n / \partial \theta \partial \theta'$ exists and is continuous in an open, convex neighborhood of θ_0 .

(B) $n^{-1}(\partial^2 L_n / \partial \theta \partial \theta')$ θ_n^* converges to a finite positive definite matrix $A(\theta_0) = \lim E n^{-1}(\partial^2 L_n / \partial \theta \partial \theta') \theta_0$ in probability for any sequence θ_n^* such that $\text{plim } \theta_n^* = \theta_0$.

(C) $n^{-1/2}(\partial L_n / \partial \theta) \theta_0 \rightarrow N[0, B(\theta_0)]$, where $B(\theta_0) = \lim E n^{-1}(\partial L_n / \partial \theta) \theta_0 \times (\partial L_n / \partial \theta') \theta_0$.

(D) $n^{-1}L_n(\theta)$ converges to a nonstochastic function $L(\theta)$ in probability uniformly in θ in an open neighborhood of θ_0 .

(E) $\text{plim } n^{-1}\partial^2 L_n / \partial \theta \partial \theta'$ exists and is continuous in a neighborhood of θ_0 .

A transformation of the original data matrix $X(n)$ is used to verify the stated conditions. Partition $X(n)$ as $(X_1', X_2', X_3', \dots, X_g', \dots)'$, where X_g is the observed $n_g \times p$ data matrix for group g . The eigenvalues of each $n_g \times n_g$ diagonal block of $R(\rho)$ are $[1 + (n_g - 1)\rho]$ with multiplicity 1 and $(1 - \rho)$ with multiplicity $(n_g - 1)$. Let T_g be a nonsingular $n_g \times n_g$ matrix with all the elements in the first column equal to $n_g^{-1/2}$ and the remaining columns being orthonormal without ρ and Σ involved. Define $Y_g = T_g' X_g'$. Then, all the rows of Y_g are independent. The first row of each Y_g has a normal distribution with

mean zero and covariance matrix $[1 + (n_g - 1)\rho] \Sigma$. The remaining rows also have a normal distribution with zero means and covariance matrix $(1 - \rho) \Sigma$. Denote $(Y_1', Y_2', Y_3', \dots, Y_g', \dots)'$ as $Y(n)$. $Y'(n) = (y_1, y_2, y_3, \dots, y_{n_1}, y_{n_1+1}, y_{n_1+2}, y_{n_1+3}, \dots, y_{n_1+n_2}, y_{n_1+n_2+1}, y_{n_1+n_2+2}, \dots, y_{n_1+n_2+n_3}, \dots, y_{n_1+n_2+n_3+\dots+n_g}, \dots)$. Each y_i is a column vector of order p .

All the rows of Y are independently, but non-identically distributed. Although there are theorems for consistency and asymptotic normality of the MLE from independently but not identically distributed observations (e.g., Bradley & Gart, 1962; Hoadley, 1971; Philippou & Roussas, 1973), Amemiya's theorems (1985) are applied to prove the consistency and asymptotic normality of the MLE's of ρ and Σ for the sake of simplicity.

Let us denote the y_i , whose distribution is $N(0, (1 + (n_g - 1)\rho)\Sigma)$, $g = 1, 2, 3, \dots, G$, as u_g . Let the $G \times p$ matrix U represent the transpose of the collection of all the u_g 's. Note that $(1 + (n_g - 1)\rho)^{-1/2} u_g \sim N(0, \Sigma)$ for every g . The remaining y_i 's are independently, identically, normally distributed with zero means and covariance matrix $(1 - \rho)\Sigma$. Denote the transpose of the collection of all these y_i 's as V , an $(n - G) \times p$ matrix. The i.i.d. property plays an important role in the demonstration of asymptotic properties of the MLE's, $\hat{\rho}$ and $\hat{\Sigma}$. The negative Log-likelihood for Y , or equivalently for U and V , is

$$L_n(Y, \theta) =$$

$$2^{-1} \left\{ n p \ln(2\pi) + p \sum_{g=1}^G \ln[1 + (n_g - 1)\rho] + p(n - G) \ln(1 - \rho) + n \ln|\Sigma| + \sum_{g=1}^G \frac{1}{1 + (n_g - 1)\rho} u_g' \Sigma^{-1} u_g + \frac{1}{1 - \rho} \text{tr} \left\{ V \Sigma^{-1} V' \right\} \right\}. \tag{1}$$

The first and second derivatives of $L_n(\theta)$ are as follows. The derivative with respect to Σ refers only to the p^* distinct elements in the matrix. The product $A \otimes B$ represents the right Kronecker product with typical element $[a_{ij}B]$.

$$\frac{\partial L_n(\theta)}{\partial \sigma} = 2^{-1} K_p^- \left\{ n \text{ Vec } \Sigma^{-1} - \sum_{g=1}^G \frac{1}{1+(n_g-1)\rho} \text{ Vec}(\Sigma^{-1} u_g u_g' \Sigma^{-1}) - \frac{1}{1-\rho} \text{ Vec}(\Sigma^{-1} V' V \Sigma^{-1}) \right\} \quad (2)$$

$$\frac{\partial L_n(\theta)}{\partial \rho} = 2^{-1} \left\{ p \sum_{g=1}^G \frac{n_g-1}{1+(n_g-1)\rho} - p(n-G) \frac{1}{1-\rho} - \sum_{g=1}^G \frac{n_g-1}{[1+(n_g-1)\rho]^2} u_g' \Sigma^{-1} u_g + \frac{1}{(1-\rho)^2} \text{tr} V \Sigma^{-1} V' \right\} \quad (3)$$

$$\begin{aligned} \frac{\partial^2 L_n(\theta)}{\partial \sigma \partial \sigma'} &= 2^{-1} K_p^- \left\{ -n(\Sigma^{-1} \otimes \Sigma^{-1}) \right. \\ &\quad + \sum_{g=1}^G \frac{1}{1+(n_g-1)\rho} (\Sigma^{-1} \otimes \Sigma^{-1} u_g u_g' \Sigma^{-1} + \Sigma^{-1} u_g u_g' \Sigma^{-1} \otimes \Sigma^{-1}) \\ &\quad \left. + \frac{1}{1-\rho} (\Sigma^{-1} \otimes \Sigma^{-1} V' V \Sigma^{-1} + \Sigma^{-1} V' V \Sigma^{-1} \otimes \Sigma^{-1}) \right\} K_p^{-1} \quad (4) \end{aligned}$$

$$\frac{\partial^2 L_n(\theta)}{\partial \sigma \partial \rho} = 2^{-1} K_p \left\{ \sum_{g=1}^G \frac{n_g - 1}{[1 + (n_g - 1)\rho]^2} \text{Vec}(\Sigma^{-1} u_g u_g' \Sigma^{-1}) - \frac{1}{(1-\rho)^2} \text{Vec}(\Sigma^{-1} V' V \Sigma^{-1}) \right\} \quad (5)$$

$$\frac{\partial^2 L_n(\theta)}{\partial \rho^2} = 2^{-1} \left\{ -p \sum_{g=1}^G \left[\frac{n_g - 1}{1 + (n_g - 1)\rho} \right]^2 - p(n - G) \frac{1}{(1-\rho)^2} + 2 \sum_{g=1}^G \frac{(n_g - 1)^2}{[1 + (n_g - 1)\rho]^3} u_g' \Sigma^{-1} u_g + 2 \frac{1}{(1-\rho)^3} \text{tr} \left[V' \Sigma^{-1} V \right] \right\} \quad (6)$$

The definition of matrix K_p is given in the text.

Verification of Condition (A)

Condition (A) is met trivially.

Verification of Condition (B)

It is to be shown that for any sequence $\theta_n^* = (\rho_n^*, \sigma_n^*)$ such that $\text{plim } \theta_n^* = \theta_0$, $n^{-1}(\partial^2 L_n / \partial \theta \partial \theta')$ converges to a finite positive definite matrix $A(\theta_0) = \lim E n^{-1}(\partial^2 L_n / \partial \theta \partial \theta')_{\theta_0}$ in probability.

Recall that $(1 + (n_g - 1)\rho_0)^{-1/2} u_g$'s are i.i.d., so are the rows of V .

Therefore, we have

$$\frac{1}{n - G} V' V \xrightarrow{p} (1 - \rho_0) \Sigma_0 ;$$

$$E \left[\frac{1}{G} \sum_{g=1}^G \frac{1}{1+(n_g-1)\rho_0} u_g u_g' \right] = \Sigma_0; \text{ and}$$

$$E \left[\frac{1}{n-G} V' V \right] = (1-\rho_0) \Sigma_0.$$

It can be shown that under the assumption $|\rho_n^* - \rho_0| < \varepsilon^* < K^{-2}$,

$$\frac{1}{G} \sum_{g=1}^G \frac{1}{1+(n_g-1)\rho_n^*} u_g u_g' \xrightarrow{p} \Sigma_0.$$

The results follow based on the above statements.

$$\text{plim } n^{-1} \left(\frac{\partial^2 L_n}{\partial \sigma \partial \sigma'} \right) \theta_n^* = 2^{-1} K_p^- (\Sigma_0^{-1} \otimes \Sigma_0^{-1}) K_p^-,$$

$$= \lim E n^{-1} \left(\frac{\partial^2 L_n}{\partial \sigma \partial \sigma'} \right) \theta_0.$$

$$n^{-1} \left(\frac{\partial^2 L_n}{\partial \sigma \partial \rho} \right) \theta_n^* = \frac{1}{2n} K_p^- \sum_{k>1}^K \frac{k-1}{[1+(k-1)\rho_n^*]^2} \text{Vec} \left[\Sigma_n^{*-1} \left(\sum_{\{g:n_g=k\}} u_g u_g' \right) \Sigma_n^{*-1} \right]$$

$$- \frac{1}{2n(1-\rho_n^*)^2} K_p^- \text{Vec} (\Sigma_n^{*-1} V' V \Sigma_n^{*-1})$$

$$\xrightarrow{p} 2^{-1} \delta \sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} K_p^- \text{Vec} \Sigma_0^{-1}$$

$$- 2^{-1} (1-\delta) \frac{1}{1-\rho_0} K_p^- \text{Vec} \Sigma_0^{-1}$$

$$= 2^{-1} \left[\delta \sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} - \frac{1-\delta}{1-\rho_0} \right] K_p^- \text{Vec } \Sigma_0^{-1}.$$

$$\begin{aligned} \text{En}^{-1} \left(\frac{\partial^2 L_n}{\partial \sigma \partial \rho} \right) \theta_0 &= \frac{1}{2n} \sum_{k>1}^K m_k \frac{k-1}{1+(k-1)\rho_0} K_p^- \text{Vec } \Sigma_0^{-1} \\ &\quad - \frac{n-G}{2n(1-\rho_0)} K_p^- \text{Vec } \Sigma_0^{-1} \end{aligned}$$

$$\text{So, } \lim \text{En}^{-1} \left(\frac{\partial^2 L_n}{\partial \sigma \partial \rho} \right) \theta_0 = \text{plim } n^{-1} \left(\frac{\partial^2 L_n}{\partial \sigma \partial \rho} \right) \theta_n^*.$$

$$n^{-1} \left(\frac{\partial^2 L_n}{\partial \rho^2} \right) \theta_n^* = -\frac{p}{2n} \sum_{k>1}^K m_k \left[\frac{k-1}{1+(k-1)\rho_n^*} \right]^2 - \frac{n-G}{2n} p \frac{1}{(1-\rho_n^*)^2}$$

$$+ \frac{1}{n} \sum_{k>1}^K \frac{(k-1)^2}{[1+(k-1)\rho_n^*]^3} \text{tr} \left[\left(\left\{ \begin{matrix} \Sigma \\ \{g:n_g=k\} \end{matrix} \right\} u_g u_g' \right) \Sigma_n^{*-1} \right]$$

$$+ \frac{1}{n(1-\rho_n^*)^3} \text{tr } V' V \Sigma_n^{*-1}$$

$$\rightarrow -\frac{1}{2} \delta p \sum_{k>1}^K f_k \left[\frac{k-1}{1+(k-1)\rho_0} \right]^2 - \frac{1}{2} (1-\delta) p \frac{1}{(1-\rho_0)^2}$$

$$+ \delta p \sum_{k>1}^K f_k \left[\frac{k-1}{1+(k-1)\rho_0} \right]^2 + (1-\delta) p \frac{1}{(1-\rho_0)^2}$$

$$= \frac{1}{2} p \left\{ \delta \sum_{k>1}^K f_k \left[\frac{k-1}{1+(k-1)\rho_0} \right]^2 + \frac{1-\delta}{(1-\rho_0)^2} \right\}$$

$$= \lim E n^{-1} (\partial^2 L_n / \partial \rho^2)_{\theta_0}$$

Therefore, $n^{-1}(\partial^2 L_n / \partial \theta \partial \theta')$ converges to $A(\theta_0) = \lim E n^{-1}(\partial^2 L_n / \partial \theta \partial \theta')$ in probability for any sequence θ_n^* such that $\text{plim } \theta_n^* = \theta_0$.
 $A(\theta_0) =$

$$\frac{1}{2} \begin{bmatrix} K_p^{-1} (\Sigma_0^{-1} \otimes \Sigma_0^{-1}) K_p^{-1} \left[\delta \sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} - \frac{1-\delta}{1-\rho_0} \right] K_p^{-1} \text{Vec} \Sigma_0^{-1} \\ \text{Sym.} \quad \left[\delta \sum_{k>1}^K f_k \left(\frac{k-1}{1+(k-1)\rho_0} \right)^2 + \frac{1-\delta}{(1-\rho_0)^2} \right] p \end{bmatrix}$$

The positive definiteness of $A(\theta_0)$ is to be shown next.

$$\text{Let } A(\theta_0) = 2^{-1} \begin{bmatrix} K_p^{-1} & 0 \\ 0 & 1 \end{bmatrix} \bar{A}(\theta_0) \begin{bmatrix} K_p^{-1'} & 0 \\ 0 & 1 \end{bmatrix},$$

$$\bar{A}(\theta_0) = \begin{bmatrix} \Sigma_0^{-1} \otimes \Sigma_0^{-1} \left[\delta \sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} - \frac{1-\delta}{1-\rho_0} \right] \text{Vec} \Sigma_0^{-1} \\ \text{Sym.} \quad \left[\delta \sum_{k>1}^K f_k \left(\frac{k-1}{1+(k-1)\rho_0} \right)^2 + \frac{1-\delta}{(1-\rho_0)^2} \right] p \end{bmatrix}$$

$$= \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix}$$

Since $\bar{A}_{11} = \Sigma_0^{-1} \otimes \Sigma_0^{-1}$ is positive definite, $|\bar{A}(\theta_0)| = |\bar{A}_{11}| |\bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^{-1} \bar{A}_{12}|$ (see, e.g., Graybill, 1983: 184), and $\bar{A}(\theta_0)$ is positive definite if and only if $|\bar{A}(\theta_0)| > 0$ (see, e.g., Basilevsky, 1983: 135).

$$\begin{aligned}
 |\bar{A}(\theta_0)| &= |\bar{A}_{11}| \left| \bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^{-1} \bar{A}_{12} \right| \\
 &= p |\Sigma_0|^{-2p} \left\{ \frac{\delta(1-\delta)}{(1-\rho_0)^2} + 2 \frac{\delta(1-\delta)}{1-\rho_0} \sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} \right. \\
 &\quad \left. + \delta \left[\sum_{k>1}^K f_k \left(\frac{k-1}{1+(k-1)\rho_0} \right)^2 - \left(\sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} \right)^2 \right] \right\} \\
 &> \delta p |\Sigma_0|^{-2p} \left[\sum_{k>1}^K f_k \left(\frac{k-1}{1+(k-1)\rho_0} \right)^2 - \left(\sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} \right)^2 \right] \\
 &= \delta p |\Sigma_0|^{-2p} \sum_{k>1}^K f_k \left[\frac{k-1}{1+(k-1)\rho_0} - \sum_{k>1}^K f_k \frac{k-1}{1+(k-1)\rho_0} \right]^2 \geq 0.
 \end{aligned}$$

So, $|\bar{A}(\theta_0)| > 0$ and $\bar{A}(\theta_0) > 0$. Note that $\begin{bmatrix} K_p^{-1} & 0 \\ 0 & 1 \end{bmatrix}$ is of full column rank $p^* + 1$. Therefore, $A(\theta_0)$ is positive definite.

□□

Verification of Condition (C)

The first derivative of L_n with respect to θ at θ_0 is rewritten before the asymptotic distribution of $n^{-1/2} (\partial L_n / \partial \theta)_{\theta_0}$ is shown.

$$\begin{aligned} \left(\frac{\partial L_n}{\partial \sigma} \right)_{\theta_0} &= 2^{-1} K_p^- \left\{ n \text{Vec} \Sigma_0^{-1} - (\Sigma_0^{-1} \otimes \Sigma_0^{-1}) \text{Vec} \sum_{k>1}^K \sum_{\{g:n_g=k\}} \frac{u_g u_g'}{1+(k-1)\rho_0} \right. \\ &\quad \left. - (\Sigma_0^{-1} \otimes \Sigma_0^{-1})(n-G) \text{Vec} \left[\frac{V'V}{(n-G)(1-\rho_0)} \right] \right\} \\ &= 2^{-1} K_p^- \left\{ n \text{Vec} \Sigma_0^{-1} - G \text{Vec} \Sigma_0^{-1} \right. \\ &\quad \left. - (\Sigma_0^{-1} \otimes \Sigma_0^{-1}) \left[\sum_{k>1}^K m_k \text{Vec} \left(\frac{1}{m_k} \sum_{\{g:n_g=k\}} \frac{u_g u_g'}{1+(k-1)\rho_0} - \Sigma_0 \right) \right] \right. \\ &\quad \left. - (n-G)(\Sigma_0^{-1} \otimes \Sigma_0^{-1}) \text{Vec} \left[\frac{V'V}{(n-G)(1-\rho_0)} - \Sigma_0 \right] - (n-G) \text{Vec} \Sigma_0^{-1} \right\} \\ &= -2^{-1} K_p^- (\Sigma_0^{-1} \otimes \Sigma_0^{-1}) \left\{ \sum_{k>1}^K m_k \text{Vec} \left[\frac{1}{m_k} \sum_{\{g:n_g=k\}} \frac{u_g u_g'}{1+(k-1)\rho_0} - \Sigma_0 \right] \right. \\ &\quad \left. + (n-G) \text{Vec} \left[\frac{V'V}{(n-G)(1-\rho_0)} - \Sigma_0 \right] \right\}. \\ \left(\frac{\partial L_n}{\partial \rho} \right)_{\theta_0} &= 2^{-1} \left\{ p \sum_{g=1}^G \frac{n_g - 1}{1 + (n_g - 1)\rho_0} - p(n-G) \frac{1}{1 - \rho_0} \right\} \end{aligned}$$

$$\begin{aligned}
 & - \sum_{k > 1}^K m_k \frac{k-1}{1+(k-1)\rho_0} \text{tr} \left[\left(\frac{1}{m_k} \begin{matrix} \Sigma \\ \{g:n_g = k\} \end{matrix} \frac{u_g u_g'}{1+(k-1)\rho_0} - \Sigma_0 \right) \Sigma_0^{-1} \right] \\
 & - p \sum_{k > 1}^K m_k \frac{k-1}{1+(k-1)\rho_0} \\
 & + \frac{n-G}{1-\rho_0} \text{tr} \left[\left(\frac{V'V}{(n-G)(1-\rho_0)} - \Sigma_0 \right) \Sigma_0^{-1} \right] + p(n-G) \frac{1}{1-\rho_0} \left. \right\} \\
 = & -2^{-1} \text{Vec}' \Sigma_0^{-1} \left\{ \sum_{k > 1}^K m_k \frac{k-1}{1+(k-1)\rho_0} \text{Vec} \left[\frac{1}{m_k} \begin{matrix} \Sigma \\ \{g:n_g = k\} \end{matrix} \frac{u_g u_g'}{1+(k-1)\rho_0} - \Sigma_0 \right] \right. \\
 & \left. - \frac{n-G}{1-\rho_0} \text{Vec} \left[\frac{V'V}{(n-G)(1-\rho_0)} - \Sigma_0 \right] \right\}.
 \end{aligned}$$

$$\frac{1}{\sqrt{n}} \left(\frac{\partial L_n}{\partial \theta} \right)_{\theta_0} = \frac{1}{\sqrt{n}} \begin{bmatrix} \frac{\partial L_n}{\partial \sigma} \\ \frac{\partial L_n}{\partial \rho} \end{bmatrix}_{\theta_0} = -\frac{1}{2\sqrt{n}} \Delta \quad D, \quad \text{Where}$$

$\Delta =$

$$\left[\begin{array}{c} K_p^-(\Sigma_0^{-1} \otimes \Sigma_0^{-1}) \left[(\sqrt{m_2} \ \sqrt{m_3} \ \dots \sqrt{m_K} \ \sqrt{n-G}) \otimes I_{p^2} \right] \\ \text{Vec}' \Sigma_0^{-1} \left[\left(\sqrt{m_2} \frac{2-1}{1+\rho_0} \ \sqrt{m_3} \frac{3-1}{1+2\rho_0} \ \dots \sqrt{m_K} \frac{K-1}{1+(K-1)\rho_0} \ - \frac{\sqrt{n-G}}{1-\rho_0} \right) \otimes I_{p^2} \right] \end{array} \right],$$

and,

$$D = \left[\begin{array}{c} \sqrt{m_2} \text{Vec} \left[\frac{1}{m_2} \quad \Sigma_{\{g:n_g=2\}} \frac{u_g u_g'}{1+(2-1)\rho_0} - \Sigma_0 \right] \\ \sqrt{m_3} \text{Vec} \left[\frac{1}{m_3} \quad \Sigma_{\{g:n_g=3\}} \frac{u_g u_g'}{1+(3-1)\rho_0} - \Sigma_0 \right] \\ \vdots \\ \sqrt{m_K} \text{Vec} \left[\frac{1}{m_K} \quad \Sigma_{\{g:n_g=K\}} \frac{u_g u_g'}{1+(K-1)\rho_0} - \Sigma_0 \right] \\ \sqrt{n-G} \text{Vec} \left[\frac{V'V}{(n-G)(1-\rho_0)} - \Sigma_0 \right] \end{array} \right]$$

All the u_g 's and rows of V are mutually independent. For every k , $\sqrt{m_k} \text{Vec}$

$$\left[\frac{1}{m_k} \Sigma_{\{g:n_g=k\}} \frac{u_g u_g'}{1+(k-1)\rho_0} - \Sigma_0 \right] \rightarrow N[0, 2M_p(\Sigma_0 \otimes \Sigma_0)], \text{ where } M_p =$$

$K_p K_p^-$ is a $p^2 \times p^2$ symmetric idempotent matrix (see, e.g., Browne, 1974). And

$\sqrt{n - G} \text{ Vec} \left[\frac{V'V}{(n - G)(1 - \rho_0)} - \Sigma_0 \right]$ has the same asymptotic distribution.

The joint asymptotic distribution of D is $N[0, I \otimes 2M_p(\Sigma_0 \otimes \Sigma_0)]$.

It can be shown through matrix algebra that $\frac{1}{\sqrt{n}} \left(\frac{\partial L_n}{\partial \theta} \right)_{\theta_0}$ is asymptotically distributed as $N[0, B(\theta_0)]$, and $B(\theta_0)$ is identical to $A(\theta_0)$.

□□

Verification of Condition (D)

Theorem 4.2.2 of Amemiya (1985) is used to show condition (D) on uniform convergence of the likelihood function.

Define $\tilde{\Theta}$, a compact subset of q -dimensional Euclidean space, as

$\tilde{\Theta} = \{ \theta = (\Sigma, \rho) : \lambda_2 I_p \leq \Sigma \leq \lambda_1 I_p, -\varepsilon \leq \rho \leq \gamma, \text{ for } 0 < \lambda_2 < \lambda_1, \text{ and some } \varepsilon < K^{-1} \text{ and } \gamma < 1 \}$. Note $\Theta \subset \tilde{\Theta}$. Let $h_i(y_i, \theta) = \ln f_i(y_i, \theta) - E_{\theta_0} \ln f_i(y_i, \theta)$, where $f_i(y_i, \theta)$ is the normal density function for y_i . Note $E_{\theta_0} h_i(y_i, \theta) = 0$. It is to be shown that for every i , there exists some $\delta > 0$

such that $E \sup_{\theta \in \tilde{\Theta}} |h_i(y_i, \theta)|^{1 + \delta} < \infty$. Then, $n^{-1} \sum_{i=1}^n h_i(y_i, \theta)$ converges to

0 in probability uniformly in $\theta \in \tilde{\Theta}$ according to Theorem 4.2.2 of Amemiya (1985). So $n^{-1} L_n(\theta)$ converges to a nonstochastic function $L(\theta)$ in probability uniformly in θ in an open neighborhood of θ_0 .

Let $n_0 = 0$.

For $i = \sum_{k=0}^{g-1} n_k + 1$, where $g = 1, 2, \dots, G$,

$$h_i(y_i, \theta) = \frac{1}{1 + (n_g - 1)\rho} \text{tr}\{y_i y_i' - [1 + (n_g - 1)\rho_0] \Sigma_0\} \Sigma^{-1}.$$

Otherwise,

$$h_i(y_i, \theta) = \frac{1}{1 - \rho} \text{tr}\{y_i y_i' - (1 - \rho_0) \Sigma_0\} \Sigma^{-1}.$$

Let us first discuss the case where $i = \sum_{k=0}^{g-1} n_k + 1$, with $g = 1, 2, \dots, G$.

Note $0 < \frac{1}{1 + (n_g - 1)\rho} < K$.

$$\begin{aligned} |h_i(y_i, \theta)| &< K \left| \text{tr}\{y_i y_i' - [1 + (n_g - 1)\rho_0] \Sigma_0\} \Sigma^{-1} \right| \\ &= K \left| y_i' \Sigma^{-1} y_i - [1 + (n_g - 1)\rho_0] \text{tr}(\Sigma_0 \Sigma^{-1}) \right|. \end{aligned}$$

If $y_i' \Sigma^{-1} y_i \geq [1 + (n_g - 1)\rho_0] \text{tr}(\Sigma_0 \Sigma^{-1})$,

$$\begin{aligned} |h_i(y_i, \theta)| &< K y_i' \Sigma^{-1} y_i \\ &= K (\Sigma_0^{-1/2} y_i)' \Sigma_0^{1/2} \Sigma^{-1} \Sigma_0^{1/2} (\Sigma_0^{-1/2} y_i). \end{aligned}$$

Let $z_i = \Sigma_0^{-1/2} y_i$, $z_i \sim N(0, [1 + (n_g - 1)\rho_0] I_p)$.

Because $\lambda_2 I_p \leq \Sigma \leq \lambda_1 I_p$, we have $\lambda(\Sigma_0^{1/2} \Sigma^{-1} \Sigma_0^{1/2}) \leq \lambda_2^{-1} \lambda_{\max}(\Sigma_0) = M$, say.

So, $\Sigma_0^{1/2} \Sigma^{-1} \Sigma_0^{1/2} \leq M I_p$.

Therefore, $|h_i(y_i, \theta)| < K z_i' (M I_p) z_i$
 $= K M (z_i' z_i)$.

$$\text{Sup}_{\theta \in \tilde{\Theta}} |h_i(y_i, \theta)|^2 < K^2 M^2 (z_i' z_i)^2.$$

$$E \sup_{\theta \in \tilde{\Theta}} |h_i(y_i, \theta)|^2 < K^2 M^2 E(z_i' z_i)^2 \\ \leq 3p^2 K^2 M^2 [1 + (n_g - 1)\rho_0]^2 < \infty \text{ for every } g.$$

If $y_i' \Sigma^{-1} y_i < [1 + (n_g - 1)\rho_0] \text{tr}(\Sigma_0 \Sigma^{-1})$,

$$|h_i(y_i, \theta)| < K[1 + (n_g - 1)\rho_0] \text{tr}(\Sigma_0 \Sigma^{-1}) \\ \leq pKM[1 + (K - 1)\rho_0] < \infty \text{ for every } g.$$

Therefore, when $i = \sum_{k=0}^{g-1} n_k + 1$, with $g = 1, 2, \dots, G$, there exists some $\delta > 0$

such that $E \sup_{\theta \in \tilde{\Theta}} |h_i(y_i, \theta)|^{1+\delta} < \infty$.

A similar relationship can be derived for other $h_i(y_i, \theta)$'s, noting

$$\frac{K}{K+1} < \frac{1}{1-\rho} \leq \frac{1}{1-\gamma} < \infty. \text{ Details are given in Weng (1990). Therefore, for}$$

every i , there exists some $\delta > 0$ such that $E \sup_{\theta \in \tilde{\Theta}} |h_i(y_i, \theta)|^{1+\delta} < \infty$. This

completes the proof. □□

Verification of Condition (E)

$$\text{plim} \frac{1}{n} \frac{\partial^2 L_n}{\partial \sigma \partial \sigma'} = \text{plim} \frac{1}{2n} K_p^- \left\{ -n(\Sigma^{-1} \otimes \Sigma^{-1}) \right.$$

$$+ \sum_{k>1} \frac{K}{1+(k-1)\rho} \left[\Sigma^{-1} \otimes \Sigma^{-1} \left(\sum_{\{g:n_g=k\}} u_g u_g' \right) \Sigma^{-1} + \Sigma^{-1} \left(\sum_{\{g:n_g=k\}} u_g u_g' \right) \Sigma^{-1} \otimes \Sigma^{-1} \right] \\ + \frac{1}{1-\rho} (\Sigma^{-1} \otimes \Sigma^{-1} V' V \Sigma^{-1} + \Sigma^{-1} V' V \Sigma^{-1} \otimes \Sigma^{-1}) \left. \right\} K_p^{-1}$$

$$= 2^{-1} K_p^{-1} \left\{ -(\Sigma^{-1} \otimes \Sigma^{-1}) \right.$$

$$+ \delta \sum_{k>1} f_k \frac{1+(k-1)\rho_0}{1+(k-1)\rho} (\Sigma^{-1} \otimes \Sigma^{-1} \Sigma_0 \Sigma^{-1} + \Sigma^{-1} \Sigma_0 \Sigma^{-1} \otimes \Sigma^{-1})$$

$$\left. + (1-\delta) \frac{1-\rho_0}{1-\rho} (\Sigma^{-1} \otimes \Sigma^{-1} \Sigma_0 \Sigma^{-1} + \Sigma^{-1} \Sigma_0 \Sigma^{-1} \otimes \Sigma^{-1}) \right\} K_p^{-1}.$$

$$\text{plim} \frac{1}{n} \frac{\partial^2 L_n}{\partial \sigma \partial \rho} = \text{plim} \frac{1}{2n} K_p^{-1} \left\{ \sum_{k>1} \frac{k-1}{[1+(k-1)\rho]^2} \text{Vec} \left[\Sigma^{-1} \left(\sum_{\{g:n_g=k\}} u_g u_g' \right) \Sigma^{-1} \right] \right. \\ \left. - \frac{1}{(1-\rho)^2} \text{Vec}(\Sigma^{-1} V' V \Sigma^{-1}) \right\}$$

$$= 2^{-1} \left\{ \delta \sum_{k>1}^K f_k \frac{(k-1)[1+(k-1)\rho_0]}{[1+(k-1)\rho]^2} - (1-\delta) \frac{1-\rho_0}{(1-\rho)^2} \right\} \mathbf{K}_p^- \text{Vec}(\Sigma^{-1} \Sigma_0 \Sigma^{-1}).$$

$$\text{plim} \frac{1}{n} \frac{\partial^2 L_n}{\partial \rho^2} = \text{plim} \frac{1}{2n} \left\{ -p \sum_{k>1}^K m_k \left[\frac{k-1}{1+(k-1)\rho} \right]^2 - p(n-G) \frac{1}{(1-\rho)^2} \right.$$

$$+ 2 \sum_{k>1}^K \frac{(k-1)^2}{[1+(k-1)\rho]^3} \text{tr} \left(\sum_{\{g:n_g=k\}} u_g u_g' \right) \Sigma^{-1}$$

$$\left. + 2 \frac{1}{(1-\rho)^3} \text{tr} \mathbf{V}' \mathbf{V} \Sigma^{-1} \right\}$$

$$= 2^{-1} \left\{ -\delta p \sum_{k>1}^K f_k \left[\frac{k-1}{1+(k-1)\rho} \right]^2 - (1-\delta) p \frac{1}{(1-\rho)^2} \right\}$$

$$+ \left\{ \delta \sum_{k>1}^K f_k \frac{(k-1)^2 [1+(k-1)\rho_0]}{[1+(k-1)\rho]^3} + (1-\delta) \frac{1-\rho_0}{(1-\rho)^3} \right\} \text{tr} \Sigma_0 \Sigma^{-1}.$$

Since $\text{plim} n^{-1} \partial^2 L_n / \partial \theta \partial \theta'$ exists and is continuous for every $\theta \in \Theta$, Condition

(E) is satisfied. □□

大陸臺商工廠待獲利期間之分析： 存活模型之應用*

林祖嘉** 李紀珠***

壹、緒言

廠商自投資設廠到生產獲利，需要一段時間，此待獲利期間長短與風險有關。一般而言，期間越長者，風險越大。因此，廠商投資之初，必然會對開始獲利期間的長短做詳細的評估。然而，該段期間不易掌握，因為有許多因素會影響該期間，例如投資額大小、產業別、員工人數、投資環境、法令規章等。所以，如何去掌握該段期間，對每一有意投資的廠商，都是十分重要的。

1987年底政府開放人民赴大陸探親以來，開始出現臺商大陸投資的熱潮。依投審會資料顯示，至1994年底為止，赴大陸投資廠商超過12000家，投資金額超過35億美元。若依大陸方面資料，廠商家數與投資金額更高出許多。雖然大多數學者相信造成臺商赴大陸投資熱潮的主因在於賺錢的機會較多，¹但不可否認的，赴大陸投資的確比其他

* 作者們感謝兩位審稿人寶貴意見，當然，文中若還有任何錯誤，都應由作者們自行負責。

** 政治大學經濟系教授

*** 政治大學經濟系教授

東南亞地區投資有更多的風險。畢竟大陸的法令規章、基礎建設、人治色彩等諸多因素，都較東南亞其他國家為差。因此，赴大陸投資的臺商一方面不願意冒然投下大筆資金，一方面也希望已投下去的資金能及早回收。然而，投資的待獲利期間固然主要決定於廠商經營能力的高低之外，但也往往受到經濟環境、投資環境、及其他許多外在因素的影響。因此，如何正確的掌握赴大陸投資的待獲利期間，可說是每一位有意於赴大陸投資的廠商所必須首先面對的課題。

由於待獲利期間具有不確定因素，因此一個存活分析模型 (survival analysis model) 就是最能掌握期間分析的利器。本研究擬採用存活模型對於廠商待獲利期間加以分析，一方面我們要探究影響赴大陸廠商投資待獲利期間的可能因素為何？另一方面我們也想了解另外一個更重要的課題，即：待獲利期間是否具有時間依賴性 (time dependence)？因為，當一個廠商在投資一段期間仍然無法獲利時，他面臨的問題是：無法獲利期間將會立即結束，或是未來可能獲利的機率會愈來愈小？不同的答案顯然將帶給該投資者非常不一樣的決策！

存活模型早期多應用於生物醫學及工業工程領域，如 Kao (1958, 1959), Pike (1966), Cox (1972), Crowley and Hu (1977)，及 Dhillon (1979) 等。晚期則有學者將其應用在經濟及商學領域，如有關探討失業問題者，有 Lancaster (1979, 1985), Nickell (1979), Abraham and Farber (1987)，林祖嘉 (1991)，張清溪與駱明慶 (1991) 等，關於遷徙問題探討者，有 Morrison (1967)、Ginsberg (1971)，林祖嘉 (1990) 等。此外，在產業及金融方面的應用亦相當廣泛，如 Lane, et al. (1986), Green and Shoven (1986), Giliberto and Thibodeau (1989), Schwartz and Torous (1989)，林祖嘉與方世調 (1992)，李紀珠 (1992, 1993a, 1993b, 1994)，李紀珠與蔡龍學 (1992)、林惠玲 (1993) 等。

廠商投資的主要目的在於獲取利潤，殆無疑問。獲利能力除決定於廠商的經營能力以外，一般而言尚與規模大小、投資型態、產業種

類、總體經濟環境都有關係。²自投資設廠之始至獲利為止的期間長短，與獲利能力有關，因此也會與上述影響因素有關，但是由於臺商赴大陸投資的時間並不很久，因此有關於大陸臺商經營狀況相關文獻並不多見，高希均等 (1992, 1995) 及高長及嚴宗大 (1992) 有較為完整的相關探討。高希均等 (1992) 之間卷調查，除了瞭解個別廠商獲利情況外，並曾利用迴歸方法，進一步探討決定廠商獲利之因素。高希均等人發現投資金額、員工人數、廠房面積、廠齡、成本降低能力、銷售額、及產業別等因素都與廠商的獲利能力有關。³

本文利用高希均等 (1992) 問卷資料，進一步探討影響大陸臺商待獲利期間之因素。除了瞭解那些因素決定廠商待獲利期間長短之外，我們還要進一步探究該期間是否有時間相依性 (time dependence) 的存在。亦即，隨著廠商待獲利期間越長情況下，究竟是表示廠商將更不容易獲利，或更容易即將獲利？這個問題對於瞭解廠商赴大陸投資風險有很大助益。如果，隨著待獲利期間愈長，顯示出以後將更不容易獲利，則這時廠商應可能考慮儘速結束投資；反之，如果待獲利期間愈長，表示即將獲利的機會愈大，則廠商顯然應繼續投資，這時也表示未來將面臨投資無法獲利的風險較小。⁴

本文第二節先討論存活模型與期間分析的基本架構，並建立計量模型。第三節先說明本研究使用的變數及資料的基本性質，然後分析實證結果。第四節是結論及說明進一步的研究方向。

貳、大陸臺商待獲利期間估計模型

首先我們令臺商在大陸設廠至開始獲利之期間 (或稱待獲利期間) 為一隨機變數 T ，並假定所有廠商的待獲利期間面對相同的分配及相同可能參數，⁵ 並令其機率密度函數 (pdf) 及累積分配函數 (cdf) 分別為 $f(T)$, $F(T)$ 。因此，臺商在某一時期 t 尚未獲利的累積機率函數 $S(t)$

爲

$$S(t) \equiv p(T > t) \equiv 1 - F(t). \quad (1)$$

而臺商在 t 期前雖尚未獲利，但在 t 期開始獲利之獲利狀態條件轉換函數 (或稱轉換率) $h(t)$ 爲

$$h(t) = \lim_{\Delta t \rightarrow 0^+} \frac{p(t < T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}. \quad (2)$$

根據式 (1)、(2) 兩式可推導出 $h(t)$ 與 $S(t)$ 具下列關係

$$h(t) = \frac{-d \log S(t)}{dt}, \quad (3)$$

$$f(t) = h(t) \exp\left(-\int_0^t h(\mu) d\mu\right). \quad (4)$$

如果廠商的獲利狀態轉換率與廠商設廠期間無必然關連，則其 $h(t)$ 將是一個與廠商設廠期間 (t) 無關的常數，即 $h(t) = \lambda$ ，則由式 (4) 知，其 pdf 將爲：

$$f(t) = \lambda \exp(-\lambda t). \quad (5)$$

此時隨機變數 T 是呈 exponential 分配。但本文認爲較合理的假設，爲並不事先排除獲利狀態之轉換率具相關特質的可能，因此，我們將令待獲利期間變數 T 爲一較 exponential 分配更一般化的 Weibull 分配。在 Weibull 分配下，發生獲利狀態轉換之條件轉換率爲

$$h(t) = \lambda \alpha (\lambda t)^{\alpha-1}, \quad (6)$$

式中 λ 爲 exponential 分配下之獲利狀態轉換之轉換函數，而 α 參數與 Weibull 分配下之規模參數 (scale parameter) σ ，具下列關係 $\sigma = \sigma^{-1}$ 。再由 (6) 式得知，規模參數 σ 與獲利狀態轉換函數 $h(t)$ 之期間相依性 ($dh(t)/dt$) (time dependence)，具下列關係：即

$$\begin{cases} \frac{dh(t)}{dt} \geq 0, & \text{當 } \sigma \leq 1 \text{ 時,} & (7a) \\ \frac{dh(t)}{dt} < 0, & \sigma > 1 \end{cases}$$

亦即於 Weibull 分配中，若估計得 σ 小於 1，則獲利狀態之轉換率將隨廠商設廠期間之加長而上升，即獲利狀態轉換與設廠期間具正期間相依性。反之，若規模參數 σ 大於 1，則廠商獲利狀態轉換率，將隨廠商設廠期間愈久而愈下降，即獲利狀態轉換率與設廠期間具負期間相依性。而當規模參數 σ 等於 1 時，廠商獲利狀態轉換率與廠商設廠期間長短無關，即其不具期間相依性，且此時 Weibull 分配將退化為 exponential 分配。據此，在廠商待獲利期間呈 Weibull 分配假定下，我們可透過對規模參數 σ 之估計與檢定，來判定獲利狀態轉換率是否具有期間相依性，亦即推估廠商獲利狀態轉換率是否與廠商設廠期間有關，且若有關，則其是呈正相關或負相關。

另外，為更精確描述廠商待獲利期間的特性，本文也將估計一般化 gamma 分配 (generalized gamma distribution, GGD) 下之結果。因為 GGD 是 Weibull 分配的一般化形式，因此，GGD 必定會比 Weibull 有更好的估計結果。⁶

廠商發生獲利狀態轉換率高低，一般除了可能受到廠商設廠期間長短影響外，尚可能受到其它因素 (Z) 的影響，如廠商規模、產業特性、設廠區域、投資型態等因素。若假定此時潛在因素 Z 對廠商獲利狀態轉換率是呈現指數形式的影響，則此時式 (6) 廠商獲利狀態轉換函數可表達成如下形式：

$$h(t; Z) = \lambda \alpha (\lambda t)^{\alpha-1} \exp(Z' \beta), \quad (6')$$

$Z' \equiv (Z_1, Z_2, \dots, Z_s) \equiv s$ 個潛在影響因素，

$\beta \equiv (\beta_1, \beta_2, \dots, \beta_s) \equiv s$ 個迴歸參數。

根據 (4) 式，則此時 T 之 pdf 爲

$$f(t; Z') = \lambda \alpha (\lambda t)^{\alpha-1} \exp(Z' \beta) \exp[-(\lambda t)^\alpha \exp(Z' \beta)] \quad (8)$$

若令 $X = \log T$ ，則 $T = \exp X$ 代入式 (8)，並令 $\beta_0 = -\log \lambda$ ， $\alpha = \alpha^{-1}$ ， $\beta^* = -\sigma \beta$ ，則 $\log T$ 的 pdf 爲

$$f(X) = \frac{1}{\sigma} \exp\left[-\frac{X - \beta_0 - Z' \beta^*}{\sigma} - \exp\left(\frac{X - \beta_0 - Z' \beta^*}{\sigma}\right)\right], -\infty < X < \infty. \quad (9)$$

令 $W = [(X - \beta_0 - Z' \beta^*) / \sigma]$ ，則由 (9) 式可求得 W 爲標準極值分配 (standard extreme value distribution) 的隨機變數，其 pdf 爲

$$f(W) = \exp(W - \exp W), -\infty < W < \infty. \quad (10)$$

且 $\log T$ 可表達成：

$$X = \log T = \beta_0 + Z' \beta^* + \sigma W. \quad (11)$$

比較式 (9)、(10) 兩式可得知， $f(\log T) = \sigma^{-1} f(W)$ 。

此外，(11) 式亦可改寫爲如下：

$$T = \exp(\beta_0 + Z' \beta^* + \sigma W) = \exp(\beta_0 + Z' \beta^*) T_0. \quad (12)$$

$T_0 = e^{\sigma W}$ ，表示基準 (baseline) 狀態下廠商待獲利期間。

由 (12) 式知，潛在影響因素對廠商待獲利期間之影響具倍數作用 (multiplication)，而具有此類特性之期間模型，文獻上又稱之爲加速生命模型 (accelerated life model)，參見 Cox and Oakes (1983)。

關於其它潛在影響因素 (Z)，本文將就廠商員工人數 (LABOR)、大陸臺商使用原料來源 (MAT)、成本下降比例 (COST)、投資形式 (TYPE)、投資地區 (AREA) 及產業類別 (IND) 等可能因素進行分析。⁷ 因此本文的 Z 函數爲：

$$Z = Z'(LABOR, MAT, MACH, COST, TYPE, AREA, IND). \quad (13)$$

在處理期間資料 (duration data) 時，一般人常會遇到一種困擾，即在觀察期間結束時，某些樣本仍停留在原狀態，並未產生狀態轉換現象，因此，部份樣本於原狀態所停留的完整存續期間，並無法真正被觀察到，亦即在觀察樣本中有部份樣本發生截斷 (censored) 的現象。以本文所觀察的大陸臺商待獲利期間分析為例，當觀察期間結束時 (1991年10月)，許多大陸臺商尚未開始獲利，因此，其完整的待獲利期間為多長不得而知，此即為觀察期間不完全或遭截斷的樣本，且屬於右方截斷 (right censored) 的樣本。⁸

若觀察樣本中存在有被截斷的樣本，雖然可提供之訊息不若未被截斷樣本來的精確，但仍可提供部份訊息，因此，我們需先將所收集到大陸臺商樣本中，同時涵蓋未截斷以及右方截斷的樣本，提供不同訊息的現象加以處理。

令 L_i 表第 i 個觀察樣本 (即第 i 家大陸臺商) 的截斷時間 (censoring time)，而 T_i 表第 i 家大陸臺商實際開始獲利期間，因此只有當 $T_i \leq L_i$ ，該廠商實際待獲利的完整期間才會被觀察到，故我們令資料中第 i 個廠商的待獲利對數期間 ($\log t_i$) 為：

$$\log t_i = \text{Min}(\log T_i, \log L_i),$$

$$\text{且 } \delta_i = \begin{cases} 1, & \log T_i \leq \log L_i \\ 0, & \log T_i > \log L_i \end{cases} \quad i = 1, 2, \dots, n. \quad (14)$$

即 $\delta_i = 0$ ，表示待獲利期間有被截斷現象；反之則無，令 $\delta_i = 1$ 。而當資料被右方截斷時

$$\text{Pr}(\log t_i = \log L_i) = \text{Pr}(\log T_i > \log L_i) = S(\log L_i) = S(\log t_i). \quad (15)$$

若資料未遭截斷 (即 $\log t_i < \log L_i$)，則其機率密度函數為⁹

$$\text{Pr}(\log t_i, \delta_i = 1) = f(\log t_i), \quad t_i < L_i. \quad (16)$$

綜合 (14) 與 (15) 式得

$$Pr(\log t_i, \delta_i) = f(\log t_i)^{\delta_i} S(\log t_i)^{1-\delta_i}, i = 1, 2, \dots, n. \quad (17)$$

若每一個觀察值的 (t_i, δ_i) 皆獨立，則涵蓋全部右方截斷及未遭截斷樣本之概似函數為：

$$L(\beta, \sigma) = \prod_{i=1}^n f(\log t_i)^{\delta_i} S(\log t_i)^{1-\delta_i}. \quad (18)$$

因 $W_u = (\log T_i - \beta_0 - Z' \beta^*) / \sigma$ ，故臺商待獲利期間之對數最大概似函數可寫為：

$$\text{Max } \log L(\beta, \sigma) = \sum_{i=1}^n [\delta_i \log \sigma^{-1} f(W_i) + (1 + \delta_i) \log S(W_i)] \quad (19)$$

將 Z' 向量代入 (11) 式，並將 (11) 式之機率分配代入對數最大概似函數 (19) 式，利用 Newton--Raphson 反覆求解，即可求得最適之 $\hat{\beta}$ 及 $\hat{\sigma}$ ，從而可再進一步求得影響大陸臺商待獲利期間之最適估計 $\log T$ 及對應之獲利狀態轉換函數 \hat{h} 。

參、實證分析

本文實証資料取自高希均等 (1992) 對大陸臺商工廠營運調查問卷結果，其回收有效樣本為 319 家，就目前國內研究臺商投資相關調查而言，為資料較完整的一次調查。該問卷調查樣本不論在產業分佈或地域分佈上，皆與經濟部投審會登記的 2503 家母體樣本分佈相當接近，¹⁰ 因此樣本有相當代表性，經去除解釋變數資料不完整之樣本後，計得 60 家廠商可供本研究使用。¹¹

本研究使用之變數說明如下：

DUR：待獲利期間，依廠商填答之待獲利期間為準。若尚未獲利，則待獲利期間以問卷日期減去設廠投資日期計算。

AREA：投資地區，若投資地點為城市，¹² 則 AREA=1；否則 AREA=0。

COST：成本降低率，與臺灣母公司相比，大陸工廠生產成本減少的比例。

IND：產業型態，若為資本密集型產業，則 IND=1；否則 IND=0。¹³

LABOR：員工人數。

MAT：原料來自臺灣比例，即大陸臺商工廠使用原料購自臺灣的比例。

MACH：機器來自臺灣比例，即臺商工廠使用機器購自臺灣的比例。

TYPE：投資型態，若為獨資，則 TYPE=1；否則 TYPE=0。

表 1：大陸臺商工廠基本統計量

	平均數	標準差
已獲利廠商比例 (%)	60.00	49.40
平均待獲利期間 (DUR) (月)	13.45	8.17
員工人數 (LABOR) (人)	82.87	95.97
原料來自臺灣比例 (MAT) (%)	56.75	37.93
機器來自臺灣比例 (MACH) (%)	33.39	25.52
成本降低率 (COST) (%)	24.83	16.56
投資型態 (TYPE) (%)	61.67	49.03
投資地區 (AREA) (%)	38.33	49.03
產業型態 (IND) (%)	26.67	44.59

表 2：大陸臺商待獲利期間分析

被解釋變數 (log DUR)			
解釋變數	Exponential	Weibull	Generalized Gamma
INTERCEPT	3.8241 (0.0001)**	3.4646 (0.0001)**	3.7008 (0.0001)**
LABOR	-0.0006 (0.7329)	-0.0002 (0.8369)	-0.0008 (0.0767)*
MAT	-0.0052 (0.3420)	-0.0037 (0.1477)	-0.0047 (0.0424)**
MACH	-0.0118 (0.0854)*	-0.0113 (0.0004)**	-0.0120 (0.0001)**
COST	-0.0104 (0.3043)	-0.0081 (0.0825)*	-0.0071 (0.0515)*
TYPE	0.3663 (0.3363)	0.3173 (0.0668)*	0.5260 (0.0001)**
AREA	0.2539 (0.4790)	0.3553 (0.0340)**	0.4185 (0.0030)**
IND	-0.1867 (0.6555)	-0.2547 (0.1730)	-0.3130 (0.0069)**
SCALE	1	0.4474	0.0183
SHAPE	-	(0.0615) ^a	(0.0085) ^a
LLV	1	1	32.6725
	-	-	(15.1221) ^a
NO.of Obs.	60	60	60

說明：

1. 括弧內為P 值 (P value)。
2. *,**分別表該係數滿足10%和5%的顯著水準。
3. LLV 為 log likelihood Value。
4. 有 a 者，為 standard error。

經對樣本初步的觀察，參見表 1，在 60 家廠商中，已開始獲利廠商有 36 家，平均待獲利期間 (DUR) 為 13.45 個月，比高長等 (1993) 調查的 14.04 個月略短。¹⁴ 廠商平均雇用勞工人數為 82.87 人，原料平均有 56.75% 來自臺灣，機械設備平均亦有 33.39% 來自臺灣。而就觀察樣本中的廠商，在大陸設廠平均成本下降率達 24.83%，此與目前文獻中所提及比率相當接近，也可以初步解釋為什麼會有這麼多廠商前往大陸投資。而樣本中廠商有 61.67% 屬獨資，38.33% 為合資或合作型態經營。其中 38.33% 於城市地區投資，61.67% 則在都市以外地區投資。最後，這些廠商有 26.67% 屬於資本密集型產業，大部分 (73.33%) 仍屬於勞力密集型產業。

將 60 家大陸臺商資料放入前節的存活模型當中，再利用 SAS 軟體，對本文 (19) 式加以估計，迴歸估計結果列於表 2。表 2 顯示 exponential 分配的估計結果並不理想，但在 Weibull 分配時，估計結果則有顯著改善，不但 LLV 明顯上升，且估計係數亦有多個顯著。由於本文主要目的之一在分析待獲利期間相依性的特質，因此此處我們針對 Weibull 分配的結果加以說明。

首先是員工人數 (LABOR) 的係數為負 (-0.0002)，表示勞力使用較多的廠商其待獲利期間較短，這與臺商赴大陸投資的主要目的在使用大陸廉價勞工的說法一致。MAT 之係數亦為負 (-0.0037)，表示從臺灣進口原料比例較高之廠商其待獲利期間較短。一般而言，由於大陸地區提供的原料品質及交貨期間都較不穩定，因此許多廠商寧願直接從臺灣進口原料以控制品質與出貨期間，這些廠商通常都是生產能力較佳的廠商，故其待獲利期間較短。MACH 之係數負且顯著 (-0.0113)，其理由與原料來自臺灣相近。成本降低比例 (COST) 扮演獲利與否的重要角色，當成本降低愈多，表示獲利機會愈大，待獲利期間愈短，此處負且顯著 (-0.0081) 的估計結果驗證上述觀點。

投資型態 (TYPE) 亦對待獲利期間有正且顯著的影響 (0.3173)。一般而言，當廠商赴國外獨資設廠時，必須花費較多時間去瞭解當地法令、規章、及交易習慣，因此需要較長時間才能獲利。¹⁵ 至於投資地區 (AREA) 的係數為正且顯著 (0.3553)，表示臺商工廠設在鄉鎮地區反而容易獲利。最後，產業型態 (IND) 之係數為負 (-0.2547)，則表示資本密集型產業的待獲利期間較短。再進一步觀察 Weibull 分配下之規模參數，我們發現其顯著的異於1 (0.4474)，表示大陸臺商獲利狀態之轉換率確實具期間相依性。由於其值顯著小於1，根據本文 (7) 式，顯示臺商獲利狀態之轉換率具正期間相依性，即隨著廠商設廠時間加長，廠商開始獲利的條件機率會上升，且由於其規模參數不但顯著異於1，且其倒數顯著大於2，因此其獲利狀態之轉換率，不但隨廠商設廠時間加長而上升，且呈加速的上升。¹⁶

若我們進一步使用較 Weibull 分配更一般化之 gamma 分配 (generalized gamma distribution, GGD) 進行分析，則發現在 GGD 分配下，規模係數亦顯著異於1 (0.0183)，再一次確定廠商獲利狀態轉換率確實具期間相依性。¹⁷ 同時觀察 GGD 分配下迴歸的結果，也可以發現其各解釋變數的符號方向皆與 Weibull 分配相同，且不但機械來源 (MACH)、成本下降比例 (COST)、投資形式 (TYPE) 及投資地區 (AREA) 對廠商待獲利期間有如在 Weibull 分配下顯著影響外，其員工人數 (LABOR)、原料自臺灣來的比例 (MAT) 以及產業型態 (IND) 等解釋變數，對大陸臺商待獲利期間亦皆有相當顯著的影響。

仔細的讀者也許會發現 LABOR 與 IND 的係數都是負的，似乎有些矛盾，一方面表示勞動多的廠商待獲利期間較短，而另一方面，高資本密集產業的廠商的待獲利期間較短。我們的解釋是：LABOR 代表臺商使用勞動的多少，這可做為廠商規模大小的指標。由於大陸勞工十分廉價，因此當使用勞動愈多時，臺商賺錢的機會就愈大，所以待獲利期間就愈短。另一方面，IND 固然代表資本密集與勞力密集產業，但我們用的是虛擬變數 (而不是用資本勞動比)，所以，此處當

IND 的係數為負時，表示待獲利期間較短，我們的解釋是這些「資本密集型」的產業在大陸可能會更有機會，包含機械、電子、紡織等等。因此，如果把 LABOR 與 IND 兩種變數估計結果合併解釋，就是以大規模 (使用較多勞動) 方式進入大陸的機械、電子、紡織等產業的臺商，他們的待獲利期間會較短，會比大多數進入大陸的傳統臺商產業更短，如製鞋、玩具、成衣等等。我們認為可能的理由是臺商在機械、電子、紡織等產業的生產技術優勢可能更高於傳統的製鞋、玩具、成衣等產業，使得他們在大陸有更好的表現機會 (即待獲利期間更短)。¹⁸

最後，為比較待獲利期間實際分配與不同函數下的估計結果的差異，我們把 PDF、CDF、及尚未獲利期間 S (即 1-CDF) 的實際值與估計值列在表 3。¹⁹ 表 3 顯示 Exponential 與 Weibull 分配下的估計值很相近，但卻與實際值有一些出入。若再看圖 1 與圖 2 的 CDF 與 S 函數，可以更清楚的看到，Exponential 分配與 Weibull 分配的估計結果相當接近，但與實際值分配有明顯差異。

造成這種差異的原因可能有二：第一個原因是廠商在實際填寫問卷時造成的誤差。由於有些廠商可能無法清楚的記住他們的待獲利期間是某一確定期間，例如 5 個月或 10 個月，而只能籠統的說是半年或一年。因此，我們在觀察實際資料時，我們發現半年和一年的 PDF 都很高，分別為 13.3% 及 18.3%。由圖 3 的 PDF 圖形中也可以看到，實際分配值分別在 6 月與 12 月有兩個特別高的值。如果廠商待獲利期間真的是隨機分配，或符合某一分配，則不應在半年或一年的地方有特別高的觀察值，因此我們認為主要的原因在於廠商填答問卷時，所造成的誤差。

另外一個差異則是來自於實際觀察值中有部份被截斷的資料，而在 Exponential 與 Weibull 分配的估計值當中，則已將這些被截斷的觀察值加以修正。因此我們看到 Exponential 和 Weibull 兩個分配的

表 3：待獲利期間之實際與估計分配值

期間 (月)	實際觀察值			EXPONENTIAL估計值			WEIBULL估計值		
	PDF	CDF	S	PDF	CDF	S	PDF	CDF	S
0	1.7	1.7	98.3	0	0	100.0	0	0	100.0
1	0	1.7	98.3	0	0	100.0	0	0	100.0
2	1.7	3.3	96.7	0	0	100.0	0	0	100.0
3	0	3.3	96.7	0	0	100.0	0	0	100.0
4	0	3.3	96.7	0	0	100.0	0	0	100.0
5	3.3	6.7	93.3	1.7	1.7	98.3	1.7	1.7	98.3
6	13.3	20.0	80.0	1.7	3.4	96.6	1.7	3.4	96.6
7	1.7	21.7	78.3	3.3	6.7	93.3	1.7	5.0	95.0
8	5.0	26.7	73.3	1.7	8.4	91.6	1.7	6.7	93.3
9	3.3	30.0	70.0	5.0	13.4	86.6	5.0	11.7	88.3
10	10.0	40.0	60.0	5.0	18.4	81.6	3.3	15.0	85.0
11	1.7	41.7	58.3	5.0	23.4	76.6	5.0	20.0	80.0
12	18.3	60.0	40.0	5.0	28.4	71.6	1.7	21.7	78.3
13	3.3	63.3	36.7	0	28.4	71.6	10.0	31.7	68.3
14	6.7	70.0	30.0	6.7	35.0	65.0	6.7	38.4	61.6
15	3.3	73.3	26.7	8.3	43.3	56.7	6.7	45.0	55.0
16	3.3	76.6	23.4	5.0	48.3	51.7	0	45.0	55.0
17	0	76.6	23.4	6.7	55.0	45.0	5.0	50.0	50.0
18	6.7	83.3	16.7	8.3	63.3	36.7	11.7	61.7	38.3
19	0	83.3	16.7	5.0	68.3	31.7	8.3	70.0	30.0
20	1.7	85.0	15.0	6.7	75.0	25.0	6.7	76.7	23.3
21	0	85.0	15.0	5.0	80.0	20.0	0	76.7	23.3
22	0	85.0	15.0	5.0	85.0	15.0	10.0	86.7	13.3
23	3.3	88.3	11.7	5.0	90.0	10.0	3.3	90.0	10.0
24	3.3	91.7	8.3	3.3	93.3	6.7	0	90.0	10.0
25	1.7	93.3	6.7	0	93.3	6.7	1.7	91.7	8.3
26	0	93.3	6.7	0	93.3	6.7	3.3	95.0	5.0
27	0	93.3	6.7	0	93.3	6.7	1.7	96.7	3.3
28	0	93.3	6.7	1.7	95.0	5.0	0	96.7	3.3
29	0	93.3	6.7	1.7	96.7	3.3	0	96.7	3.3
30	3.3	96.6	3.4	1.7	98.3	1.7	0	96.7	3.3
31	0	96.6	3.4	0	98.3	1.7	0	96.7	3.3
32	1.7	98.3	1.7	0	98.3	1.7	1.7	98.4	1.6
33	0	98.3	1.7	0	98.3	1.7	1.7	100.0	0
34	0	98.3	1.7	0	98.3	1.7	0	100.0	0
35	0	98.3	1.7	0	98.3	1.7	0	100.0	0
36	1.7	100.0	0	1.7	100.0	0	0	100.0	0

圖 1 CDF 分配圖形

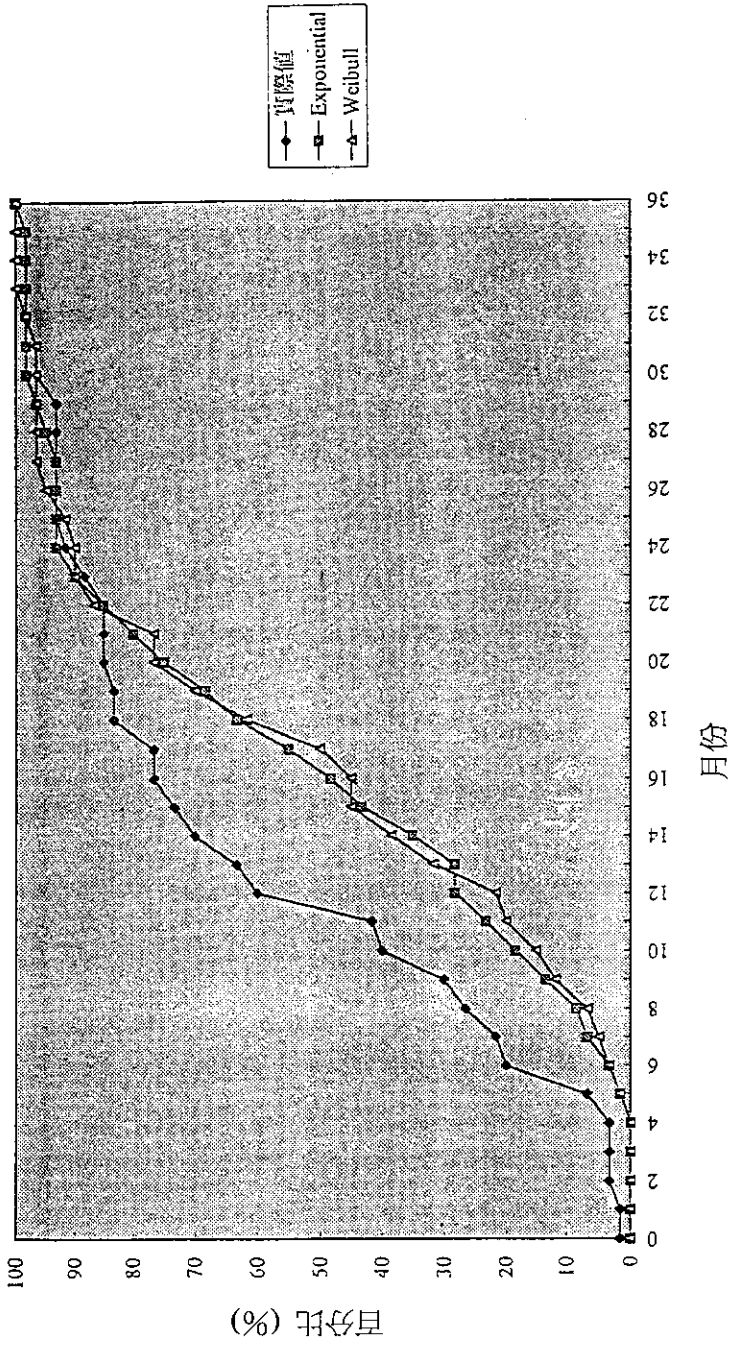


圖 2 S 分配圖形

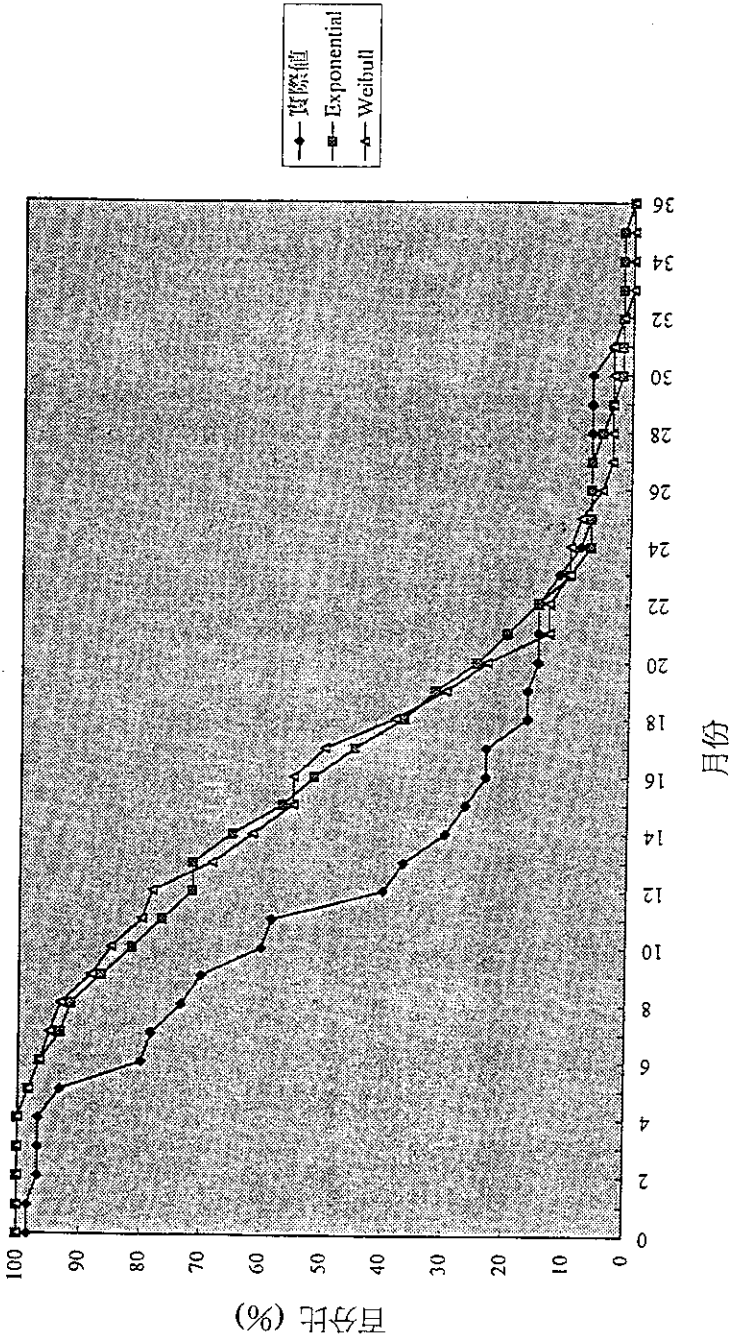
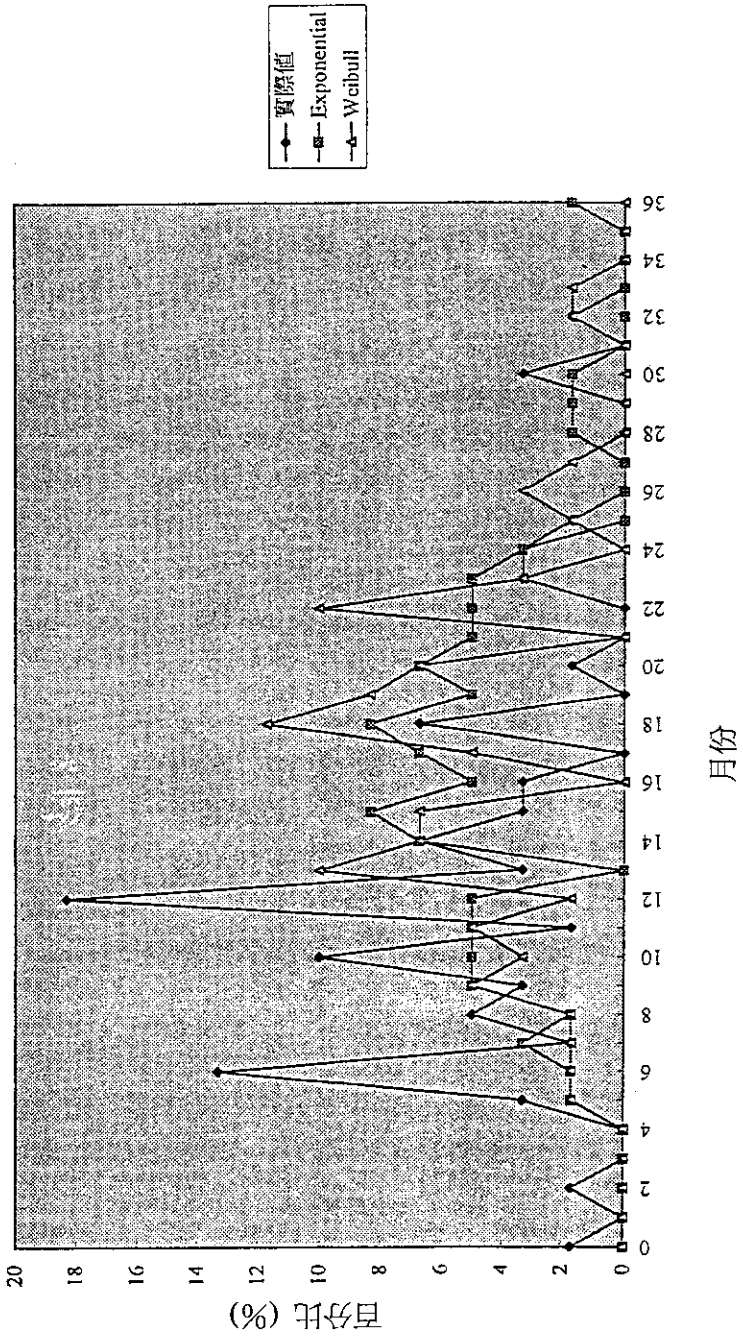


圖 3 PDF 分配圖形



CDF 與 S 都很接近，但卻都與實際值有所差異。同時由於實際值中有被截斷的資料，故若以平均數來看，實際估計值的平均資料待獲利期間只有13.45個月；而 Exponential 與 Weibull 分配估計的平均待獲利期間較長，分別為15.81個月及16.22個月。

肆、結 論

待獲利期間長短與廠商投資風險大小有密切關係。與東南亞國家相比，赴大陸地區投資風險更大，因此，對有意赴大陸投資的廠商而言，如何正確估計待獲利期間，是首先要解決的課題。但由於待獲利期間具有不確定性，並非完全能由廠商掌握，因此，本文建立一個存活模型，來描述廠商待獲利期間的動態變化。我們一方面找出決定待獲利期間的主要因素以外，我們也想知道對一個已設立一段時期但尚未獲利的廠商而言，其獲利機率將減少或增加？這對於瞭解廠商的投資風險有很大的助益。

應用 60 家大陸臺商廠商營運資料，並使用三種不同分配 (exponential, Weibull, GGD) 估計，我們發現廠商員工人數、自臺灣進口原料比例、自臺灣進口機器比例、成本降低比例、投資型態、投資地區、及產業型態等因素都對大陸臺商的投資待獲利期間有顯著的影響。另一方面，模型估計結果發現待獲利期間出現明顯的正時間相依效果，即隨廠商待獲利期間增加，廠商轉為獲利狀態的條件機率會增加，且是呈加速的增加。此一結果顯示，赴大陸投資的獲利風險會隨著待獲利期間的增加而降低。

由於資料限制，使得本文的樣本只有60家廠商，對於比較不同分配之間的差異而言，樣本數目略嫌不足，這可能是本研究主要限制之一。另外，除了本文使用的三種分配以外，是否還有其他種分配更能描述廠商的待獲利期間，值得進一步研究。最後，由於大部分的臺商都是在1988年之後才進入大陸，對大陸市場都相當陌生，待獲利期間

自然受到影響。最近這兩年進入大陸的廠商，他們擁有的訊息要遠超過當時最初進入大陸的廠商，因此他們的待獲利期間是否會有明顯的差異？這是另一個值得深思的課題。

註 釋

- 1 見高希均等 (1992) 第五章的探討。
- 2 見 Samuels and Smith (1968) 及 Mueller (1976) 探討規模大小與訊息多寡對廠商獲利的影響。
- 3 見高希均等 (1992)，第五章，頁185-200。
- 4 在先驗上，我們認為廠商在投資之前，一定認為有獲利的可能，才會去投資設廠。因此，我們假設廠商最後一定會獲利，應該是可以被接受的。
- 5 由於在統計上，我們很難區分每家廠商是面對不同的分配，或相同分配但不同參數，故既有文獻在實證上的一般作法，是假定這些廠商是面對相同分配，故於此我們亦做相同假定。
- 6 然而，GGD 的規模參數 (scale parameter) 與形狀參數 (shape parameter) 並不能用來證實其期間相依性為正或負，故本文在解釋迴歸估計結果時，仍將採用 Weibull 分配來說明。關於 Weibull 與 GGD 的關係，可參見 Kalbfleish and Prentice (1980)，第二章。
- 7 關於這些變數的定義，請參考下一節的說明。
- 8 所謂右方截斷 (right censored) 是指當觀察期間結束時，經濟體系仍於該狀態繼續，而所謂左方截斷 (left censored) 是指在觀察期間開始以前經濟體系即已進入目前狀態。當此二種情況發生時，我們皆無法完整的觀察到經濟體系停留於該狀態的完整期間，因此必須另作處理。有關樣本遭截斷形式及處理方法，可參考 Miller (1981)，頁2-38，或 Lawless (1982)，頁31-44。

- 9 為說明簡便，我們以 $\text{Pr}(\cdot)$ 表達 t 於某種狀況下 (此為 $t < Li$) 的機率密度函數。
- 10 有興趣者，請參閱高希均等 (1992)，第五章說明。
- 11 此處廠商家數較少的主要原因在於原始資料中，關於獲利方面之相關變數的有效樣本較少。見高希均等 (1992)，頁190。
- 12 城市地區主要是指：深圳市、廈門市、東莞市、上海市、南京市、重慶市、廣州市、福州市、珠海市及汕頭市等。
- 13 本文所指資本密集產業包含機械、電子、紡織、造船、精密儀器、基本金屬、塑膠、及運輸工具。此一定義參考 Lin (1994)，頁16。
- 14 因為我們樣本中還有部份廠商尚未獲利，當其待獲利期間完成，則本文整個樣本的待獲利期間就會較長。另外，高長等 (1993) 調查顯示赴大陸投資的廠商中，新加坡、日本、港澳、美國，及其他國家廠商的待獲利期間分別為 13.31, 15.96, 16.20, 18.12 及 19.80 個月，都比臺商工廠要長。見高長 (1993)，表 8-1，頁8-13。
- 15 關於臺商投資型態選擇的探討，可以參見 Lin and Png (1995)。
- 16 Weibull 分配之危險函數 (hazard function) 為：

$$h(t; Z) = Pt^{P-1} X,$$

其中 $X = \exp\left\{-\frac{Z_0\beta_0}{\sigma}\right\}$ ， $\sigma = P^{-1}$ 。所以

$$\frac{dh(t; Z)}{dt} = P(P-1)t^{P-2} X,$$

$$\frac{d^2h(t; Z)}{dt^2} = P(P-1)(P-2)t^{P-3} X.$$

因此：

(1) 當 $0 < P < 1$ ；則 $\frac{dh(t; Z)}{dt} < 0$ ， $\frac{d^2h(t; Z)}{dt^2} > 0$ 。(2) 當 $1 < P < 2$ ；則

$\frac{dh(t; Z)}{dt} > 0$ ， $\frac{d^2h(t; Z)}{dt^2} < 0$ 。(3) 當 $P > 2$ ，則 $\frac{dh(t; Z)}{dt} > 0$ ， $\frac{d^2h(t; Z)}{dt^2} > 0$

- > 0 。而因 SAS 估計樣本的規模參數 (σ) 等於 P 的倒數 ($1/P$)，因此，表 2 中估計出 $\sigma = 0.4474$ ，故 $P > 2$ ，因此，其待獲利期間之轉換率，不但隨廠商存活期間增加而上升，且呈加速上升。
- 17 若我們採用概似值檢定 (likelihood value test)，可以發現 GGD 與 Weibull 之間有顯著差異，也就是說，我們應該採用 GGD 之分配。然而在 GGD 分配下，不易判定其期間相依性的特性，因此，為了強調分析待獲利期間是具有正或負之期間相依性，我們採用 Weibull 的係數來說明。如果，我們研究的重點在於對待獲利期間的推估，則顯然採用 GGD 是比較適當的。
 - 18 我們非常謝謝評審的指正與提醒，讓我們能更清楚表達估計結果。
 - 19 此處的估計值係以各解釋變數觀察值的樣本平均數為基礎，再分別乘上 Exponential 及 Weibull 所估計到的係數，然後就可以得到估計值。但是，我們利用 SAS 程式推估 GGD 下之估計 PDF、CDF、及 S 時，出現一些問題，即在 GGD 分配中，有些值為 0，取對數之後就不存在，因此無法得到 GGD 下的估計 PDF 及 CDF，故表 2 中只列出 Exponential 與 Weibull 分配下的估計值。

參考資料

李紀珠

- 1992 〈貨幣持有期間與貨幣支用傾向期間相依性分析——以臺灣為例〉，《中國經濟學會年會論文集》，頁 131-155。
- 1993a 〈金融機構失敗預測模型—加速失敗時間模型之應用〉，《經濟論文叢刊》21: 355-379。
- 1993b 〈美國景氣循環期間穩定性、相依性及政策影響力——另一種實證分析方法〉，《經濟論文叢刊》21: 99-122。
- 1994 〈臺灣股票市場持股行為之實證研究—民國七十五年至七十九年〉，《中國財務學刊》2: 1-28。

李紀珠、蔡龍學

- 1992 〈臺灣股票上市公司變更交易行為之研究〉，《台灣銀行季刊》43: 227-245。

林祖嘉

- 1990 〈隨機程序下的住屋需求與遷徙〉，《中國經濟學會年會論文集》頁269-317。
- 1991 〈失業期間與工作搜尋模型——臺灣地區大專畢業生之經驗〉，《經濟論文》18: 137-158。

林祖嘉、方世調

- 1992 〈臺北市紡織業與食品工業廠商存活期間之分析〉，《經濟論文》20: 59-91。

林惠玲

- 1993 〈廠商之退出率與存活時間之計量模型——臺灣電力及電子機械器材製造業的驗證〉，《經濟論文叢刊》21: 355-379。

高希均、李誠、林祖嘉

- 1992 《臺灣突破：兩岸經貿追蹤》。臺北：天下文化出版公司。

高希均、林祖嘉、林文玲、許彩雪

- 1995 《臺商經驗：投資大陸的現場報導》。臺北：天下文化出版公司。

高長、季聲國、吳世英

- 1993 《臺商與外商在大陸投資經驗之調查研究——以製造業為例》，經濟部研究計畫報告。臺北：中華經濟研究院。

高長、嚴宗大

- 1992 《兩岸經濟交流之現況及其發展趨勢研究》，行政院陸委會研究計畫報告。臺北：中華經濟研究院。

張清溪、駱明慶

- 1991 〈臺灣勞動力失業期間的研究〉，《勞動市場與勞資關係》，頁75-100。臺北：中央研究院中山人文社會科學研究所。

Abraham K.G. and H.S. Farber

1987 "Job Duration, Seniority, and Earnings," *American Economic Review* 77 (3): 278-297.

Cox, D. R.

1972 "Regression Models and Life-Tables," *Journal of Royal Statistical Society* 34: 187-220.

Cox, D.R., and D.Oakes

1983 *Analysis of Survival Data*, Monographs on Statistics and Applied Probability. Chapman and Hall.

Crowley, J. and M. Hu

1977 "Covariance Analysis of Heart Transplant Data," *Journal of American Statistic Association* 72: 27-36.

Dhillon, B. S.

1979 "A Hazard Model," *IEEE Transactions on Reliability* 26: 150-165.

Giliberto, S. M., and T. G. Thibodeau

1989 "Modeling Conventional Residential Mortgage Refinancing," *Journal of Real Estate Finance and Economics* 2: 285-299.

Ginsberg, R. B.

1971 "Semi-Markov Process and Mobility," *Journal of Mathematical Sociology* 1: 233-262.

Green, J., and J. B. Shoven

1986 "The Effect of Interest Rates on Mortgage Prepayment," *Journal of Money, Credit and Banking* 41-59.

Kalbfleish, J. D., and R. L. Prentice

1980 *The Statistical Analysis of Failure Time Data*. New York: John Wiley and Sons, Inc.

Kao, J. H. K.

1958 "Computer Methods for Estimating Weibull Parameters in Reliability

- Studies," *Transactions of IRE-Reliability and Quality Control* 13: 15-22.
- 1959 "A Graphical Estimation of Mixed Weibull Parameters in Life-Testing Election Tubes," *Technometrics* 1: 389-407.
- Lancaster, T.
- 1979 "Econometric Methods for the Duration of Unemployment," *Econometrica* 47: 939-956.
- 1985 "Generalized Residuals and Heterogeneous Duration Models: with Applications to the Weibull Model," *Journal of Econometrics* 28: 155-169.
- Lane, W. R., S. Leoney and J. W. Wansley
- 1986 "An Application of the Cox Proportional Hazards Models to Bank Failure," *Journal of Banking and Finance* 511-531.
- Lawless, J. F.
- 1982 *Statistical Model and Method for Life Time Data*. New York: John Wiley and Sons, Inc.
- Lin, C. C.
- 1994 "Production Function, Factor Substitutibility, and Direct Foreign Investment: A Case Study of Taiwan," *Asian Economic Journal* 9: 193-203.
- Lin, C. C., and I. Png
- 1995 "Kinship, Control, and Incentives in Cross-Border Investment," Presented at the 7th World Congress of the Econometric Society, Tokyo, August 26-29.
- Miller R. G.
- 1981 *Survival Analysis*. New York: John Wiley and Sons Inc.
- Morrison, R. A.
- 1967 "Duration of Residence and Prospective, Migration : the Evidence of a Stochastic Model," *Demography* 4: 553-561.

Mueller, D.

1976 "Information, Mobility, and Profit," *Kyklos* 29: 419-448.

Nickell, S. J.

1979 "Estimating the Probability of Leaving Unemployment,"
Econometrica 47: 1249-1266.

Pike, M.C.

1966 "A Method of Analysis of Certain Class of Experiments in Carcinogenesis," *Biometrika* 22: 142-161.

Samuels, J. M., and D. J. Smith

1968 "Profits, Variability of Profits and Firm Size," *Economica* 35: 127-139.

Schwartz, E. S., and W. N. Torous

1989 "Prepayment and Valuation of Mortgage-Backed Securities,"
Journal of Finance 94: 375-392.

貨幣數量學說的競爭性：遺傳規畫在 知識發掘上的應用*

陳樹衡** 葉佳炫***

壹、緒論

實証研究工作者在從事資料分析中，所要面臨的問題，雖然是千頭萬緒，但是一個最根本的問題仍然是：如何在現有的資料中去做好知識發掘 (knowledge discovery) 的工作。¹ 這個問題不僅構成了社會科學與統計學一個交匯的領域，同時也是人工智慧與統計學中一個交匯熱絡的領域。很遺憾地，以上兩個交匯彼此之間似乎卻缺乏再交匯。舉個例子而言，以人工智慧與統計學會 (Society for Artificial Intelligence and Statistics) 1995年的年會所發表的62篇論文來看，幾乎所有有關應用方面的論文，都是分散在生物統計、醫療計量、語言學…等領域裡。應用在社會科學方面的論文，可以說是鳳毛麟角。而本文的目的就是希望能夠對此交匯有所貢獻。換句話說，我們希望藉著這篇文章，能讓社會科學中的實証工作者體認到：以資料庫為導向的知識發掘工作，是一個橫跨社會科學、統計學與人工智慧，一種高

* 本文承蒙管中閔教授以及兩位匿名評審提出寶貴的意見，在此一併致謝。

** 政治大學經濟系副教授

*** 政治大學經濟系博士班研究生

度科技整合性的工作，如果我們能更深入地融入在這個整合的領域中，或許將有助於社會科學實証工作的提昇。

有鑑於此，本文的構想如下：首先，我們將在第二節中，介紹一個在人工智慧領域中，最新發展的工具，即遺傳規畫 (Genetic Programming)。遺傳規畫在晚近人工智慧與統計學就有關以資料庫為導向的知識發掘工作中，扮演著一個很有潛力的角色。其次，我們要問的就是，遺傳規畫作為一個發掘知識的工具，它的功能性如何？對這樣的問題要有一個一般性的答案，在現階段而言，似乎並不容易。因此，本文的著眼點只是一個個案性的評估。我們希望能在一連串個案經驗的累積中，增加我們對遺傳規畫在發掘知識功能性上的認識，而奠定未來更一般性理論發展的基礎。

本文所使用的個案，就是在經濟學上歷久不衰的貨幣數量學說。我們的作法是在第三節中，建立一個由與貨幣數量學說相關變數所組成的小型資料庫。然後在「假裝」我們對該資料庫中缺乏任何知識的情況下，來測試遺傳規畫發掘此一知識的能力。如果我們對它的這個能力有信心，那麼將來在我們對某一資料庫「真正」缺乏知識時，遺傳規畫才會被有信心地使用作為發掘知識的工具。因此針對這點，我們的個案研究所要問的是：如果貨幣數量學說基本上是隱藏在我們的資料庫中，則遺傳規畫能不能發現此一學說。

第二，當我們對資料庫中的某一知識雖然掌握正確，但卻不是很完全時，那麼遺傳規畫是否能發掘並完備該知識。譬如說，我們雖然知道經濟資料庫中，擁有貨幣數量學說的知識，但是當我們並不確定流通速度的形式時，遺傳規畫在發掘出貨幣數量學說的同時，是否能告訴我們流通速度的形式？這點是要了解遺傳規畫對知識完備性上的能力。

第三，遺傳規畫所發掘(或完備)出來的「知識」究竟能不能算知識？譬如說，當我們資料庫中並不蘊涵貨幣數量學說，此時遺傳規畫

可能不會發現或完備該一學說。相反地，它可能為我們找尋出它認為對的知識。但是這個知識對我們有沒有用呢？這個問題在知識發掘的領域中，已逐漸受到重視 (Lubinsky, 1995)。此乃是因為人工智慧所使用的工具對其所發掘知識的表達形式，經常缺乏可理解性。因此，從這個理解的角度而言，這些知識並沒有用。但是，若就預測的眼光而言，則這些不易消化的知識，往往卻有益於提高預測能力。所以，從預測的角度上來說，這些知識又顯得有用。因此，要怎麼樣在可理解與可預測之間決定知識的有用性，並非一件容易的事。本文會在後面繼續對這點加以探討。

在這樣的構想下，我們將如下安排本文的章節。在第二節中，我們將對遺傳規畫做一個簡單的介紹。在第三節中，我們將敘述我們的經濟資料庫及在該資料庫中一個歷史悠久的經濟知識——即貨幣數量學說。貨幣數量學說若從洛克 (John Locke) 算起，已有三百年的歷史。其中歷代經濟學者的貢獻，已大大豐富它的內容。本文所用來代表貨幣數量學說的貨幣數量方程式 (Quantity Equation)，只是貨幣數量學說中的一種版本。第四節裡，我們將開始探討遺傳規畫作為一種在資料庫中發掘知識的工具，它能否發掘出此一學說。第五節中我們將以預測績效來衡量遺傳規畫所發掘出來的知識。最後一節則是結論。

貳、什麼是遺傳規畫？

要了解遺傳規畫之前，我們先要了解什麼是演化計算 (Evolutionary Computation)。顧名思義，演化計算是將生物學中達爾文「物競天擇，適者生存」的進化論，運用到「計算」的領域來。由於自人工智慧領域正式誕生之前，學者對於「智慧」(intelligence) 的概念幾乎就與「計算」平行。最有名的例子就是 Turing 檢定 (Turing Test) (Crevier, 1993)。所以，就某個意義而言，演化計算可以視為生物學與人工智慧一個整合性研究下的產物。然而，更基本的問題是為

什麼這樣的整合有意義？要回答這個問題之前，就必須先了解什麼是計算？計算可以視為一個尋找答案的過程。² 這個尋找是要在解答空間 (solution space) 中，去找尋正確的答案。然而，一般而言，由於解答的數目可能相當龐大，靠著一股盲勁 (brutal force) 去找，或是隨機亂找，都是不切實際 (impractical) 或不可行 (infeasible) 的。因此，程式學 (Theory of Algorithms) 的目的，就是希望能研究出「有效」尋找的途徑。而演化型計算就是將演化的運作，明顯地使用到尋找答案的設計上來。這些運作包含了複製、交配與突變。而最重要的是它們創造了一個競爭性的環境，並彰顯了「弱肉強食，適者生存」的原則。

演化型計算概念的提出已有30年的歷史。剛開始時，它是出現在不同的領域，由不同的人用來處理不同的問題，在文獻上也分別有不同的名詞，如 Rechenberg (1965) 的演化策略 (Evolutionary Strategies)，Fogel 等 (1966) 的演化規畫 (Evolutionary Programming)，Holland (1975) 的遺傳程式 (Genetic Algorithms) 及 Koza (1989) 的遺傳規畫。早年這些支派的整合性很低。直到最近，對它們彼此之間關連性的研究才起步。而演化型計算的名詞的出現，也是象徵這一個綜合性研究的開始。誠如前面提到的，本文所使用遺傳規畫就是演化型計算下的一個支派。從淵源上來講，遺傳規畫是遺傳程式的後續發展，但是也正因為此，就它的格局而言，它似乎是目前演化性計算各種支派中，最具一般性者。

遺傳規畫基本上是將整個解答空間中的每一個解答，用一個電腦程式 (computer program) 來代表。而又因為如果每個電腦程式都用 LISP (List Programming) 語言來表示的話，則它們都具有 LISP 的一種樹狀展開表現法。³ 所以，在 LISP 語言下，每一個解答都表現成一棵樹，而整個解答空間就是一個森林。遺傳規畫的任務，就是要在這個森林中找尋一棵最好的樹。它的作法首先是透過一個隨機程序產生一個原始森林。這個原始森林像一般生物一樣，具有繁衍的能力，也就是原始森林能繁衍出第一代森林，而第一代森林又能繁衍出第二代森

林。繁衍的方式也不外是複製 (reproduction) 與交配 (crossover)。問題是並不是所有的樹都有一樣的機會能使用複製或交配去繁衍它的下一代。相反地，複製與交配的機會是有選擇性的。凡是對環境有良好適應力 (fitness) 的樹，自然就會較那些適應力較差的樹有更多的機會被複製，或被選來與其他樹交配而繁衍後代。除此之外，在繁衍的過程中，基因也可能發生突變。至於複製、交配與突變的實際執行，就構成了遺傳規畫程式設計中的細節，這些將會在附錄A中詳細交代。至於遺傳規畫概觀性的簡略流程圖，可以參考 Koza (1992a) 一書第五章中的圖 5.1。

根據以上的描述，遺傳規畫可以說是將進化原則應用到一個以電腦程式所組成的社會中。在優勝劣敗的競爭性環境下，自行繁衍與成長。將這個方法用到知識發掘的領域裡，則每一棵樹就代表一種假說或說是一個模型。而森林就是各種假說 (或模型) 的集合。原始森林則是最起初假說 (或模型) 的集合。而用以決定每個假說 (或模型) 複製或交配機會的適應力，則可定義為該假說 (或模型) 對資料庫中資料的解釋能力。複製的意義在這裡很明顯。至於交配，則可視為兩種假說 (或模型) 的結合。突變則是某一假說 (或模型) 中的某一部份，在未經交配的情況下，自動性地變化。而在這樣的對比下，知識發掘的工作就自然可視為一個「物競天擇，適者生存」的競爭性過程。這個過程將淘汰掉不好的假說，而保留並尋找好的假說。這個過程也是使遺傳規畫能成為知識發掘工具的一個基本條件。至於成功的交配與突變，則可視為對現存優良假說的一個完備或補強。

除了知識的發掘與完備之外，從另一個角度來講，遺傳規畫也提供對任何假說一個檢試的環境。一個假說若要能夠被接受，必須是它有強大的競爭或存活力。遺傳規畫透過參數的設計，如競爭的模型數 (森林規模)，或競爭的代數 (時間)，可以增加或減弱一個環境的競爭程度。例如，森林規模愈大時間愈長的競爭愈是激烈。而在愈激烈的

環境中，能存活下來的模型，其競爭力便愈強，而可接受性便也就愈高。

從以上的說明中，我們也可以看出遺傳規畫和一般統計模型的關係。一般的統計模型無論是母數或是無母數在估計它的一些參數時，都會使用到某種數值方法作為尋找過程的設計依據。但是，這些尋找方式基本上而言仍是「單點到單點」(single-point-to-point)的尋找，而遺傳規畫(或遺傳程式)則是「多點到多點」(multiple-points-to-points)的尋找。因此，在克服區域最適(local optimum)的陷阱上，後者應有較佳的效率。⁴然而，這並不是意謂遺傳規畫與一般統計處於對立競爭的地位。比較合宜的看法應是遺傳規畫是提供一種新的尋找典範，這種典範可以應用來提高處理一般統計模型中尋找上的效率。譬如說，它可以用改進逐步迴歸中的尋找效率，也可以用來改良神經網路(Neural Networks)中的尋找績效(Angeline, Saunders and Pollack (1994))。

參、遺傳規畫與貨幣數量方程式

Koza (1992b) 是遺傳規畫在經濟學上最早的應用。在他的那篇文章中，Koza 強調遺傳規畫可以用來發現潛在的知識。他並用貨幣數量方程式 $P = \frac{MV}{T}$ 做例子，說明這樣一個蘊藏在資料中的知識，可以透過遺傳規畫重新發現。本文以下將重新執行 Koza 的研究，這樣做有幾個目的，第一、遺傳規畫是個隨機尋找過程，每次尋找的答案未必是一樣的。然而，Koza 對這點似乎沒有提到。在他整篇文章中，他只展示了「一次」執行的結果。這使人難以了解這次究竟是「一次中的一次(100%)」，還是許多次中的一次。畢竟，我們所關切的是：遺傳規畫是不是每一次都能學到貨幣數量方程式。如果不是，則學到的機率或困難有多高？這些都是我們在評價遺傳規畫發掘知識的能力

時，不可或缺的了解。第二、Koza 所執行的遺傳規畫不僅成功地發現了貨幣數量方程式，同時，還得到流通速度是常數的結果。只是這個結果經過驗証，似乎並不正確，Koza 的結論是根據他文中70頁(圖 7)的展示的樹狀程式推導而來。但是，這個樹狀程式所代表的一個如下的函數：

$$P = \frac{M2[1.234366 + h(.)]}{GNP}$$

其中，

$$h(.) = \frac{M2(0.47 * GNP + 0.00235)}{0.47 * GNP^2 - 0.3749 * GNP + 0.0003008}$$

很明顯地，這個函數並無法簡化成一個常數。Koza 似乎想用此結果來說明遺傳規畫所發現的與他所熟悉的計量文獻「竟然」完全一致。我們認為這裡有點誇張的嫌疑。因此，重新執行他的作業，也是想再度確認這點。最後，我們也想了解 Koza 的這套方法執行在臺灣的資料會得到些什麼結果。爲了確定我們所使用的方式與 Koza 相同，重新再執行 Koza 的作業是有必要的。而且只有在我們執行的方式與 Koza 相同時，遺傳規畫在美國資料與臺灣資料對貨幣數量方程式發掘過程的比較才會有意義。

所以，基於以上這些理由，我們將在第四節中重新執行 Koza 的遺傳規畫，然後在確定我們的遺傳規畫與 Koza 是相同的版本後，再將這套程式執行在臺灣。同時，並比較它與美國經驗的異同。在此之前，讓我們先對資料來源做一簡單的說明。

本文中所使用的資料是從 EPS 資料庫中取得，相關資訊請參見表 1 與表 2。除了貨幣供給量與臺灣的利率是月資料外，其餘皆是季資料。而貨幣供給量與臺灣利率的季資料，是以3個月的平均數作爲其季資料。美國的資料是從 1963: 1 到 1992: 4，而臺灣的資料則是從 1962: 1 到 1991: 4。因此兩國皆有120個觀察值。

表 1：美國資料

代號	意義	位址
PGNP	GNP 平減指數	USCEN
GNP87	實質 GNP (基期 = 1987)	USCEN
MNY2	M2 貨幣供給	USCEN
60c	國庫券利率 (三個月)	IMF 之 IFS

表 2：臺灣資料

代號	意義	位址
PGNP	GNP 平減指數	NIAQ
GNP86	實質 GNP (基期 = 1986)	NIAQ
M2	M2 貨幣供給	FSM
RMTD3@1ST	第一商業銀行存款牌告利率一定定期存款 (三個月)	FSM

由於貨幣數量方程式是以 $MV = PT$ 來表示，當我們在等號兩邊皆取 \log 後，可以得到下列的線性關係：

$$\log (P) + \log (T) = \log (M) + \log (V) \quad (1)$$

而方程式 (1) 又可改寫成

$$\log (P) = \log (M) - \log (T) + \log (V) \quad (2)$$

因為在我們資料庫中 V 不是直接可觀察的，所以，遺傳規畫在執行中，就會自動用我們所提供的其他變數來代替。此時，遺傳規畫實際上所要發掘的知識應是方程式 (3)。

$$\log (P) = \log (M) - \log (T) + k (M, T, IR) \quad (3)$$

其中「 IR 」表示利率。方程式 (3) 等號的右邊可以分成兩部份，其中一部份是 $\log (M)$ 與 $\log (T)$ 的線性部份。而另一部份則是原來 $\log (V)$

的替代變數， k 。如上所述， k 一般而言是資料庫其他變數的函數。至於，這個函數的形式並沒有任何限制。一般而言，它可以是非線性。在方程式 (3) 的基礎上，如果遺傳規畫成功地發現了方程式 (3) 中的線性部份，我們便說遺傳規畫發掘到了貨幣數量學說的知識。在這種情況下，我們也可以說貨幣數量學說在我們資料中具有競爭性。而此時的 $k(\cdot)$ 則視為遺傳規畫對貨幣數量學說中流通速度函數形式的補充。反之，若是遺傳規畫的學習並沒有發現方程式 (3) 中的線性部份，則我們便說遺傳規畫並沒有發掘到貨幣數量學說的知識，或者說貨幣數量學說在我們的資料中，並沒有競爭性。以下我們便根據表1與表2的資料來比較貨幣數量學說在美國與臺灣的競爭性。亦即在我們的分析中，我們是用「PGNP」代表 P ，「GNP87」(臺灣則以「GNP86」) 代表「 T 」，「 $M2$ 」代表「 M 」，三個月國庫券利率(臺灣則以第一商業銀行三個月定期存款牌告利率) 代表「 IR 」。

表 3：線性迴歸模型估計的係數 (美國)

參數估計				
變數	估計值	標準差	t-值	p-值
Intercept	-0.941802	0.39298943	-2.397	0.0182
$\log(GNP)$	-0.773090	0.06621536	-11.675	0.0001
$\log(M2)$	0.930928	0.02095407	44.427	0.0001
$\log(IR)$	0.045321	0.00672047	6.744	0.0001

$$SSE=0.07001, R^2=0.9977, \bar{R}^2=0.9977, DW=0.152$$

肆、實證結果

首先，我們對 $\log(P)$ 做迴歸分析，結果列於表 3 中。而 $\log(GNP)$ 、 $\log(M2)$ 和 $\log(IR)$ 為自變數，其中適合性 (fitness) 定義為誤差平方和 (SSE)。此處迴歸分析得到的 SSE 為 0.07。在當代計量經濟

學發展中，當我們去跑這樣一個簡單的線性迴歸時，有許多問題要注意。其中，在近來文獻發展上較引人矚目的就是 Granger 與 Newbold (1974) 所提出的假迴歸 (spurious regression)。假迴歸的發展到了 80 年代在計量經濟學中，產生了一些重要的概念，尤其是單根 (unit root) 與共整合 (cointegration)。除了這些之外，發展於 70 年代早期，而在晚近已與共整合納入一個體系的 Granger causality 都會使得我們所使用的簡單迴歸方程式呈現地太簡單。所以，在這裡我們必須指出我們不是要用這條簡單迴歸方程式來「檢定」貨幣數量學說。我們只是用這條迴歸方程式中所產生的 SSE 來決定遺傳規畫這種自動尋找機能所得到的適合性是否可以接受。如果遺傳規畫的表現連我們可順手撿來的線性知識都不如的話。它的可接受性無疑地將受到很大的保留。至於如何將遺傳規畫配上當代的計量工具來使用，則自是未來應再研究的。⁵ 本研究基本上是平行於 Koza (1992b)，所以在希望能與 Koza 的結果做一個直接比較的前提下，對更複雜卻更嚴謹的層面暫且省略。

接下來，我們便按照第二節的敘述，用遺傳規畫在表 1 所代表的資料中，發動一個發掘知識的假說 (或模型) 競爭環境。表 4 與表 5 是該環境的參數設定。表 4 與表 5 的差別，只在於函數集合的不同。在表 4 中，函數集合包含了 \sin 與 \cos 函數，然而在表 5 中卻不包含這兩個函數。這些參數基本上可讓我們選擇演化 (競爭) 的規模及演化的過程。在規模方面包含了競爭對手的多寡 (族群個體總數)，及開放競爭的時間 (子孫代數)。而在過程方面則已包含了基本結構 (函數集合與終點集合)、創造 (完整法與成長法)、及演化運作元 (複製、交配、突變與新生嬰兒) 的進行方式。雖然遺傳規畫的結果會受設定參數的影響，但是這方面的研究尚無進展。⁶ 所以目前尚沒甚麼理論能告訴我們一個參數的設定是太高還是太低。所以，我們對參數值的設定基本上是依照 Koza (1992a, 1992b) 的方式。

表 4：模擬 a1-a5 的參數設定

子空間個體總數	500
原始子空間深度等級	2, 3, 4, 5, 6
每一等級完整法創造個體的數目	50
每一等級成長法創造個體的數目	50
函數集合的元素	{+, -, ×, %, exp, Rlog, Sin, Cos}
終點集合的元素	{GNP, M2, IR, R}
複製的個數	50
新生嬰兒的個數	50
突變的個數	100
突變的機率	0.2
個體長度的限制	17
終點被選為交配點的機率	0.5
子孫代數	100
配適值標準	SSE

對於表 4 與表 5，我們分別執行 5 次的模擬，其結果列於表 6 中。表 6 分別陳列各次模擬其 *SSE* 在第 0 代、50 代與 100 代的表現。其中 a1 到 a5 是根據表 4 所得到的結果，而 b1 到 b5 則是根據表 5 所得到的結果。從表 6 中，我們可以看到，雖然進化的速度不同，但進化的現象在每次的模擬中，都明顯地陳列。此外，我們也可以看到，在經過 100 代的演化後，依據表 4 所執行的 5 次模擬中，有兩次的結果優於線性迴歸所得到的結果。模擬 a2 與 a3 的 *SSE* 分別是 0.066 與 0.060。而在這十次的模擬中，模擬 a3 所選擇的模型是最具有競爭力的。其函數形式列於方程式 (4) 與 (5) 中。至於其它各次模擬所選擇的最佳模型，則列於附錄 B.1 中。為了便於比較，我們將模擬 a3 與 OLS 對 $\log(P)$ 的配適殘差值列於圖 1。

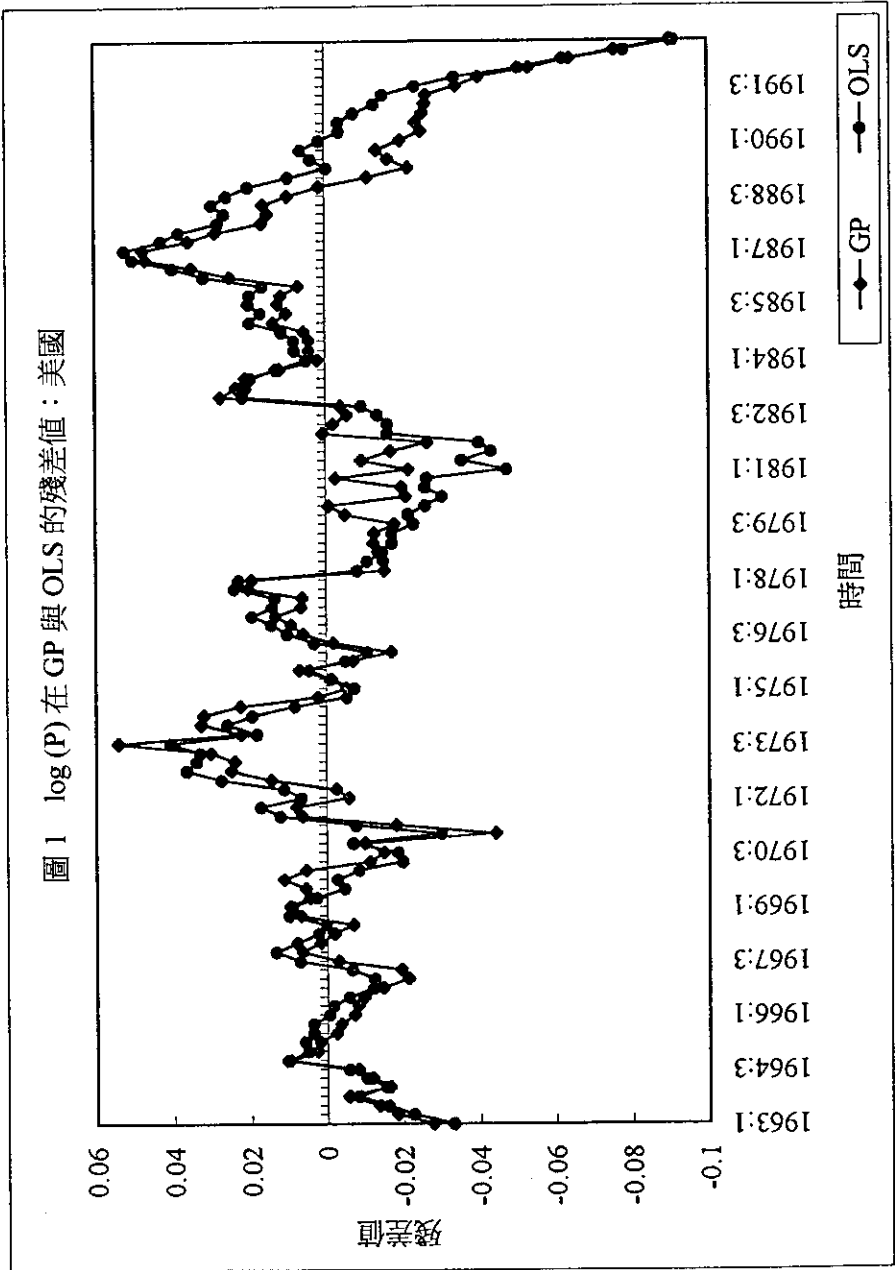
表 5：模擬 b1-b5 的參數設定

子空間個體總數	500
原始子空間深度等級	2, 3, 4, 5, 6
每一等級完整法創造個體的數目	50
每一等級成長法創造個體的數目	50
函數集合的元素	{+, -, ×, %, exp, Rlog}
終點集合的元素	{GNP, M2, IR, R}
複製的個數	50
新生嬰兒的個數	50
突變的個數	100
突變的機率	0.2
個體長度的限制	17
終點被選為交配點的機率	0.5
子孫代數	100
配適值標準	SSE

表 6：各次模擬在 Gen= 0,50 及 100 時的 SSE (美國)

模擬	GEN = 0	GEN = 50	GEN = 100
a1	16.6014356945	1.7560194444	1.5877132321
a2	4.5280953476	0.0943477508	0.0669074593
a3	18.4662807460	0.0662950995	0.0607039044
a4	14.1192882072	1.9335507499	1.5578049443
a5	9.7651900811	2.6750861048	1.6254611716
b1	24.2477124200	0.1209872472	0.0950912770
b2	7.4295429800	0.1248445692	0.1222076943
b3	6.7823155886	0.6041591037	0.5166354455
b4	24.7274992134	0.5533928640	0.1564576940
b5	24.4970863778	0.5991962402	0.2972428306

圖 1 $\log(P)$ 在 GP 與 OLS 的殘差值：美國



$$\log(P) = \log(M2) - \log(GNP) + k(GNP, M2, IR) \quad (4)$$

其中

$$\begin{aligned} k(GNP, M2, IR) = & ((gnp / (((((gnp / ((m2 - gnp) - gnp)) + gnp) - \\ & (-0.56)) + ((gnp + (ir * (((sin(m2 / gnp) + gnp) * sin(-9.54)) - \\ & (sin(m2 / ((gnp * (-0.56 - gnp)) * (gnp / gnp)))) + gnp)))) * sin \\ & ((ir / gnp) * (gnp - m2)))) - (-9.54)))) \end{aligned} \quad (5)$$

且

$$gnp = \log(GNP) \quad m2 = \log(M2) \quad ir = \log(IR)$$

由於(4)在線性的部份完全符合貨幣數量學說的形式，因此我們基本上認為遺傳規畫發掘到了貨幣數量學說的知識。這點與 Koza (1992b) 的發現是一致的。我們也必須注意到：遺傳規畫並不是在每一次都學到了貨幣數量學說，這點是 Koza (1992b) 所沒有指出的。事實上根據附錄 B.1，在我們所做的十次模擬中，只有模擬 a3、b1 與 b2 三次成功地學到貨幣數量學說。換句話說，遺傳規畫在美國資料中，成功發掘貨幣數量學說的機率是30%。

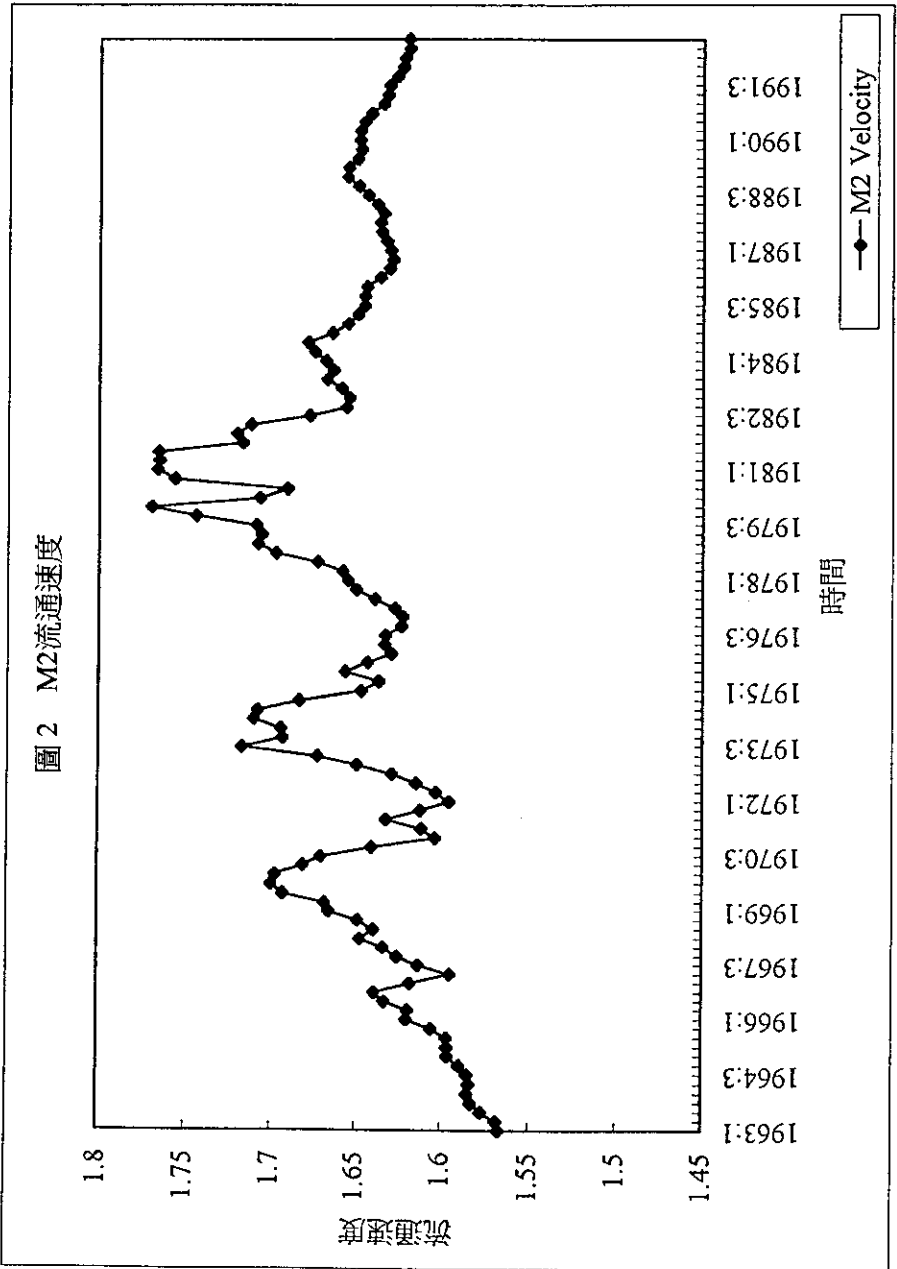
此外，若是將貨幣數量學說中流通速度的函數形式分為常數與非常數，則我們所得到的結果與 Koza 並不一致。Koza 得到的流通速度是一個常數項，即 1.6527。而流通速度為常數的貨幣數量學說，則只是我們三個學到貨幣數量學說模擬中的一個，即模擬 b2。根據附錄 B.1，模擬 b2 所學到的正好就是流通速度為常數 1.6610 的貨幣數量學說。問題是根據表 6，模擬 b2 所產生的最佳模型，在十次模擬中並不是最具有競爭力的模型。相反地，在最具競爭力的模擬 a3 中，我們所得到的流通速度則如方程式(5)所示，它是一個利率、GNP 與 M2 的非線性函數，而非一個常數。

在圖 2 中，我們將方程式 (5) 所估計出的流通速度陳列出來。我們可以看出雖然流通速度不是一個常數，但它是相當穩定地在 1.5 與 1.8 之間波動。而它的平均數是 1.6508，與 Koza 的也非常接近。Hallman et al. (1989) 與 Humphrey (1989) 均認為美國在戰後 M2 流通速度是相當穩定的。我們遺傳規畫上的發現，基本上是支持他們的觀點。

然而這個結果在其他的個案是否一樣會成立呢？我們以相同的程序測試臺灣的資料。線性迴歸的結果列於表 7 中，此時的 *SSE* 為 0.706。同樣地，我們分別依據表 4 與表 5 的參數設計對遺傳規畫進行 5 次的模擬。表 8 則是各次模擬中，最佳模型在第 0 代、50 代與 100 代的 *SSE*。其中 a1 到 a5 是在表 4 下所得到的結果，而 b1 到 b5 則是在表 5 下所得到的結果。與執行美國資料的結果相似，在 100 代的演化後，表 4 中的參數所執行的 5 次模擬中，有兩次的結果優於線性迴歸所得到的結果。模擬 a4 與 a5 的 *SSE* 分別是 0.634 與 0.506。最具競爭力的模型是由模擬 a5 所得到的，其函數形式列於方程式 (6)，而每次模擬所選擇的最佳模型則列於附錄 B.2 中。同樣爲了比較，我們將模擬 a5 與 OLS 對 $\log(P)$ 的配適殘差值列於圖 3。

$$\begin{aligned} \log(P) = & (\log(((gnp + ((\log(((((-7.19 - gnp) - (((m2 - gnp) - (gnp \\ & + gnp)) - gnp) * ir)) / gnp) - ir) - (m2 * m2)) + gnp) - \log((((gnp + gnp) \\ & - m2) - ((m2 - gnp) * (((gnp + ir) + ir) - (m2 - m2)))))) / gnp) - \\ & ir))) + gnp) - (m2 * m2)) - \exp.\log.\sin.\log((m2 - gnp) - (((m2 - \\ & \cos(m2 + (m2 - (m2 - m2)))) - \cos(m2 + (m2 - (m2 - (-4.25)))))) * \\ & ((-4.25 + ((gnp + gnp) - (gnp * m2))) - (((m2 + ir) + ir) - (m2 - gnp)) * \\ & m2)))))) \end{aligned} \quad (6)$$

圖 2 M2流通速度



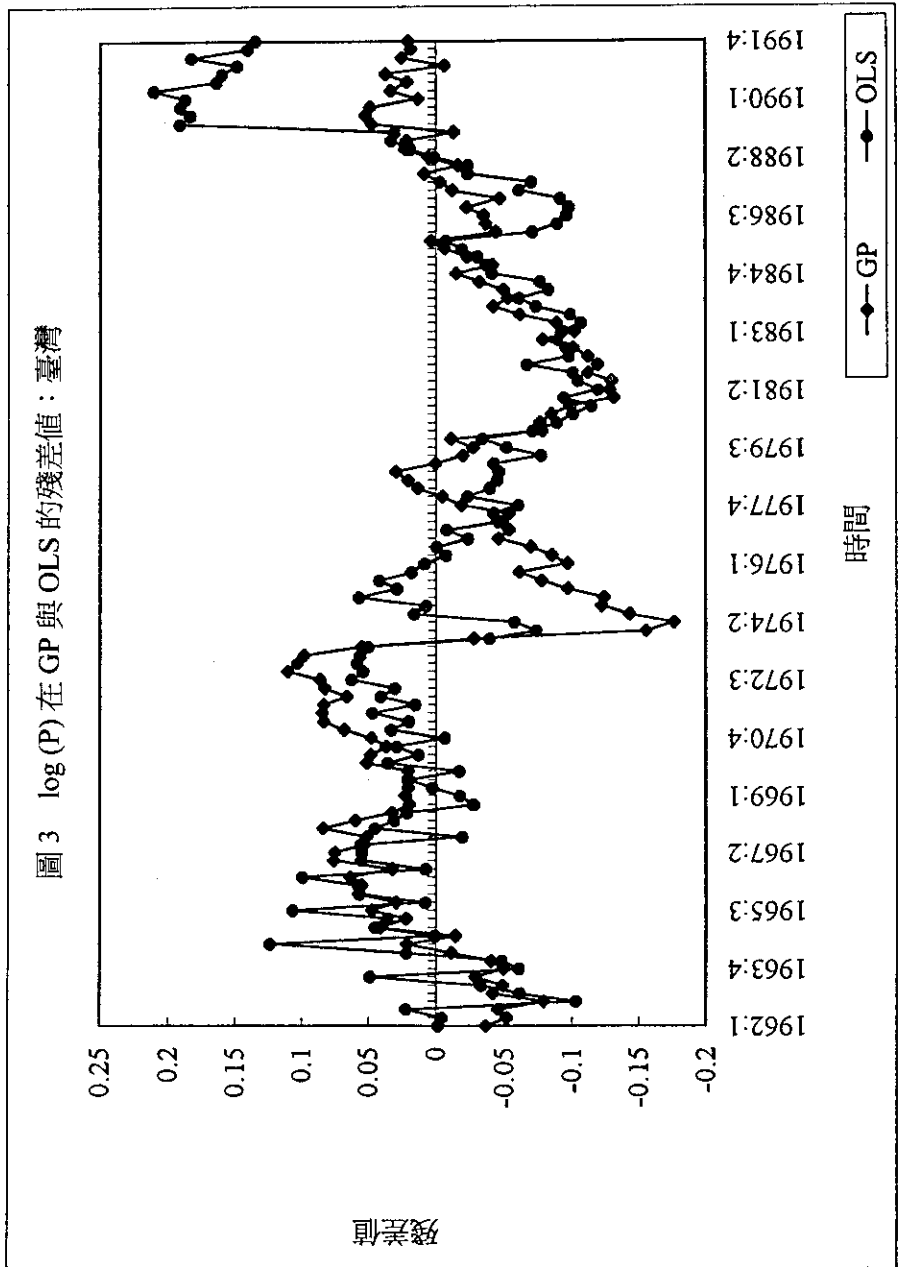


表 7：線性迴歸模型估計的係數 (臺灣)

參數估計				
變數	估計值	標準差	t-值	p-值
Intercept	5.150583	0.94570982	5.446	0.0001
$\log(GNP)$	-0.753168	0.12869377	-5.852	0.0001
$\log(M2)$	0.615813	0.05337688	11.537	0.0001
$\log(IR)$	0.245289	0.02838309	8.642	0.0001

$$SSE = 0.70620, R^2 = 0.9821, \bar{R}^2 = 0.9816, DW = 0.283$$

表 8：各次模擬在 $Gen = 0, 50$ 及 100 時的 SSE (臺灣)

模擬	$GEN = 0$	$GEN = 50$	$GEN = 100$
a1	39.4751786496	1.6001926406	1.6001926406
a2	41.3487536682	0.7455070149	0.7315891104
a3	26.8567646949	1.8955262774	1.1959025085
a4	33.0949060660	0.8526333263	0.6346563918
a5	22.6192243478	1.2386248062	0.5062178911
b1	32.3602415962	0.8990328900	0.8445017865
b2	35.7581759220	1.2207678276	1.0602512173
b3	12.4748416357	1.3310489601	1.3310489601
b4	22.4495660415	1.5862027886	1.3700185785
b5	43.4232369828	0.8667889255	0.7646361182

方程式 (6) 整個來講，是個相當複雜的非線性函數。它並沒有方程式 (3) 所需要的線性部份。所以，十次模擬中的最佳模型並不代表貨幣數量方程式。至於，其他九次的模擬又如何呢？從附錄 B.2 中，我們發現其他九次模擬所產生的最佳模型也沒有一個具有方程式 (3) 的線性部份。從這點來講，遺傳規畫在臺灣資料中，成功發掘貨幣數量學說的機率是零。所以不像美國資料，貨幣數量學說在臺灣的資料中並不具競爭力。這點發現也許不太令人意外。事實上，臺灣貨幣數量

成長率與通貨成長率之間一個不成比例的關係，早已受到學者普遍地注意，有關這點最近的研究，可參考 Yang & Shea (1994)。而遺傳規畫所發掘的只是再度將這個問題的面貌呈現出來。

伍、遺傳規畫是否會產生過度配適？

在上一節中，我們用美國與臺灣的資料，說明了遺傳規畫在發掘知識上所發動的假說競爭過程。透過大規模（在我們的例子是五十萬次）的演化尋找中，遺傳規畫找出這之間最有競爭力的模型。然而，在這整個過程上一個更基本的問題則是：到底遺傳規畫所發掘出來的模型究竟能不能算知識？譬如說，以美國的個案而言，遺傳規畫在模擬 a3、b1 與 b2 中，均發掘出貨幣數量方程式。但是從附錄 B.1 中，我們可以清楚地看見，除了 b2 以外，a3 與 b1 在理解上都有相當的困難。因此，在這種情況下，我們面臨到的問題是：這樣成堆而難以理解的數學方程式能不能算知識。這個問題不僅僅是遺傳規畫的問題，也是其他一些人工智慧自動尋找方法所同樣面臨的問題。這個問題也許並不是一個容易解答的問題，但是，解決的前提是我們必須正確了解它的本質。

這個問題的本質就是複雜。用 LISP 語言所表現的樹狀展開來講，樹的層級（或深度）就是一個複雜的概念。籠統地說，愈深的樹（或枝葉愈茂密的樹）所表現的就是愈複雜的模型。然而，在資訊理論中，複雜（或複雜度）並不是一個「繁瑣」或「冗長」的概念。相反地，它和經濟學中成本函數中的成本有一個非常相似的地方，亦即它們都是衡量在「最有效」的臨界邊（frontier）上。在經濟學中的成本是為生產某一單位的產量，所必須付出的最少投入（最有效的投入）。在程序資訊理論中（Algorithmic Information Theory），複雜度（或是程序複雜度（Algorithmic Complexity））則是說，某一物件（object）在萬能圖靈機器（Universal Turing Machine）展現下最短程式的長度（最有效程式的長

度)。⁷ 在這樣的定義下，複雜度的概念和我們日常用語的方式有些出入。最重要的，除非我們願意犧牲一點精確度或做「差不多推理」(Approximate Reasoning)，複雜度不是任意可以降低的。⁸ 因此，從這個角度來講，如果複雜度是造成我們理解上的障礙，那麼要增加理解度的代價就是要犧牲精確度。或者說，如果知識的前提必須是它能用語言、文字直覺來理解，而此理解又受到了複雜度的影響時，那麼精確的假說(或模型)就不太容易成為知識。

這樣一個認識是否能幫我們解決問題？以美國的例子來講，當我們很重視理解能力時，只有模擬 b2 所發掘的算是一個知識。但是，這並不代表當我們很重視精確度時，模擬 a3 所發掘的就能算是知識。因為這裡還牽涉到另外一個問題，即什麼是「精確」？以模擬 a3 而言，它的 SSE 最小，但是「過度精確」的背後，往往是統計學家所擔心的「過度學習」或「過度配適」(overfitting)。所謂過度配適，就是當我們用模型來解釋過去時，它可以解釋的天衣無縫，但是要用它來預測未來時，它的表現卻是很差。由於過度學習或過度配適所展現的另一面就是模型的過度複雜。因此統計學中對過度配適的處理基本上是將模型的複雜化視為一個要「懲罰」的對象。在實際操作上，則是建立一個懲罰函數 (penalty function)。而在目前文獻裡，幾乎所有的懲罰函數都可溯其基礎到資訊理論上。所以，由這些懲罰函數所構成的模型選擇判準又稱為「資訊判準」(Information Criterion)，如 AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), FIC (Fisher Information Criterion)。此外，MDLP (Minimum Description Length Principle) 也是種不折不扣的資訊判準。在真實模型存在的前提下，對這些資訊判準的大樣本研究已有很多。如 Scolve (1994) 與 Wei (1992)。但在小樣本的情況下又如何呢？Lauritzen (1995) 在比較這些準則時，曾下了這樣的結論：

No small sample results of any significance are known to me. (p.26)

除了小樣本是個限制外，截至目前為止，我們也不很清楚如何將資訊判準應用到遺傳規畫的架構中來，從本文的立場，這自是一個值得深入的研究方向。因此，在以下的討論中，我們仍沿用一般常用的「樣本內」(in-sample)「樣本外」(out-of-sample)的配適比較。亦即，遺傳規畫中所產生的複雜模型要成為知識之前，必須說明它的精確度絕非過度配適下的過度精確。也就是不僅它的 *SSE* 很低，而且在預測未來時，它的 *SSPE* (sum of squared prediction errors) 也應低得合理。依循 Koza (1992b) 的方式，我們將全部的資料分成兩部份，即樣本內與樣本外兩部份。樣本內的資料，是用來開啓演化的過程，而樣本外的資料，是用來測試演化的績效。在我們的實驗中，將120季的資料中的前80季作為樣本內的資料，亦即美國資料從1963: 1 到 1982: 4，臺灣資料從 1962: 1 到 1981: 4。而後40季則作為樣本外資料，美國資料從 1983: 1 到 1992: 4，臺灣資料從 1982: 1 到 1991: 4。

利用美國前80個季資料所產生的線性迴歸結果，列於表 9 中。至於遺傳規畫方面，我們則仿照前面的作法，根據表 4 與表 5 的參數分別對美國前80季的資料執行5次模擬，其結果列表 10 中。表 10 的第二欄是各次模擬在進化到第100代時，最佳模型的 *SSE*。第三欄則是該最佳模型在後樣本的 *SSPE*。根據 *SSE*，模擬 a4 所選擇的最佳模型是最具有競爭力的。其函數形式列於方程式 (7) 中。

$$\log(P) = \log(M2) - \log(GNP) + k(GNP, M2, IR) \quad (7)$$

其中

$$k(GNP, M2, IR) = \sin(2.72 + (((ir + (ir + \sin(-1.48))) * (-1.48)) / ((4.33 + \sin((gnp - (\sin(gnp / ((m2 * m2) / ((m2 * (-1.48)) + gnp))) - gnp)) / (((\sin(m2) / m2) * \sin(gnp)) - ((\cos(m2) - (ir * (gnp + m2))) + m2)))) * m2)))$$

表 9：線性迴歸模型估計的係數 (美國)：1963: 1 — 1982: 4

Model-1 參數估計				
變數	估計值	標準差	t-值	p-值
Intercept	0.635123	0.28049390	2.264	0.0264
$\log(GNP)$	-1.036667	0.04699530	-22.059	0.0001
$\log(M2)$	1.003368	0.01539021	65.195	0.0001
$\log(IR)$	0.078958	0.00686719	11.498	0.0001

$SSE = 0.01667$, $SSPE = 0.07574$, $R^2 = 0.9984$, $R^2 = 0.9983$, $DW = 0.72$

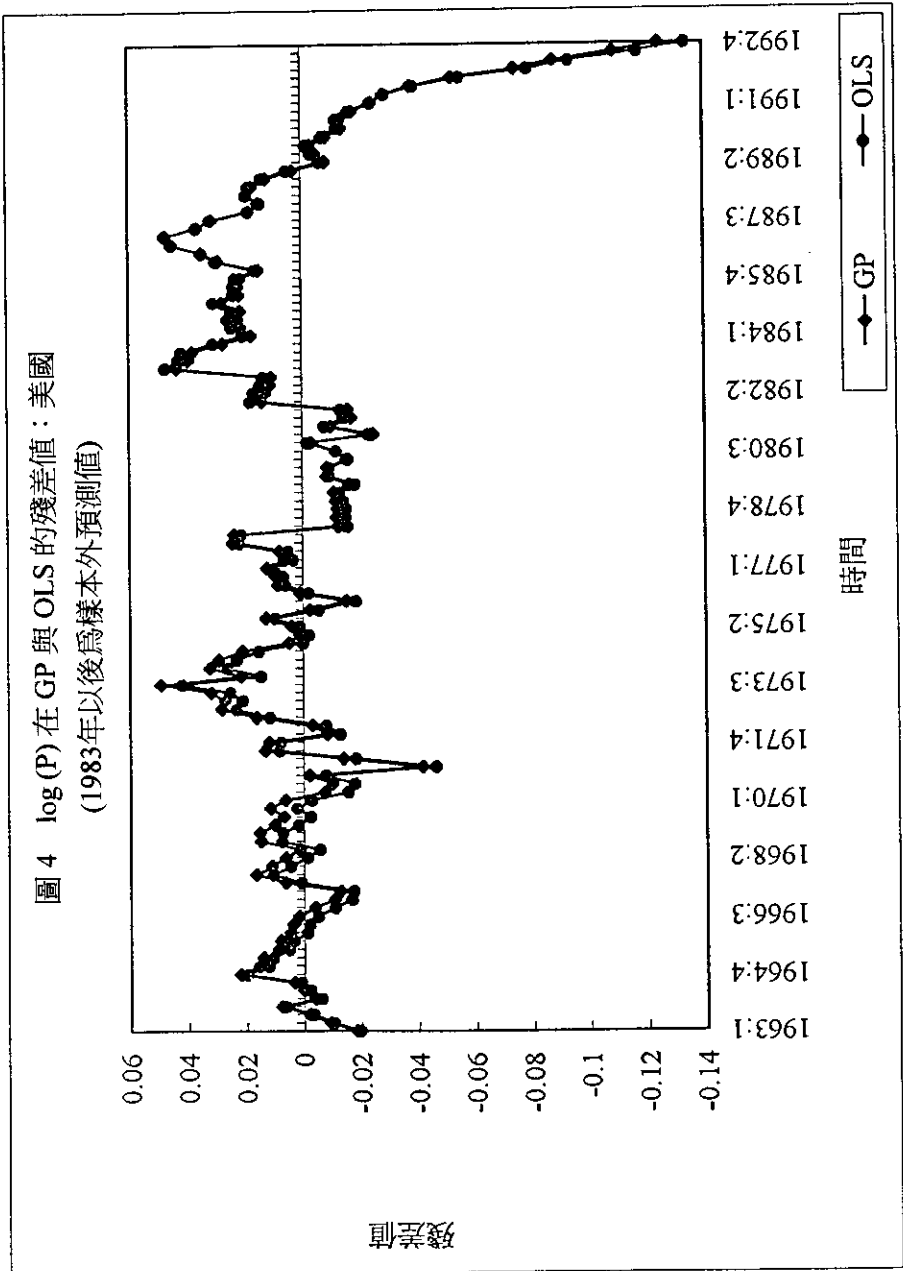
表 10：各次模擬在 Gen = 100 時的 SSE 及 SSPE (美國)：1963: 1 — 1982: 4

模擬	SSE	SSPE
a1	1.5999298828	1.7465597347
a2	1.6255382926	1.8543774187
a3	0.0717859184	2.1047432318
a4	0.0192405101	0.0672607563
a5	0.0284926297	0.0831830847
b1	0.1670751781	3.0964667583
b2	0.9946509059	15.3993457101
b3	0.0821156711	0.2780130295
b4	0.3717427121	2.2893105894
b5	0.0861279756	0.0349934493

表 11：線性迴歸模型估計的係數 (臺灣)：1962: 1 — 1981: 4

Model-1 參數估計				
變數	估計值	標準差	t-值	p-值
Intercept	4.479476	0.62668134	7.148	0.0001
$\log(GNP)$	-0.704005	0.08648681	-8.140	0.0001
$\log(M2)$	0.608745	0.03811986	15.969	0.0001
$\log(IR)$	0.325145	0.02979381	10.913	0.0001

$SSE = 0.15509$, $SSPE = 0.70262$, $R^2 = 0.9899$, $R^2 = 0.9895$, $DW = 0.895$



由於方程式 (7) 在線性的部份完全符合貨幣數量方程式 (3) 中的線性部份，因此和前面用120個觀測值的結果一樣，遺傳規畫再次在美國樣本內的資料中，發現了貨幣數量學說。至於，複雜的 $k(\cdot)$ 函數是否能視為一種「知識」，或是否有必要將流通速度視為變數而非常數，則取決於方程式 (7) 有無過度配適的嫌疑。我們將用方程式 (7) 來做樣本外40季的條件性預測，⁹ 並將此結果與線性迴歸模型相比。根據表 10 中的第三欄，我們可以看見，方程式 (7) 的 $SSPE$ 為 0.067，較線性迴歸的 0.075 為佳。¹⁰ 為了進一步了解兩種方法在條件預測上的差異，圖 4 中同時展現 GP 與 OLS 對 $\log(P)$ 的預測誤差。從圖 4 中，我們發現 GP 在 1983 年以後的 40 季中，其對 $\log(P)$ 的預測，有四分之三以上的時間都較 OLS 為佳。

另外，為了更嚴謹起見，我們也依照 Diebold and Mariano (1995) 中所提 S_2 統計量，在假設損失函數是絕對誤差的情況下，對 GP 與 OLS 預測的準確度進行近似檢定 (asymptotic test)。為了確定 S_2 統計量使用的方式，我們先對兩個預測的損失誤差進行 Box-Pierce X_2^2 檢定。結果並無法拒絕原始假設 ($X_{12}^2=6.227$)，亦即序列不相關的假設成立。然後，我們再進行 S_2 檢定。因為 $S_2 = 10$ ，所以在 0.05 型一誤差的雙尾檢定下，我們拒絕原始假設，而 GP 的預測準確度因而顯著優於 OLS。

所以，這點也代表了方程式中的 $k(\cdot)$ 並沒有過度配適 M2 的流通速度。因此，綜合前一節的結果下來，我們可以做這樣一個結論：遺傳規畫運用在美國的資料中，所發現的不僅是一個具競爭力的模型，而且也是個具競爭力的知識。這個模型 (或知識) 就是貨幣數量方程式 (學說)。它的流通速度，是隨著時間而變動的，並非常數。

同樣地，我們將相同的程序運用在臺灣的資料上。表 11 與表 12 分別是線性迴歸與遺傳規畫的結果。從表 12 「 SSE 」一欄中，我們可以看出最具競爭力的模型是由模擬 a5 所選出的最佳模型。其 SSE 值為

0.056，較線性迴歸模型的0.15509為低。而它的函數形式則列於方程式(8)。

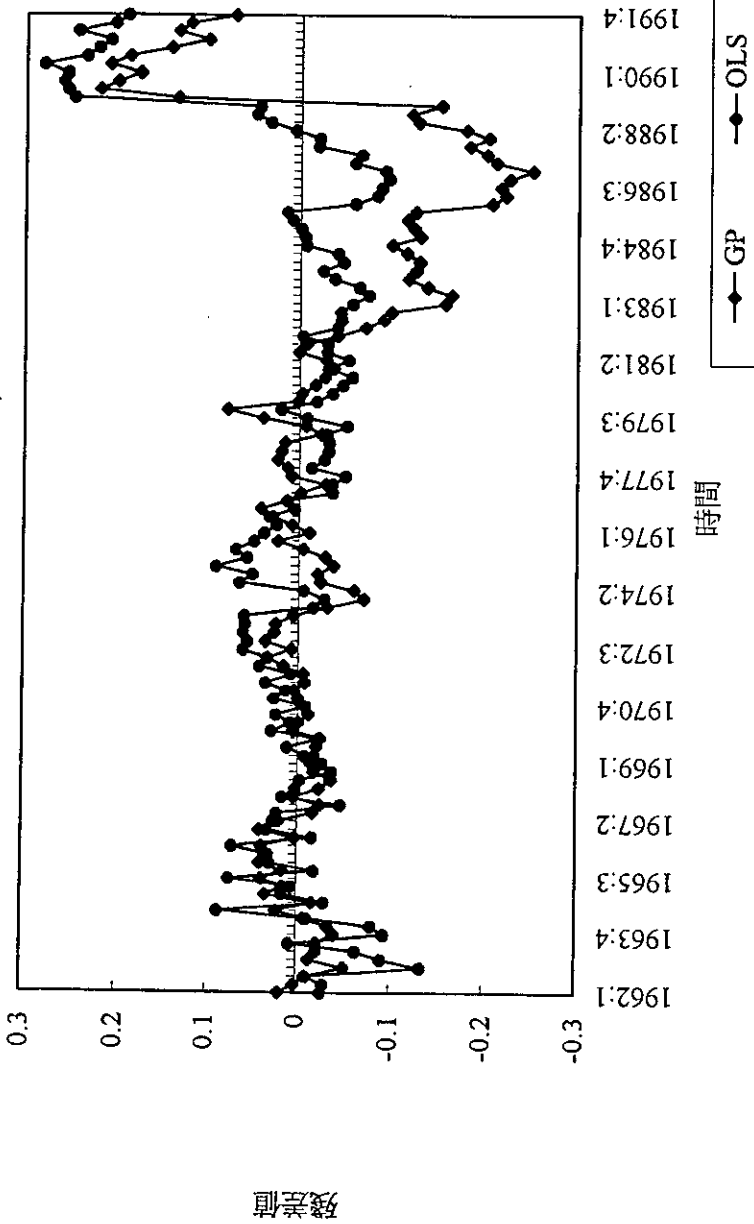
$$\log(P) = \log(((\sin(\sin(\log.\exp.\log.\exp.\sin(m2) + \sin.\log.\exp.\log.\exp.\sin(m2))) + \sin(gnp)) * \log(((\log(m2) * gnp) * \log.\exp(m2)) + \exp(\exp.\log.\exp.\sin(m2) + (\sin(\sin(gnp + \sin(m2))) + \sin(m2)) * \log((((gnp + \sin(m2)) * m2) * \log.\exp.\log.\exp(m2)) + \exp(1.72 + \sin(1.72)))))) - (-7.69)) + (m2 / \sin.\log(\log.\exp.\log.\exp(m2) + (((4.86 / \log.\exp(gnp + \sin.\sin(gnp + \sin(m2)))) + ir) + ir)))) \quad (8)$$

表 12：各次模擬在 Gen=100 時的 SSE 及 SSPE (臺灣)：1962:1 — 1981:4

模擬	SSE	SSPE
a1	0.1924329652	7.4097310516
a2	0.5120331812	5.3415378651
a3	0.1634411429	0.7357428760
a4	0.7196929084	0.6155675044
a5	0.0567330853	1.0089169895
b1	0.6988117015	31.8570062237
b2	0.6229769314	3.7542788341
b3	2.8936244352	4.2058834105
b4	0.3024963573	0.9325835964
b5	0.3027790330	3.9627926138

方程式(8)與前一節的方程式(6)有兩個相似點。第一、它們都是相當複雜的非線性函數。第二、它們都沒有方程式(3)所需要的線性部份，所以貨幣數量方程式在臺灣樣本內的資料中，並沒有被發掘到。除此之外，在其他九次模擬中，也沒有任何一次發掘到貨幣數量

圖 5 $\log(P)$ 在 GP 與 OLS 的殘差值：臺灣
(1982 年以後為樣本外預測值)



方程式。¹¹唯一比較接近的是模擬 b3。模擬 b3 所得到的是一個線性函數(方程式(9))，但是它並沒有方程式(3)線性部份所要的 m_2 。而且從表 12 中，也可以看出它的 SSE 是十次模擬中「特別」差的。所以，整體來講，不像美國，貨幣數量方程式在臺灣資料中，並沒有被遺傳規畫發掘到。

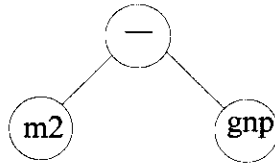
$$\log(P) = (gnp - 8.66) \quad (9)$$

至於模擬 a5 所發現的模擬能不能算知識，有沒有過度配適？從表 12 的 $SSPE$ 一欄中，我們可以看到模擬 a5 對樣本外 40 個季資料的 $SSPE$ 是 1.008，比線性迴歸模型的 0.7026 高出一些。¹² 爲了進一步了解這個差異，圖 5 中我們比較 GP 與 OLS 在預測 $\log(P)$ 上的差異。圖 5 可告訴我們從 1982 年第一季以後，GP 對 $\log(P)$ 的預測較之於 OLS 而言，一直都處於劣勢，直到 1989 第二季後的 13 個季，GP 才顯出它的優勢來。另外，我們也就 Diebold 與 Mariano 的 S_2 進行近似檢定。首先，Box-Pierce 的 X^2 檢定並無法拒絕序列不相關假定 ($X^2_{12} = 4.8799$)，所以就正常程序而言， S_2 是 29。在 0.05 型一誤差的機率下，我們拒絕原始假設。也就是方程式(8)所代表的 GP 預測是劣於 OLS 的預測。這點雖然不利於方程式(8)成爲一種知識，但是，由於十次模擬中的九次都傾向非線性的模型。而且，其中表現最佳的模擬 a3，其 SSE 優於 OLS，而且它的 $SSPE$ 與 OLS 相差無幾。所以，就「非線性」這點而言，遺傳規畫所發現的仍是有知識的特性。因此，我們的結論是遺傳規畫運用在臺灣資料中，所發現「非線性」現象，不僅是一個有競爭力的模型，也是個有競爭力的知識。這個發現使得貨幣數量學說在臺灣的資料不具競爭力。

六、結 論

將遺傳規畫應用在以資料庫爲導向的知識發掘裡，在社會科學中仍是一個相當新的研究工作。因此，在這個嘗試性的起步裡，本文藉

著具體的應用來試圖說明遺傳規畫如何可能成爲一個知識發掘的有力工具。而在這個說明裡，我們考慮了貨幣數量學說一種最簡單形式，即數量方程式。用 LISP 的樹狀結構而言，這個等於是在檢視遺傳規畫能否發現以下這個簡單的樹：



由於這棵樹的深度只有2，在遺傳規畫的植樹過程中，應是相當容易被創造出來的樹。這棵樹在美國的資料中出現了，而且能存活下來。而在臺灣的資料中也可能是出現了，但卻沒有存活下來。所以，這其間的差別就說明了貨幣數量方程式在美國資料中比在臺灣資料中較具競爭力。當然，這並不意謂著貨幣數量學說在臺灣資料中得不到支持。也許臺灣資料中存在著某種更複雜形式的貨幣數量學說。譬如說，具有「共整合」關係下的動態貨幣數量學說。由於共整合牽涉到落差元的運作，而要将落差元的運作融入遺傳規畫的程式設計中，尚有一些困難。若能克服這些困難，則遺傳規畫除了能發掘隱藏的非線性關係外，還可被用來發掘隱藏的動態結構。這將是本文未來繼續深入研究的方向。

註 釋

- 1 在計量經濟學正統的訓練中，這樣一個基本的工作，卻有一個惡名昭彰的名稱，那就是「資料濫墾」(data mining)。有關這點，可參考 Charemza 與 Deadman (1992)，第二章。

- 2 這只是從本文應用的觀點來思考「計算」，而並不是對「計算」的一個定義。有關對計算精確的定義，請參考 Davis (1978)。
- 3 LISP 是用在1956到1962年間由數學家 McCarthy 所創造。它與 Prolog (programming in Logic) 同是人工智慧中最常用到的兩種語言。有關 LISP 的歷史，可參考 McCarthy (1981)。
- 4 自遺傳程式裡，Rudolph (1994) 已經證明，在加入「保留精英」(elitist election) 的運作元後，用來代表 GA 馬可夫鏈中唯一的滯點 (Absorbing States) 就是全域性最適點。所以，不論原始條件為何，任何尋找途徑最終都能收斂到全域均衡。至於在收斂速度上的實証研究，請參考 Gelsema (1995)。
- 5 暫且不論遺傳規畫，就以其動態模型較遺傳規畫為簡單的遺傳程式而言，目前對參數值設定的理論依據，在研究成果上也非常有限。Shapiro, Prugel-Bennett and Rattray (1994) 是作者所僅知的文獻。在那篇文章中，他們用統計力學的方式探討了複製、交配與突變之參數值對配適度分配各階動差的影響。雖然，它們對最適複製參數值的研究有較一般的結果，但是其對交配與突變的研究仍只針對了特定的問題而並不具一般性。
- 6 有關這點一些例子，可參考陳樹衡、林祖嘉與葉佳炫 (1995)。
- 7 程序複雜度在程序資訊理論中，也稱為 Kolmogorov 複雜度 (Kolmogorov Complexity)。它是整個程序資訊理論中，最重要的概念。Cover 與 Thomas (1991) 一書中，對此有專章 (第七章) 介紹。
- 8 有關這點，也請參考 Martin (1995) 的討論。
- 9 詳細結果，請參見陳樹衡與葉佳炫 (1996)。
- 10 條件性預測即指在使用方程式 (7) 來對 $\log(P)$ 進行預測時，我們是在方程式 (7) 右邊的 $M2$ 、 GNP 與 k 均已為確知的情形下進行的。也就是，我們是將 $M2$ 、 GNP 與 k 的真實值代入式 (7)，來對 $\log(P)$ 進行預測。
- 11 但是依照 Hochberg-Tamhane 的檢定 (Hochberg and Tamhane (1987), Table 7)，在 $a1$ - $a5$ 與 OLS 六個模型一起檢定的結果，其差異並不具顯著性 (顯著性水準 0.05)。

12 但在統計上而言，這差異亦不顯著。

參考資料

陳樹衡、林祖嘉、葉佳炫

1995 〈空屋率的模型選擇及其穩定性：遺傳規畫的應用〉，《住宅學報》3: 73-98。

陳樹衡、葉佳炫

1996 〈貨幣數量學說的競爭性：遺傳規畫在知識發掘上的應用〉，Working Paper Series No. 9604，政治大學經濟系。

Angeline, P. J., G. M. Saunders and J. B. Pollack

1994 “An Evolutionary Algorithm that Construct Recurrent Neural Networks,” 5(1): 54-65.

Charezma, W. W. and D. F. Deadman

1992 *New Directions in Econometric Practice: General to Specific Modelling, Cointegration and Vector Autoregression*. Edward Elgar.

Chen, Shu-Heng and Chia-Hsuan Yeh

1996a “Toward a Computable Approach to the Efficient Market Hypothesis: An Application of Genetic Programming,” Forthcoming in *Journal of Economic Dynamics and Control*.

1996b “Bridging the Gap Between Nonlinearity Tests and the Difficult Market Hypothesis by Genetic Programming,” in *Proceedings of the IEEE/IAFE 1996 Conference on Computational Intelligence for Financial Engineering*, IEEE Press (ISBN: 0-7803-3236-9), pp.34-40.

1996c “Genetic Programming and the Efficient Market Hypothesis,” in J. Koza et al. (eds.), *Proceedings of the 1996 International Conference on Genetic Programming*. Cambridge: MIT Press.

Chen, Shu-Heng, J. Duffy and C. H. Yeh

- 1996 "Genetic Programming in the Coordination Game with a Chaotic Best-Response Function," in P. Angeline, T. Back, and D. Fogel (eds.), *Proceeding of the Fifth Annual Conference on Evolutionary Programming EP'96*. Cambridge: MIT Press, MA.

Cover, T. M. and J. A. Thomas

- 1991 *Elements of Information Theory*. John Wiley and Sons.

Crevier, D.

- 1993 *AI: The Tumultuous History of the Search for Artificial Intelligence*. BasicBooks.

Davis, M.

- 1978 "What is a Computation?" in L. A. Steen (ed.), *Mathematics Today: Twelve Informal Essays*. Springer-Verlag.

Diebold, F. X. and R. S. Mariano

- 1995 "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics* 13(3): 253-264.

Dworman, G., S. O. Kimbrough and J. D. Laing

- 1996 "Bargaining by Artificial Agents in Two Coalition Games: A Study in Genetic Programming for Electronic Commerce," in J. Koza et al. (eds.), *Proceedings of the 1996 International Conference on Genetic Programming*. Cambridge: MIT Press.

Fogel, L. J., A. J. Owens, and M. J. Walsh

- 1966 *Artificial Intelligence through Simulated Evolution*. John Wiley.

Gelsema, E. S.

- 1995 "Abductive Reasoning in Bayesian Belief Networks: Using a Genetic Algorithm," in D. Fisher (ed.), *Proceedings of the 5th International Workshop on Artificial Intelligence and Statistics*, pp.245-251.

Granger, C. W. J. and P. Newbold

- 1974 "Spurious Regressions in Econometrics," *Journal of Econometrics* 35: 143-159.

Hallman, J. J., R. D. Porter, and D. H. Small

- 1989 "M2 Per Unit of Potential GNP as an Anchor for the Price Level," *Board of Governors of the Federal Reserve System*. Washington, DC. Staff Study 157.

Hochberg, Y. and A. Tamhane

- 1987 *Multiple Comparison Procedures*. New York: Wiley & Sons.

Holland, John H

- 1975 *Adaptation in Natural and Artificial Systems*. University of Michigan Press.

Humphrey, T. M.

- 1989 "Precursors of the P-star Model," *Economic Review*, Federal Reserve Bank of Richmond, July-August, pp.3-9.

Koza, J. R.

- 1989 "Hierarchical Genetic Algorithms Operating on Populations of Computer Program," in *Proceedings of the 11th International Joint Conference on Artificial Intelligence*. Morgan Kaufmann.
- 1992a *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. The MIT Press.
- 1992b "A Genetic Approach to Econometric Modelling," in Bourguine, P. and B. Walliser (eds.), *Economics and Cognitive Science*, Pergamon Press.

Lauritzen, S.

- 1995 "Statistical Methods for Inducing Discrete Graphical Models from Data," AIS Tutorial in the *Fifth International Workshop on Artificial Intelligence and Statistics*.

Lubinsky, D.

- 1995 "Tree Structured Interpretable Regression," in D. Fisher (ed.), *Proceedings of the 5th International Workshop on Artificial Intelligence and Statistics*, pp.331-341.

Martin, J. K.

- 1995 "Evaluating and Comparing Classifiers: Complexity Measures," in D. Fisher (ed.), *Proceedings of the 5th International Workshop on Artificial Intelligence and Statistics*, pp.331-341.

McCarthy, J.

- 1981 "History of LISP," in R. Wexelblat (ed.), *History of Programming Languages*. Academic Press.

Rechenberg, I.

- 1965 "Cybernetic Solution Path of an Experimental Problem," *Royal Aircraft Establishment Translation No. 1122*, Ministry of Aviation, Royal Aircraft Establishment.

Rudolph, G.

- 1994 "Convergence Analysis of Canonical Genetic Algorithms," *IEEE Transactions on Neural Networks* 5(1): 96-101.

Sclove, S. L.

- 1994 "Small-sample and Large Sample Statistical Model Selection Criteria," in P. Chesseman and R. Oldford (eds.), *Selecting Models from Data: AI and Statistics IV*. Lecture Notes in Statistics 89, Springer-Verlag.

Shapiro, J., A. Prugel-Bennett and M. Rattray

- 1994 "A Statistical Mechanical Formulation of the Dynamics of Genetic Algorithms," in T. C. Fogarty (ed.), *Evolutionary Computing*. Springer-Verlag. pp.17-27.

Stark, W. R.

- 1989 *LISP, Lord and Logic*. Springer-Verlag.

Wei, C. Z.

1992 "On Predictive Least Squares Principles," *The Annals of Statistics*, 20(1): 1-42.

Yang, Y and J. Shea

1994 "Money and Prices in Taiwan in the 1980s," 開放總體經濟問題研討會，臺北：中央研究院經濟研究所主辦。

附 錄 A

A 遺傳規畫程式設計簡介

A.1 尋夫空間

遺傳規畫是一種隨機尋找程式 (random search algorithm)。要了解它程式執行的梗概，首先要了解的是它的尋找空間 (search space)。遺傳規畫的尋找空間是由程式 (programs) 所組成的空間。更明確的說，它是由 LISP 樹所組成的空間。爲了要了解這點，我們先要有一個簡單而基本的認識，即一個程式就是一棵 LISP 樹。爲了要建立這個認識，我們假設物價水準 (P) 是 M2 供給量與 GNP 的函數，且函數形式爲 $G(\cdot)$ ，

$$G(M2, GNP) = 1.65 * \left(\frac{M2}{GNP}\right) \quad (\text{A.1})$$

如果我們將 G 這個函數寫成 LISP 的 S 表現 (LISP-S expression)，則它的形式如下：

$$G = (* 1.65 (% M2 GNP)) \quad (\text{A.2})$$

若再進一步將 G 函數的 LISP-S 表現用有根有方向的分離樹 (a rooted, directed parse tree) 來表現，則可將其繪成如圖A.1的形式。

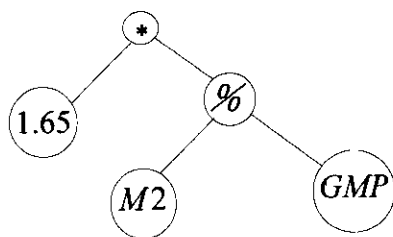


圖 A.1

所以，從方程式 (A.1), (A.2) 到圖 (A.1)，我們可以了解到為甚麼一個函數或一個程式就是一棵 LISP 樹(對這點更詳盡的介紹，可參考 Stark (1989))。

而由這些 LISP 樹所組成的空間就是遺傳規畫的尋找空間。任何一個有關物價決定的學說，若是能表現成一個函數，它就有一個 LISP-S 表現，也就有一個相對應的 LISP 樹表現。所以，凡是能被函數或程式所表現的物價學說或假設，也就自動成為 GP 尋找空間中的一員。

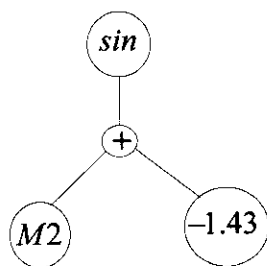


圖 A.2

但是，在實際應用上，並不是所有的 LISP 樹都可以出現在我們的尋找空間。要認識這點，我們必須再回到圖 A.1 中。在圖 A.1 中，我們

可以看到一棵 LISP 樹的組成有幾個部份。首先是所謂的函數部份，譬如說圖A.1中的 * 與 %。這些可以被允許出現在 LISP 樹中的函數的集合，就稱之為函數集合。以本文為例，我們使用的函數集合有兩類。兩者之差在於一者有 *sin* 與 *cos*，一者則沒有 (請參見與表 4 與表 5)。所以，以圖A.2所示的這棵樹而言，它可以被允許出現在表 4 的尋找空間中，但是卻不被允許出現在表 5 的尋找空間裡。因為在表 5 的尋找空間中，並沒有 *sin* 函數。當然，這裡所謂的不被允許是指不能以「直接」的形式出現。但是，一個函數如 *sin* 或 *cos* 都有它的多項式展開。所以，即使它的直接形式不允許，其間接形式 (展開逼近) 仍是有可能的。也因為此，函數集合中的函數也稱為初始函數 (initial function)。而其它的函數都是由這些初始函數所衍生而來的。

其次是所謂的常數部份，如圖A.1中的 1.65。常數部份通常是給定一個初始的浮點域。如在本文中，此一浮點域是設定在由 -9.99 到 9.99 之間所有到小數點兩位 (含兩位) 的有理數。至於其它的有理數可以再透過初始函數的運作而產生。例如，2 與 3 都在初始浮點域中，但是 $(\% 2 3) = \frac{2}{3}$ 就不在初始浮點域中，而是由初始浮點域所衍生出來的。

最後，是變數部份，如圖A.1中的 M2 與 GNP。變數部份可以允許使用 (include) 的變數也必須事先設定。如在本文中，我們設定的變數有 GNP，M2 與 IR。不在變數集合中的變數，自然不可能被用來組成 GP 尋找空間中的 LISP 樹。

一棵 LISP 樹便是由這些事先設定的初始函數，初始浮點域與變數所建構或衍生出來。由於這三者 (初始函數、初始浮點與變數) 是構成 LISP 樹的基本元素，所以它們又統稱之為原子 (atom)。另外，由於每一個 LISP 樹要有意義，必須滿足封閉的條件，也就是不能以函數集合中的元素做為 LISP 樹任一分枝展開的終點。例如圖A.3所示的就不是有意義的 LISP 樹。這樣的 LISP 樹自然不能納入我們的尋找空

間。換句話說，尋找空間的任何 LISP 樹都必須以常數或變數做為其分枝的終點。也正因為此，常數與變數集合又合稱之為終點集合。

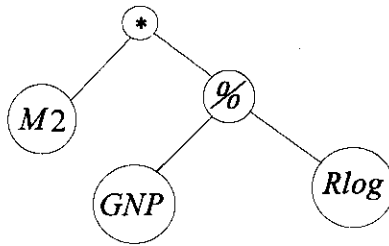


圖 A.3

A.2 尋找目標

一旦函數集合與終點集合設定，GP 的尋找空間便已確定。接下來的便是尋找的目標。尋找的目標通常視要解決的問題而定。以本文為例，我們的尋找目標是找一個模型，它的誤差平方和 (SSE) 為最小，亦即我們要找的對象是 G^* ：

$$G^* = \arg \min_{G \in \Omega(F, T)} SSE(G)$$

其中 $\Omega(F, T)$ 為由函數集合 F 及終點集何 T 所衍生的 LISP 樹的集合。換言之， Ω 就是尋找空間。而

$$SSE(G) = \sum_{i=1}^n (G(GNP_i, M2_i, IR_i, R) - P_i)^2$$

其中 $P_i, GNP_i, M2_i,$ 與 IR_i 分別代表樣本內 (in-sample) 中的第 i 個現象值。此處的 $P_i, GNP_i, M2_i,$ 與 IR_i 就是表 1 與表 2 所列的經濟資料。從 $SSE(G)$ 方程式中，我們可以看出這些經濟資料分成解釋變數與被解釋變數，透過 $SSE(G)$ 扮演著決定一個方程式良窳的標準。

A.3 尋找過程

在尋找空間與尋找對象皆已給定後，GP 所要做的便是對此一尋找對象發動一系列的隨機尋找過程。這套過程簡單來講是先是一個規模可能相當龐大的尋找空間 Ω 中，創造一個原始子空間 $\Sigma_0 (\Sigma_0 \subset \Omega)$ 。然後，再利用象徵達爾文進化論中的三個基本運作元集複製、交配與突變來移動 Σ_0 ，而產生一個序列的子空間 $\Sigma_1, \Sigma_2, \dots (\Sigma_i \subset \Omega, i \in N)$ 。對這一個過程，我們下面要分幾個步驟來說明。

首先是子空間的規模。子空間的規模可以是固定的，也可以是隨著一些條件而改變的。在目前，大多數 GP 的文獻中，子空間的規模都是固定的。本文也是如此。我們子空間的規模設定在 500 (請參見表 4 與表 5)。也就是說，我們每一次或每一代尋找的結果都是一個由 500 棵 LISP 樹所組成的森林。而這個規模固定的空間就是所有不同程式所想要競相爭奪的資源。

其次是原始子空間 Σ_0 的形成。 Σ_0 的形成是透過 Koza (1992a) 所謂的交叉對分法 ("ramped half-and-half" generative method) 所產生。這套方法也為相當多的 GP 文獻所使用。本文限於篇幅的考量，在此只做一個簡單的說明。這種方法是從 LISP 樹的根開始從函數集合與終點集合隨機選項來建立一棵樹。然而這棵樹不能任意結束，其必須滿足前面提到的封閉性。譬如說，如果我們一開始是隨機選到了終點集合中的 M2，那麼由於 M2 已經是終點集合，所以一棵 LISP 樹已經完成。但是，假設我們一開始是隨機選到了函數集合中的 *。則由於 * 函數後面必須接兩個原子 (atom) 才有意義，因此，這個植樹過程就必須繼續。在繼續的過程中，若是第一個原子抽到 1.35，第二個原子抽到 M2，那麼由於 1.35 及 M2 都是終點，所以此時，這棵樹便完成了 (圖 A.5)。但是若第二個原子抽到函數 +，則由於函數 + 並不能做終點，而且 + 必須後接兩個原子。因此，植樹的過程便還要繼續。如到圖 A.6 所示才能停止。在這樣的過程中，LISP 樹的分枝每往下一個層

次深度就加深一層。譬如說，以圖A.4, A.5與A.6所示的 LISP 樹而言，其深度分別是 1, 2, 3。



圖 A.4

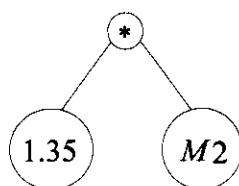


圖 A.5

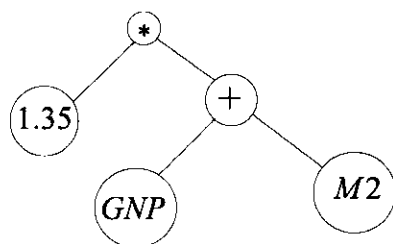


圖 A.6

至於，在植樹過程中，每一個出現的結點(即 LISP 樹中的圓圈)，其內容該如何決定則完全是隨機的。至於隨機的方式，基本上有兩種。第一種方法叫做成長法 (grow method)。成長法基本上是設定一個樹的深度 K ，然後在這個深度以內讓樹自由的成長。在這個情況下，當樹的深度還不到 K 時，任何出現的結點可以由函數來填補，也可以由常數或變數來填補。至於由何者來填補，則是以均勻分配的方式來決定。譬如說，在本文表 4 中，函數集合總共有 8 個元素，而終點集合共有 4 個元素，因此函數集合聯集終點集合一共有 12 個元素。所以，每

一個元素被選上的機率都是 $\frac{1}{12}$ 。可是當 LISP 樹的深度已經達到 K 時，則植樹過程必須要做一了結，所以此時終點集合只允許由終點集合中的元素依均勻分配的方式來填補。以本文為例，我們的終點有四個元素，所以每一個被選上的機率均為 $\frac{1}{4}$ 。

第二種方法則叫做完整法 (full method)。完整法是要要求 LISP 樹的深度一定要長到設定的深度 K 。深度沒有到 K 時，植樹的過程不能停止。基於這樣的限制，當樹的深度小於 K 時，所有的終點都必須由均勻分配的方式從函數集合中產生。也就是，在本文中每一個函數都有 $\frac{1}{8}$ 的機率被選上。而當 LISP 樹的深度到 K 時，植樹過程必須結束。所以此時就和前述的成長法一樣，所有在第 K 層的結點都是由均勻分配的方式從終點集合中產生。

在以上所描述的隨機選樣的過程中，若選上的是常數。則我們還要依照均勻分配的方式由初始浮點域 (即本文中的 $[-9.99, 9.99]$) 中去產生一個確定的數值。此外，每一個原子在進行隨機選樣時，彼此之間是獨立。而在 Σ_0 中，每一棵 LISP 樹的建立過程彼此之間也是獨立的。

至於交差對分法就是將成長法與完整法等比重地使用於給定的各種深度中。譬如說，在本文中我們依照 Koza，給定 2,3,4,5,6 五種等級的深度。而由於 Σ_0 的規模是 500。所以，在每一個深度中，用完整法創造的 LISP 樹與用成長法創造的 LISP 樹各為 50 棵 (請對照表 4)。

在初始空間 Σ_0 建立後，遺傳規畫接下來的工作就是要從 Σ_0 中去產生一個規模相等的空間 $\Sigma_1 (\Sigma_0 \rightarrow \Sigma_1)$ ，然後再從 Σ_1 去產生 $\Sigma_2 (\Sigma_1 \rightarrow \Sigma_2)$ ，依此類推，循序漸進。然而在任何一個 $\Sigma_{t+1} (t \geq 0)$ 產生之前，我們先要對 Σ_t 中的所有的 LISP 樹進行一個配適度 (fitness)

的衡量。直覺上來講，也就是我們要知道那些是「好樹」，那些是「壞樹」。至於「好」「壞」的標準，則自然隨著問題的性質或尋找的對象而不一。以本文為例，我們在前面已經提到，我們尋找的對象是一個能使誤差平方和 (*SSE*) 最小的樹。所以，在好壞認定的標準上，一個很自然的選擇就是 *SSE* (請參見與表 4 表 5)。因此，在 Σ_{i+1} 產生之前，我們必須先將 Σ_i 中所有 LISP 樹的 *SSE* 值算出來。為了討論方便，我們令 SSE_{ij} 為 Σ_i 中第 j 個樹的 *SSE*。

在 $SSE_{ij} (j = 1, 2, \dots, 500)$ 決定後，遺傳規畫下面要進行的工作就是根據 $\{SSE_{i,j}\}_{j=1}^{500}$ 對 Σ_i 去形成一個機率分配 $\{p_{ij}\}_{j=1}^{500}$ ，其中 $0 \leq p_{ij} \leq 1 (\forall j)$ 且 $\sum_{j=1}^{500} p_{ij} = 1$ 。這個機率分配不僅在 Σ_i 到 Σ_{i+1} 的過程中扮演很重要的角色，而且也是遺傳規畫能不能發揮其適當功能的一個關鍵因素，這點我們在後面還會提到。

$\{p_{ij}\}_{j=1}^{500}$ 的形式基本上是在考慮到了 LISP 樹的相對優越性後所給予它們的一個適當的權數。對於比較好的 LISP 樹而言，我們賦與其較高的權數 (機率值)，對於較差的 LISP 樹而言，我們則賦予其較低的權數。在不違反這個規則下，有很多種方式可以將 $\{SSE_{i,j}\}_{j=1}^{500}$ 轉換到 $\{p_{ij}\}_{j=1}^{500}$ 。在本文，我們使用 Koza (1992a) 在 "Symbolic Regression" 一節中所使用的轉換機制，即

$$p_{i,j} = \frac{a_{ij}}{\sum_{j=1}^{500} a_{ij}},$$

$$\text{其中 } a_{ij} = \frac{1}{1 + SSE_{i,j}}, j = 1, \dots, 500.$$

在文獻上，我們將 $SSE_{i,j}$ 稱為原始適合性 (raw fitness)，而將 a_{ij} 稱為調節適合性 (adjusted fitness)，並將 $p_{i,j}$ 稱為正規適合性 (normalized fitness)。

在 $\{p_{ij}\}_{j=1}^{500}$ 一旦形成後，遺傳規畫接下來就是要分三個步驟來進行 Σ_{i+1} 的建構。這三個步驟分別是由三個運作元所進行，即複製、交配及突變。這三個運作元將各自分擔式地創造出 Σ_{i+1} 中的一部份。這個分擔的比例是外生設定的。在本研究中，我們使用的分擔比例是由複製運作元負責10% (即50棵 LISP 樹)，交配運作元負責70% (即300棵 LISP 樹)，突變運作元則分擔10% (50棵 LISP 樹)，而另外剩下的10%是由所謂的「新生嬰兒」(new birth) 運作元來完成，新生嬰兒運作元也稱為「移民」(immigrants) 運作元，它在遺傳規畫中扮演的角色，相對於複製、交配與突變而言，是屬於比較次要、輔助性的。它的功能主要是在保持 $\Sigma_{i+1} (i \geq 0)$ 中 LISP 樹的歧異度 (diversity)，避免產生所謂的早熟型的收斂 (premature convergence)。

以下我們就分別對這幾個運作元的運作方式做一簡單的說明。首先是複製運作元。複製就是將 Σ_i 中的某一棵 LISP 樹原封不動的拷貝到 Σ_{i+1} 。所以，每一次複製都要先在 Σ_i 中選一棵 LISP 樹，而這選取的標準就是我們前面所提到的機率分配 $\{p_{ij}\}_{j=1}^{500}$ 。也就是 Σ_i 中的每一棵 LISP 樹 j 都有一個相對應的機率 p_{ij} 被選來用做複製的正本。複製要重覆的進行直到完成它應負擔的配額。以本文為例，複製需要進行50次，然而每一次複製的隨機取樣都是獨立的。

當複製完成後，我們的下一個工作就是交配。交配可以視為遺傳規畫中最核心的一個運作元。交配運作的方式是先從 Σ_i 中依據 $\{p_{ij}\}_{j=1}^{500}$ 獨立地進行兩次選樣，每一次選樣都會產生一棵 LISP 樹來，所以每一次交配運作都會先產生兩個 LISP 樹來。這兩棵 LISP 樹便為稱爲「配偶」或「父母」。以本文為例而言，假設在某次交配運作中，第一次選樣的結果方程式 (A.3)，

$$p = 3.18 + M2 (M2 * IR) \tag{A.3}$$

而第二次選樣的結果是方程式(A.4)，

$$p = 0.86 IR + \ln \left| \frac{M2}{GNP} \right| \tag{A.4}$$

則這兩條方程式寫成 LISP 程式後分別是方程式 (A.5) 及 (A.6)。

$$(+ 3.18 (* (M2 (* M2 IR)))) \tag{A.5}$$

$$(+ (* 0.86 IR) (Rlog (% M2 GNP))) \tag{A.6}$$

再將這兩棵 LISP 程式編成 LISP 樹便如圖A.7。

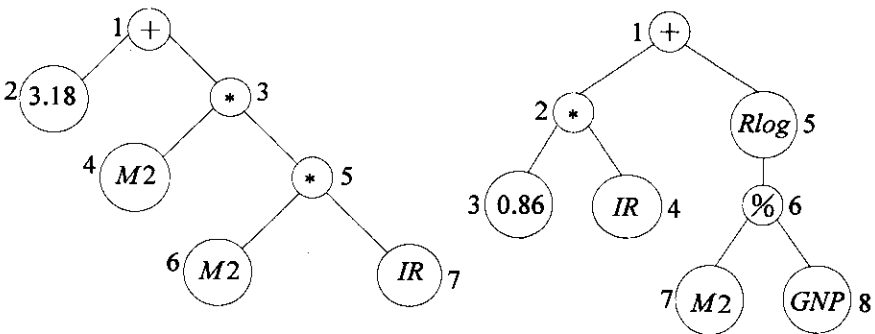


圖 A.7

圖A.7中的兩棵 LISP 樹便是父母。父母在 GP 中的責任，便是透過交配行為而產生下一代。從交配到產生下一代的過程可以分幾個步驟來說明。第一個步驟是選擇交配點 (crossover point)。這個步驟是在一棵 LISP 樹上指定一個節點，用這個節點做根，從原來的 LISP 樹中創造出一棵分離的子樹 (subtree)。這個被選定的結點便是交配點。以圖A.7為例，假設在左邊的這棵樹，我們選上了結點3做為交配點，而在右邊我們選上了點5做為交配點，則以上所說的兩棵子樹便如圖A.8所示。

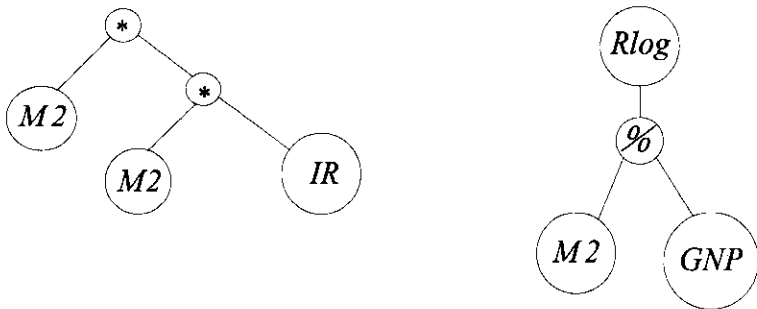


圖 A.8

至於交配點的決定也是隨機的。這個隨機的方式是先以白努力實驗 (機率1/2) 來決定交配點是終點 (常數或變數) 還是函數。一旦這個結果出來後，再進一步依據均勻分配來決定由那一個終點 (或函數點) 來做交配點。在以圖 7 為例，圖 7 右邊的 LISP 樹一共有四個終點 (即點 2,4,6,7)，另有三個函數 (即點 1,3,5)。若白努力實驗的結果是由終點做交配點，則點 2,4,6,7 將各有1/4的機率被選為交配點。若白努力實驗的結果是由函數點做為交配點，則點1,3,5將會有1/3的機率被選為交配點。

一旦交配點產生後，交配的下一個步驟便是產生下一代。產生下一代的方式便是將由交配點所決定的兩棵子樹互換至配偶子樹分離的位置。以圖A.7為例，左邊的 LISP 樹選定點3做交配點，而右邊的樹選擇點5做交配點。所以左樹中由點3所創造的子樹 (即圖A.8的左子樹) 將移接至右樹中點5以下的位置，而右樹中由點5所創造的子樹 (圖A.8的右子樹) 將移植到左樹中點3以下的位置。如此而形成的兩棵樹將如圖A.9所示。

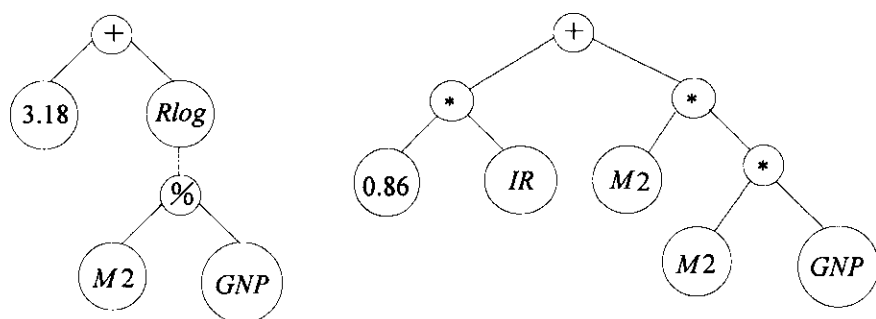


圖 9

圖A.9中的兩個子樹寫成代數形式，便是方程式 (A.7) 與 (A.8)

$$p = 3.18 + \ln \left| \frac{M2}{GNP} \right| \quad (A.7)$$

$$p = 0.86 IR + M2 (M2 * GNP) \quad (A.8)$$

而寫成 LISP 形式，便是方程式 (A.9) 與 (A.10)。

$$(+ 3.18 (Rlog (\% M2 GNP))) \quad (A.9)$$

$$(+ (* 0.86 IR) (* M2 (* M2 IR))) \quad (A.10)$$

當然，我們對樹的深度沒有任何限制，則在交配元運作不斷的運作下，樹有可能會愈長愈大(深)。太多太深的樹對整個 GP 的運算速度有相當負面的影響。因此，在本文，我們仿照 Koza 的做法將樹中的深度上限定為17(請參見表 4與表 5)。如果在交配的結果中，有任何一棵樹，其深度超過17，則我們將以1/2的機率從其父母中選擇一個做為替代，若兩個都超過17，我們將以父母做為替代。

如此，一次交配的運作便已完成，而一次交配運作的結果便是誕生兩棵 LISP 樹。這兩棵的 LISP 樹便加入到 Σ_{i+1} 中。和複製一樣的，交配必須反覆運作到完成其應用的配額。以本文為例，由於交配運作元一共要負責產生300棵樹，所以本文要重覆執行交配150次。然而每一次的隨機選樣(父母)都是獨立進行的。

交配之後的工作便是突變。突變的進行是先從 Σ_i 中依 $\{p_{ij}\}_{j=1}^{500}$ 機率分配隨機選樣產生一棵樹。然後，這棵樹上的任一結點都有一固定的機率會產生改變。在本文中，此一機率設為0.2(請參見表 4與表 5)。當一終點要變化時，我們將從終點集合中按照均勻分配的方式產生一個變數或常數來替代原有之變數或常數。至於函數點，我們則分為兩種情形來處理。其中 $\{+, -, *, \%\}$ 中的任何一函數可以被該一集合中的任一元素依均勻分配的方式所替代。至於 $(\sin, \cos, \exp, Rlog)$ 中的任何一函數也可以被該集合中的任一元素依均勻分配的方式所替代。但是，這兩個集合中的元素不可以彼此去替代另一個集合的元素。這樣子做的目的也是只在保護程式不失去其數學意義。至於，同一棵樹每一個結點其發生突變的機率在彼此之間是互相獨立的(mutually independent)。像複製與交配一樣，突變也一直要進行到它完成的配額為止。在本文中，突變要獨立的進行50次。

以上我們所介紹的三個遺傳運作元是遺傳規畫的主要運作元。它們一起完成了 $\sum_{i+1}(i \geq 0)$ 中450棵 LISP 樹的產生。在這三種運作元之外，在遺傳規畫中有時為了防止早熟型的收斂 (premature convergence) 的產生，通常還有些輔助型的運作元。其中，「新生嬰兒」(new birth) 便是一種常用的運作元。而本文也使用此一種運作元。新生嬰兒的運作方式是在 $\sum_i \rightarrow \sum_{i+1}$ 中，利用前面所談的成長法 (grow method) 在以深度不超過上限 (在本文中，此一上限設為17) 的前提下，創造一個「新」樹。這個樹稱為「新樹」是因為它的產生完全是由創造而來，而不是從 \sum_i 中衍生而來。同樣地，新生嬰兒運作元要反覆獨立地執行，直到完成其承擔的配額。在本文中，此一配額設為50 (請參見表 4與表 5)。

透過上述的四個運作元，整個序列 $\sum_0, \sum_1, \sum_2, \dots$ 便得以一個接著一個地產生。如果有一個最佳模型是 G^* ，我們希望的是存在一個 m ，使得 $G^* \in \sum_i (i \geq m)$ 。當然，如果我們不做任何額外的限制，這個尋找過程可能不會結束。所以，在本文中，如果遇到以下兩種情況，尋找過程就告結束。第一種情況是，如果存在 $G (G \in \sum_i)$ 且 $SSE(G) = 0$ ，則尋找過程便在第 i 代結束。第二種情況是，設定一個演化代數的上限，如在本文中，這個代數上限設為100 (參見表 4與表 5)。

以上便是我們對整個遺傳規畫運作過程的說明。對經濟學家而言，遺傳規畫仍是一個相當陌生的工具。限於篇幅，本文的說明可能仍然不夠詳盡。因此，我們除了推薦 Koza (1992a) (特別是第六章) 作為進階外，也提供以下幾篇目前 GP 在經濟學上僅有的應用作為參考：Chen and Yeh (1996a, 1996b, 1996c), Chen, Duffy and Yeh (1996), Dworman, Kimbrough and Laing (1996)。

附 錄 B

B 遺傳規畫模擬結果

B.1 在第100代表現最佳的個體：美國資料(1963: 1 — 1992: 4)

模擬 a1 :

$$\log(P)=\sin(((\text{gnp}-(-9.67))-m2)+\sin.\sin.\sin(((m2-9.67)-\sin.\log \\ (((\text{gnp}-(-9.67))-m2)+\text{gnp})+(-0.77-((\text{gnp}-m2)+\log(((m2-\text{gnp})-m2) \\ -(-9.67)))))))+\cos.\log((m2-(((m2-(-9.67))-m2)+\sin((\text{gnp}-(-9.67))* \\ 9.67)))/\text{gnp})))$$

模擬 a2 :

$$\log(P)=((\cos(-7.82*6.29)+\sin(m2-\text{gnp}))*\log(\log(\text{gnp}*((m2-\text{gnp}) \\ +(\text{gnp}*(m2*\exp(\cos.\cos(m2*6.29)))))))+(m2*(m2*\exp.\cos(6.29*m2))) \\))*\exp.\cos((\log(\cos(-7.82*-7.82)-m2)+(\cos(-7.82*6.29)+\sin(\\ 6.29-\text{gnp}))))+m2)))$$

模擬 a3 :

$$\log(P)=((m2-\text{gnp})+(\text{gnp}/((((\text{gnp}/((m2-\text{gnp})-\text{gnp}))+\text{gnp})-(-0.56))+ \\ ((\text{gnp}+(\text{ir}*(((\sin(m2/\text{gnp})+\text{gnp})*\sin(-9.54)))-(\sin(m2/((\text{gnp}*(-0.56- \\ \text{gnp}))*(\text{gnp}/\text{gnp}))))+\text{gnp}))))*\sin((\text{ir}/\text{gnp})*(\text{gnp}-m2))))-(-9.54))))$$

模擬 a4 :

$$\log(P)=((8.23-m2)/(m2/(-3.37)))$$

模擬 a5 :

$$\log(P)=\cos((\text{gnp}-\cos(\text{gnp}-\cos((\text{gnp}-\cos(3.47/((m2/3.47)))/(\text{gnp}/\cos.\cos((\text{gnp}-\text{gnp})+(3.47-(-4.25+m2)))))))+(3.47-(\text{gnp}+m2))))$$

模擬 b1 :

$$\log(P)=(((m2-\text{gnp})+((-0.66-(((\text{gnp}-(((0.39-(-0.66)))-(((m2-\text{gnp})+(m2-\text{gnp}))(-0.39))+(-0.66/(-5.69)))-(-0.39)))/m2)-ir))*(-0.39)/(-5.69))/(-5.69))-(-0.66-(-0.39)))$$

模擬 b2 :

$$\log(P)=((m2-\text{gnp})-(3.77/(-7.43)))$$

模擬 b3 :

$$\log(P)=(((\exp 6.80+((((((\exp(\text{gnp})/(m2-9.63))-\text{gnp})/((m2/\log.\exp(\text{gnp}/ir))+ir*(((\text{gnp}/ir)+(\exp(m2)/(gnp-\text{gnp}))/ir)/(m2-\text{gnp})))))-m2-\text{gnp})/m2)+\log(\text{gnp}))+(((\exp(m2)-\text{gnp})/(m2-\text{gnp}))/((m2/ir)+(\exp(\text{gnp})/(\exp m2-\text{gnp}))))+((m2-9.63)/((((\text{gnp}-\text{gnp})/((ir/\log((m2-6.80)/(ir-m2))))+(\exp(\text{gnp})/(m2-\text{gnp})))))-\text{gnp})-\text{gnp})+(\exp(m2/ir)/(gnp-(gnp/ir)))))))/(\exp(\text{gnp})/(m2-9.63)))$$

模擬 b4 :

$$\log(P)=\log(\log((m2+m2)-\log(m2/(5.62*(m2-\text{gnp})))))/((ir+(((5.62*(m2-\text{gnp}))/((m2/m2)*(m2/(5.62*(ir-((-0.64+m2)+m2)-\log((m2+ir)/(m2*$$

$$5.62)))))))/((m2/((m2-gnp)+\log((ir/(m2/m2))-m2))) * m2)-m2)) * (m2 - gnp)))$$

模擬 b5 :

$$\log(P) = (((m2 + (-8.39)) / (m2 + (((-9.22 - (\log(((\log(m2) - 9.36) / -6.19) / (gnp - m2)) / (gnp + (-8.39))) * gnp)) + (-8.39)) / gnp))) * \log(\log.\log.\exp(-9.22 - (((\log.\log(m2) - 9.36) / m2) / ((gnp / -6.19) / (m2 + (-8.39)))) * m2)) - gnp))$$

B.2 在第100代表現最佳的個體：臺灣資料 (1962: 1 — 1991: 4)

模擬 a1 :

$$\log(P) = (\exp.\sin.\sin.\sin.\sin.\exp.\log(gnp) + 2.83)$$

模擬 a2 :

$$\log(P) = ((\exp.\sin.\sin.\sin.\sin.\sin.\sin.\sin.\sin(6.73) / \sin.\sin.\sin.\sin.\sin.\sin.\sin(6.73)) + \sin.\sin(gnp))$$

模擬 a3 :

$$\log(P) = \log(m2 * ((-2.49 - \log(((m2 - ir) * (((\exp(\cos m2 - ((m2 - gnp) * ((ir - 3.86) / (-2.49)))) * (-2.49)) - 3.86) / ir)) + \log(((-2.49 + 3.86) - 3.86) + \log(((-2.49 + 3.86) * (-2.49) + m2)))))) + (m2 - ir))$$

模擬 a4 :

$$\log(P) = \log((\cos(gnp - ir) + ir) * (\log((m2 - \cos((\log(gnp - (-5.99 / \log(($$

$$\cos(m2)+ir)*(\log(ir)*m2))))*m2)-5.76))-(\cos(((gnp+(5.76-(\cos(ir-m2)+5.76-(m2/\sin(ir)))))))/\log(gnp))*ir)-(\cos((ir-gnp)-m2)+ir))*m2))$$

模擬 a5 :

$$\begin{aligned} \log(P)=&(\log(((gnp+((\log(((((-7.19-gnp)-(((m2-gnp)-(gnp+gnp))-gnp)* \\ & ir))/gnp)-ir)-(m2*m2))+gnp)-\log((((gnp+gnp)-m2)-((m2-gnp)* \\ & (((gnp+ir)+ir)-(m2-m2)))))/gnp)-ir))+gnp)-(m2*m2))-exp.\log.\sin. \\ & \log((m2-gnp)-(((m2-\cos(m2+(m2-(m2-m2)))))-\cos(m2+(m2-(m2-(-4.25) \\ &))))*((-4.25+((gnp+gnp)-(gnp*m2)))-((((m2+ir)+ir)-(m2-gnp))*m2)))))) \end{aligned}$$

模擬 b1 :

$$\begin{aligned} \log(P)=&(((ir*ir)*(2.82+(-2.66)))*(((m2/ir)-(ir/(m2*(m2/ir))))- \\ & (((m2/ir)-((ir*m2)/(((m2/ir)-((m2*(m2/ir))/(m2*m2))))-(2.82/(\\ & ((m2/ir)-(((m2/ir)-((ir+ir)/ir))/(ir*m2))))*(m2/ir)-((ir*4.83 \\ & -ir/(-2.66))/(ir*ir)))))))/m2))) \end{aligned}$$

模擬 b2 :

$$\begin{aligned} \log(P)=&\log((\log((ir*m2)/(5.33/((m2-\log(gnp/(\log(((gnp/gnp)/gnp)*m2)/ \\ & ((\log(ir*ir)-m2)-\log((1.38/gnp)*\log(m2)))))))-\log(5.33/(m2-5.78)) \\ &)))*m2)/(5.33/(m2-5.33))) \end{aligned}$$

模擬 b3 :

$$\log(P)=(\log.\log.\log.\exp.\log.\exp(M2)*(\log.\log.\log.\exp.\log.\exp.\exp.\log(m2))$$

(1.80*log(gnp)))

模擬 b4 :

$$\log(P)=(((m2/gnp)/(gnp/m2))+(m2/4.33))$$

模擬 b5 :

$$\begin{aligned} \log(P)=\log(&(((gnp+\log((gnp*(((gnp+5.75)-(m2+m2))+gnp)*((-4.83 \\ &+((gnp-m2)+\log(m2))))+(((-4.83+gnp)-(m2+m2))+gnp)+\log.\exp \\ &\log(m2)))))+m2)-(m2+m2))*(ir*(-4.83+((gnp-m2)+ir))))+m2) \end{aligned}$$

社會共識程度之衡量

施俊吉*

壹、前言

這是紛亂不安的年代，這是依賴社會調查作預言的年代。

在社會調查中，使用 Likert 量表的情形十分普遍。例如，當我們想瞭解，台灣社會對於「以公民投票決定台灣前途」的看法時，調查者可以提出下述之問題：

「請問您同不同意『以公民投票的方式決定台灣的前途』？」

然後讓受訪者在「非常同意」、「同意」、「沒意見」、「不同意」、「非常不同意」之間，選取一個最能代表其意見強度的選項作答。這種模式的社會調查，即是一種典型的民意測度法。

以 Likert 量表所獲得的問卷資料，經過分析以後，可以獲得諸多攸關社會意向的預言。而在所有的民意當中，有一種經常被人們掛在嘴邊，但是在衡量上，卻始終找不到明確定義的特質：社會的共識程度 (degree of consensus)。而本研究的重心，就是在探討「如何利用 Likert 量表所獲得的資料，衡量社會對於特定問題之共識程度。」

* 中央研究院中山人文社會科學研究所研究員，中央大學產經所教授。本文為國科會專題研究計畫之成果(編號：NSC 85-2415-H-001-016)。

何謂 'consensus'？牛津字典對其所下的定義是：“Agreement in opinion; the collective unanimous opinion of a number of persons.” 所以 'consensus' 的本意，應當是「完全之共識」，但是本研究的客體，則是「不完全之共識」。

本研究的起點是，不完全之共識能否分出等級與程度？本研究的終點則是，建立一項測量共識程度的指標。這項指標本文將之命名為「 s^* 指標」，它具備下述之特點：

- i) 當全體之意見一致時，不論一致同意什麼或反對什麼，皆為完全之共識，此時指標攀升到上限，上限等於 1。
- ii) 當群體均分為二，半數非常贊成，半數非常反對，意見呈兩極對峙時，指標等於 0。
- iii) 如果共識程度介於前兩者之間，例如，三分之一的人非常贊成，三分之二的人非常反對時，指標就會落在 0 與 1 之間。
- iv) 當意見呈兩極對峙時，對立程度固然激烈，但是共識程度不必然最低。例如，當一個群體的每個成員都各自堅持一種相異的意見時，這種「一盤散沙」的意見分佈型態，可能就比兩極對峙時的意見型態更缺乏共識。所以共識度指標會小於 0，以反應前述事實。

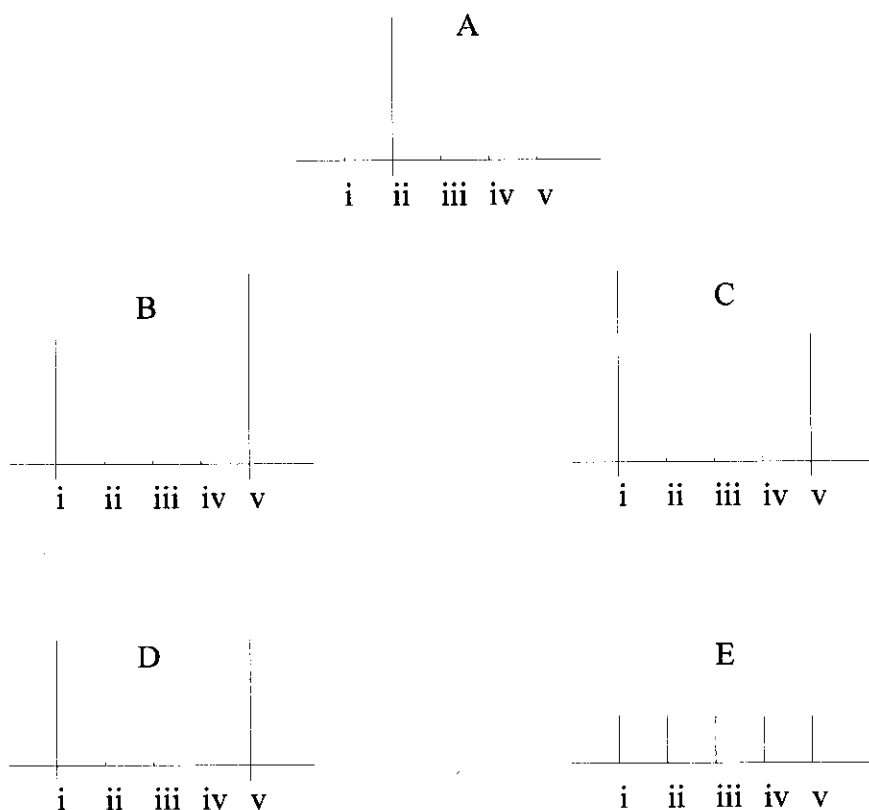
本研究為了方便社會學者利用 s^* 指標進行共識程度之分析，特地研發了一個公用軟體，專司 s^* ，以及其他相關指標的運算。這項軟體函索即贈。¹

本文之架構如下。第二節敘述共識程度的本質，並且說明其與離散度及不均度之關係。第三節以公設式分析法 (axiomatic approach) 推演共識度指標。² 第四節闡釋指標的基本性質。第五節是結論。

貳、共識程度的本質

引介一項新觀念，絕非易事。而共識程度的衡量就是一種新的觀念，所以我們必須根據直觀和概念，展開這項研究。

圖 (1)



在針對一項問題進行社會調查之後，回收的資料包括受訪者的意見，以及各種意見強度之下的統計人數。這樣的一組資料，稱為一個

「分配」(distribution)。而所謂的共識程度，必然與這個分配的型態有關。現在就讓我們以直覺來判斷圖 1 中各種分配的共識程度。

A 分配是一個完全退化的分配，因為全部的受訪者都支持同一種意見。此時應該沒有人會反對這叫「完全共識」，也不會有人反對：「完全共識是最高程度的共識，而且不論意見集中在那一點都是『完全共識』」。

準此，所謂的共識程度應該與意見的名號和稱呼無關，例如，B 分配的共識程度應該等於 C 分配的共識程度，因為這兩種分配的差別，僅止於將支持意見 i 與支持意見 v 的人數互相調換而已。

根據上述之推論，可以立即引申：共識程度在一般狀況下不等同於不平均度 (inequality)。³

此乃因為如果將原先用於代表意見強度的 i, ii, \dots, v ，重新命名為所得級距，例如：以 i 代表所得最低的一級 (少於一萬元)，以 v 代表所得最高的一級 (高於一億元)，則 B 分配的所得不平均度如果以基尼係數衡量，絕對遠超過 C 分配的不均度，雖然這兩個分配的共識程度是相同的。

如此便可獲得另外一項結論：共識度與不均度的概念不同，所以絕對不能以草船借箭的方式，從不均度指標中借用任何指標，拿來衡量共識程度。

話雖如此，但這並不表示沒有其他既存的指標足資代用。事實上，各種衡量離散程度的指標，例如變異數或均互差 (mean difference)⁴ 均有可能屏雀中選。以分配 B 和分配 C 而言，這兩項分配的共識程度相同，變異數相同，均互差也相同。因此以變異數權充共識度指標，或者是以均互差瓜代的可能性確實存在。所以我們必須藉助分配 D 與分配 E，作為進一步討論的工具。

分配 D 叫做「兩極分配」(polarization distribution)，在此狀況下，意見最弱與意見最強的受訪者，各佔半數。而分配 E 則叫做「齊一分配」(uniform distribution)，在此狀況下，支持各種意見強度的人數相等。此時可以試問：D 分配與 E 分配的共識程度，孰與強弱？

事實上，就離散程度而言，兩極分配一定比齊一分配離散，所以如果以任何一種離散度指標權充共識程度的衡量工具，則兩極分配的共識程度比較弱。

但是相信不會有太多人滿意這項答案。因為兩極分配所代表的是極端對立的社會，而齊一分配所代表的則是一盤散沙的社會。極端對立的社會比起一盤散沙的社會，我們實在很難遽下定論說，那一種社會的共識程度比較弱。所以在共識程度的衡量上，我們不能以一般的離散度指標作為代用品，武斷地認定「極端對立」較「一盤散沙」缺乏共識。

準此，Esteban 與 Ray (1994) 於晚近所發展的「兩極化指標」，顯然也不適合作為共識程度的衡量指標。因為該項指標的特色之一，即是以兩極分配作為兩極化程度最強的分配。現若採用該項指標衡量共識，無可避免地會陷入相同的武斷之中。

既然如此，我們就必須放棄舊瓶新酒的念頭，共識程度之衡量必須有其專屬之指標。

參、共識度指標的建立

3.1 受訪者之意見強度

透過前一節的分析，我們知道所謂的共識程度，乃是一個含意深邃且複雜的概念。既有的不均度指數，以及其他用於衡量離散程度的指標，都無法精確反映共識這項概念。因此，在本節當中，我們將發展出一個新的指標，作為衡量共識程度之用。

本文利用公設法建立指標。所謂的公設法，就是依據一組先驗之公設，經過逐步之演繹，推導出所需命題的數學方法。但是在提出這組公設之前，我們必須對所謂的「受訪者之意見強度」，做以下幾點說明：

- A1. 本文假定受訪者的意見強度可以用一個介於 0 與 1 之間的實數來代表。亦即，如果以 z_i 代表第 i 種意見強度，則 $z_i \in [0, 1]$ 。
- A2. 如果問卷所設計的選項，依據意見之強度分成 K 級。這 K 級之意見，假定從 L_1, L_2, \dots ，依序排列到 L_K 。例如，選項中的 L_1 代表「非常贊成」， L_2 代表「贊成」， \dots ， L_K 代表「非常反對」。則本文假定每一對相鄰之選項，其所間隔的意見距離相等，皆為 $1/(K-1)$ 。此外，任何兩個不相鄰的選項 L_m 與 L_n 之間的意見距離則為 $|m-n|/(K-1)$ 。因此， L_1 與 L_K 的意見距離就等於 1，此乃最大的意見距離。
- A3. 本文令選項 L_1 的意見強度為 0， L_2 的意見強度為 $1/(K-1)$ ， \dots ， L_m 的意見強度為 $(m-1)/(K-1)$ ，餘者依此類推。同時並假定受訪者會選擇一個最接近其意見強度的選項作答。例如當 $K=5$ ，而受訪者的意見強度為 0.6 時，此一受訪者會選擇 L_3 作答。因為此時 L_3 所代表的意見強度等於 0.5，而 L_4 所代表的意見強度則為 0.75，顯然 L_3 比較靠近 0.6。所以受訪者所選擇的答案是 L_3 。
- A4. 共識程度的高低受到意見距離的影響，但是與意見的種類沒有關係。例如，當全部的受訪者都表示「非常反對」時，其所表現的共識程度，不應該與全體皆表示「非常贊成」時有所不同。所以當問卷設計者將選項分成 K 級時，意見強度可以是依序遞減，也可以是依序遞增，我們將不容許這兩種排列方式上的差異，對結論產生實質的影響。換言之，我們可以令 L_1 代表「非常贊成」，令 L_K 代表「非常反對」；也可以顛倒過來，以 L_1 代表「非常反對」，以 L_K 代表「非常贊成」，而結論必須不變。⁵

以上四點說明乃是我們在發展共識度指標以前，必須建立的「共識」。本文將它們當作基本假設看待。

3.2 分配與共識度之衡量

現在讓我們引介一些基礎的定義與數學符號。

在針對一項問題展開社會調查之後，回收的資料包括受訪者的意見，以及各種不同意見強度之下的統計人數。這樣的一組資料，稱為一個分配，以 $(\pi, z) \equiv (\pi_1, \dots, \pi_K; z_1, \dots, z_K)$ 代表之。其中 π_i 是意見強度等於 z_i 的人數，所以受訪者的總人數即為 $\sum_{i=1}^K \pi_i$ 。此外， $z_i \in [0, 1]$ ，如前所述。而對所有的 i, j 而言， $z_i \neq z_j$ ，且 $\pi > 0$ 。

準此，所謂的共識度之衡量 (consensus measure)，就是一個從分配之空間 D 反映到實數集 R 的映射 S ，若以數學表示，即為 $S: D \rightarrow R$ 。上述之定義，也可以用白話加以表達：「所謂的共識度之衡量，就是如何以一個實數，具體反映一個分配所代表的共識程度」。

此一衡量之重要性在於：如果我們能夠畢竟其功，則一個分配的共識程度，就可以用一個實數來代表，而此一實數的大小，即代表共識程度之強弱。

3.3 共識度模型

一個分配的共識程度究竟有多強，並不是一個容易回答的問題。但是我們可以先驗地假定，共識程度的強弱，受到兩種因素的影響。第一種因素是認同感 (identification)；第二種因素是疏離感 (alienation)。⁶

所謂的認同感，就是在一組意見強度相同的人群之中，個體與個體之間，相互所激發出來的凝聚力。認同感愈強，共識程度也就愈高。至於認同感的強弱，則受到人數的影響。因為在同一群人之中，

每個成員的認同感必然隨著同志的數量而變化。當同夥的人數愈眾時，成員的認同感也就愈強。所以意見強度為 z_i 的一群人（簡稱為第 i 群人），其成員的認同感，應該是這一群人總數 π_i 的函數。所以認同感可以表示為 $I(\pi_i)$ 。

影響共識程度的另一項因素是疏離感。所謂的疏離感，指的是因為意見之差距，所造成的心理距離。疏離感愈強，共識程度也就愈弱。現在若以 $\delta_{i,j}$ 代表 z_i 與 z_j 之間的距離，亦即 $\delta_{i,j} = |z_i - z_j|$ ，則我們可以定義第 i 群人與第 j 群人之間的疏離感是 $\delta_{i,j}$ 的函數。換言之，疏離感即為 $A(\delta_{i,j})$ 。

經過上文之說明，我們知道共識程度受到認同感與疏離感之影響，而認同感與疏離感，又分別是人數與意見距離的函數。因此不難推論，當第 i 群中的一個代表性成員，面對第 j 群人中的另一個成員時，他所感覺到的共識度應該是 $I(\pi_i)$ 與 $A(\delta_{i,j})$ 的函數，亦即共識度為 $\phi(I(\pi_i), A(\delta_{i,j}))$ ，或簡寫成 $\phi(\pi_i, \delta_{i,j})$ 。準此，我們可以推論任何一個分配的共識程度，應為：

$$S(\pi, z) = \sum_{i=1}^K \sum_{j=1}^K \pi_i \pi_j \phi(\pi_i, \delta_{i,j}). \quad (1)$$

上式之由來，乃是遵循功利主義 (utilitarianism) 的精神。質言之，功利主義者將社會福利定義成爲個人福利的加總，而我們則將社會的共識程度，定義成爲個人共識感之總合。具體而言，(1) 式之來由，其演繹過程如下：

- i) 已知第 i 群中的一個代表性個人，他對第 j 群中的任何一個成員之共識感等於 $\phi(\pi_i, \delta_{i,j})$ 。現因第 j 群中共有 π_j 個人，所以此人對第 j 群全體的共識感就等於 $\pi_j \phi(\pi_i, \delta_{i,j})$ 。

ii) 所謂的「第 j 群」，可以是 K 群中的任何一群，因此第 i 群中的一個代表性個人，他對其他各群人的總合共識感即為

$$\sum_{j=1}^K \pi_j \phi(\pi_i, \delta_{i,j})。$$

iii) 因為第 i 群中，共有 π_i 個人，所以第 i 群人對其他各群人的總合共識感應當等於 $\pi_i \sum_{j=1}^K \pi_j \phi(\pi_i, \delta_{i,j})。$

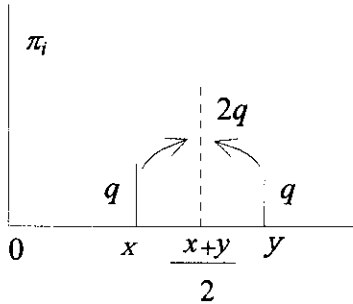
iv) 最後，所謂的「第 i 群」，可以是 K 群之中的任何一群。所以社會的共識程度即為 $\sum_{i=1}^K \pi_i \sum_{j=1}^K \pi_j \phi(\pi_i, \delta_{i,j})。$ 此式經整理後，就是上文的第 (1) 式。

3.4 共識公設

根據前一小節所敘述之模型，一個社會的共識程度就是個人共識感的總和。但是個人的共識感 $\phi(\pi_i, \delta_{i,j})$ ，不過是一個隱函數而已。所以 $\phi(\pi_i, \delta_{i,j})$ 的函數型態必須先確定下來，才有可能計算一個分配的共識程度。

現即藉助公設式分析法，將 $\phi(\pi_i, \delta_{i,j})$ 的函數型態推演出來。首先建立下述五項公設，這些公設的內容將以圖形、文字和數學詳加解釋。

公設 C1 :

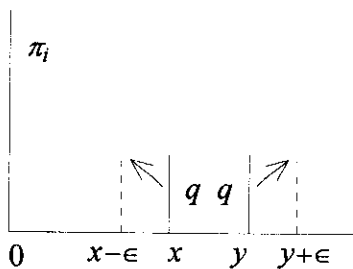


條件： $\pi_i, q \gg 0, 0 < x < y$ 。

內涵：有兩群意見強度分別為 x 與 y 的人，各群人口的總數皆為 q ；如果這兩群人的意見趨於一致，集中在 $(x+y)/2$ 的位置之上，則第 i 群人所感覺到的共識程度不會增加。

說明：這是一項直觀的公設。當右邊兩群人的意見趨於一致時，他們與第 i 群人的平均意見距離並沒有改變，所以第 i 群人所感覺到的共識程度，沒有理由增加。

公設 C2 :

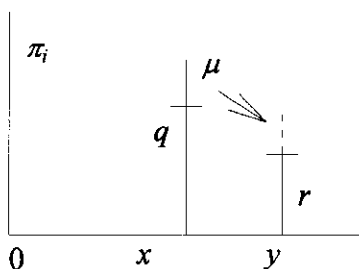


條件： $\pi_i, q \gg 0, 0 < x < y, 0 < \epsilon < x$ 。

內涵：有兩群意見強度分別為 x 與 y 的人，各群人口的總數皆為 q ，如果這兩群人的意見趨於分散，分別移動到 $x-\epsilon$ 與 $y+\epsilon$ 的位置之上，則第 i 群人所感覺到的共識程度不會增加。

說明：這也是一項直觀的公設。當右邊兩群人的意見以等距散開之後，對第 i 群人而言，平均的意見距離並沒有改變；而且向第 i 群靠攏的人，因為 $\epsilon < x$ ，所以事實上並未加入第 i 群人的行列。既然平均的意見距離不變，而且第 i 群人的總數也不變，所以這群人所感覺到的共識程度，沒有增加的道理。

公設 C3：



條件： $\pi_i, q, r \gg 0, 0 < x < y, \mu < q$ 。

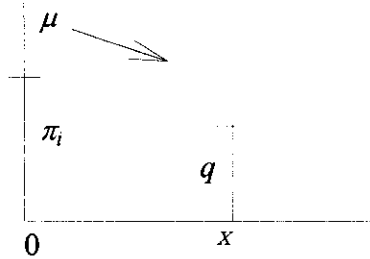
內涵：有兩群意見強度分別為 x 與 y 的人，前者的人口總數為 q ，後者的人口總數為 r ；如果在意見強度為 x 的人群當中，有 μ 人改變立場，加入意見強度為 y 的行列之中，則第 i 群人所感覺到的共識程度會減少。

說明：這項公設其理自明。

公設 C4：如果 (π, z) 分配的共識程度不亞於 $(\hat{\pi}, \hat{z})$ 分配，即 $S(\pi, z) \geq S(\hat{\pi}, \hat{z})$ ，則對所有的 $\lambda > 0$ 而言， $S(\lambda\pi, z) \geq S(\lambda\hat{\pi}, \hat{z})$ 。

說明：這是一項有關齊質性 (homotheticity) 的公設，目的在保證任意兩個分配作等比例的放大或縮小之後，其共識度指標在排序上不會發生逆轉的現象。這項公設經常出現在衡量所得分配不均度的文獻之中，例如 Esteban 與 Ray (1994)，以及 Foster (1985) 都曾經將其當作基本公設運用。

公設 C5 :



條件： $\pi_i, q \gg 0$ ， μ 趨近於 0。

內涵：如果在第 i 群人當中，有 μ 人 (μ 趨近於 0) 改變立場，加入意見強度為 x 的行列，則這兩群人所感覺到的總合共識程度，會產生如下之變化：

- i) 若 $\pi_i < q$ ，則增加；
- ii) 若 $\pi_i = q$ ，則不變；
- iii) 若 $\pi_i > q$ ，則減少。

說明：這項公設乍看之下十分複雜，但是道理卻很平常，它的意思是說：如果有人脫離少數，加入人數較多的一群，則兩群人總合的共識程度會增加；反之，如果有人脫離多數，投靠少數族群而去，則兩群人總合的共識程度會減少。

3.5 主要定理

前述的五個公設，可以演繹出下述定理。

定理 1：設 $\beta > 0$ ， $\alpha \in (0, 1)$ ，則一個分配的共識程度為：

$$S^*(\pi, z) = -\beta \sum_{i=1}^K \sum_{j=1}^K \pi_i^{1-\alpha} \pi_j \delta_{i,j}; \quad (2)$$

若且唯若公設 C1 至公設 C5 滿足。

證明：(必要性) 首先考慮公設 C1。在原分配下，第 i 群人的共識感為：

$$\phi_i^1 \equiv \pi_i q \phi(\pi_i, x) + \pi_i q \phi(\pi_i, y). \tag{3}$$

分配改變以後的共識感為：

$$\phi_i^2 \equiv 2 \pi_i q \phi(\pi_i, \frac{x+y}{2}). \tag{4}$$

現根據此公設並利用 (3)、(4) 兩式，

$$\pi_i q \phi(\pi_i, x) + \pi_i q \phi(\pi_i, y) \geq 2 \pi_i q \phi(\pi_i, \frac{x+y}{2}). \tag{5}$$

因此對任何的 $(\pi_i, q) \gg 0$ 而言，

$$\phi(\pi_i, x) + \phi(\pi_i, y) \geq 2 \phi(\pi_i, \frac{x+y}{2}). \tag{6}$$

由於 $x < y$ ，因此當 x 與 y 十分接近時，上式隱涵 ϕ 對任何的 π_i 而言，必須是 $\delta_{i,j}$ 的非凹性 (nonconcave) 函數。

其次考慮公設 C2。在原分配下，第 i 群人的共識感即為 (3) 式所示。至於分配改變以後的共識感則為：

$$\phi_i^2 \equiv \pi_i q \phi(\pi_i, x - \epsilon) + \pi_i q \phi(\pi_i, y + \epsilon). \tag{7}$$

因此對任何的 $(\pi_i, q) \gg 0$ 而言，公設 C2 要求：

$$\phi(\pi_i, x) - \phi(\pi_i, x - \epsilon) \geq \phi(\pi_i, y + \epsilon) - \phi(\pi_i, y). \tag{8}$$

由於 $x < y$ ，因此當 ϵ 是一個極小的正實數時，上式隱涵 ϕ 對任何的 π_i 而言，必須是 $\delta_{i,j}$ 的非凸性 (nonconvex) 函數。

第三，根據公設 C3，在原分配下，第 i 群人的共識感仍如 (3) 式所示。而分配改變以後的共識感則為：

$$\phi_i^2 = \pi_i(q - \mu)\phi(\pi_i, x) + \pi_i(r + \mu)\phi(\pi_i, y). \quad (9)$$

因此對任何的 $(\pi_i, q) \gg 0$ 而言，公設 C3 要求：

$$\phi(\pi_i, y) < \phi(\pi_i, x). \quad (10)$$

由於 $x < y$ ，所以上式隱涵 ϕ 對任何的 π_i 而言，必須隨 $\delta_{i,j}$ 而遞減。

證明至此我們發現，對任何的 π_i 而言， ϕ 既是 $\delta_{i,j}$ 的非凹性也是非凸性函數，而且遞減。所以 ϕ 是 $\delta_{i,j}$ 的直線型且遞減之函數。此外，由於 $\delta_{i,j} = |z_i - z_j| = 0$ ，因此 ϕ 的函數型態可以進一步確定為：

$$\phi(\pi_i, \delta_{i,j}) = -b\varphi(\pi_i)\delta_{i,j}. \quad (11)$$

其中， b 是一個大於 0 的正實數。

檢查上式後發現，唯一尚待決定的就是 φ 的函數型態。而公設 C4 與公設 C5 可以協助完成這項工作。首先就公設 C4 而言，Esteban 與 Ray (1994) 證明，利用這項公設可以演繹出一條 Cauchy 方程式，並且能夠確定 φ 是一個指數函數，⁷

$$\varphi(\pi_i) = \gamma\pi_i^{-\alpha} \quad \gamma > 0. \quad (12)$$

現在將 φ 代回(11)式，再求 ϕ 對 π_i 的偏微分可得

$$\frac{\partial \phi}{\partial \pi_i} = \alpha b \gamma \pi_i^{-\alpha-1} \delta_{i,j} \quad (13)$$

由於本模型假定，同夥的人數愈眾，認同感愈強，共識度也會隨之而上升，所以上式不能為負值。因此我們可以確定 $\alpha > 0$ ，但是 α 的上下限仍有待決定，而此則有賴公設 C5 輔助之。

根據公設 C5，以及 (11)、(12) 兩式，分配改變之後的總合共識程度為：

$$\Phi(\mu) \equiv -b\gamma x[(\pi_i - \mu)^{1-\alpha}(q + \mu) + (q + \mu)^{1-\alpha}(\pi_i - \mu)]. \quad (14)$$

取 Φ 對 μ 的微分，並且定義 $t \equiv q / \pi_i$ ，則下式成立：

$$\left. \frac{\partial \Phi}{\partial \mu} \right|_{\mu=0} = -b\gamma x \pi_i^{1-\alpha} [(1-\alpha)(t^{-\alpha} - t) - t^{1-\alpha} + 1]. \quad (15)$$

利用數值模擬，以及 $\alpha, t \gg 0$ 的條件，不難證明 C5 若欲滿足，則 $0 < \alpha < 1$ ，亦即， $\alpha \in (0, 1)$ 。最後定義 $\beta \equiv b\gamma$ ，如此即完成本定理必要條件之證明。

(充分性) 以第 (2) 式所定義的 S^* 為準，檢查每一個公設所列表述的情況，結果均能滿足，充分條件因此得證。

本定理得證。

肆、共識度指標的性質

4.1 指標之標準化

共識度指標 S^* 的數值為負，而且這項指標的絕對值，會隨著受訪者的總人數之增加而增加。這兩項性質，使得 S^* 在應用上會遭遇到如下之困擾：

- i) 一般指標通常為正值，如果有一種指標是以負數表示，將會造成判讀上的不方便。例如當甲分配的 S^* 為 -333，乙分配的 S^* 為 -1 時，甲分配的共識程度比乙分配低。這對研究者與閱讀者而言，能否立即會意且適應，不無疑問。

ii) 當受訪者的總人數增加時， S^* 的絕對值會隨之而增加。所以一個分配的共識程度，必然會隨著樣本數的大小而變化。這項性質顯然不合理，而且無法對受訪者總數不同的兩個分配，進行共識程度之比較。

解決上述困擾的途徑，就是要想辦法將 S^* 轉換成爲正值，並且將 S^* 標準化 (normalization)，使其不再受到總人數變動的影響。本節選擇標準化，作爲初步之工作。至於指標轉換的問題，則留待第 4.4 節再作研究。

首先令 $\beta \equiv (\sum_{i=1}^K \pi_i)^{-(2-\alpha)}$ ，再將 β 的數值代入 (2) 式，整理後可得

$$s = - \sum_{i=1}^K \sum_{j=1}^K f_i^{1-\alpha} f_j \delta_{i,j}. \quad (16)$$

標準化以後的共識度指標，我們將它重新命名爲 s 。至於式中的 f_i 則代表第 i 群人佔總受訪人數的比率， $f_i \equiv \pi_i / \sum_{i=1}^K \pi_i$ ，亦即第 i 群人之相對次數 (frequency)。由於新指標 s 係由分配之相對次數計算而得，所以它不會受到總受訪人數增減變化的影響。下文即以此一經過標準化以後的共識度指標 s ，替代原先的指標 S^* ，作爲研討之核心。

4.2 s 指標的上限與下限

s 的上限是 0，此一情況發生在受訪者的意見，全數集中在一點的時候。至於集中在那一點，則無關緊要。⁸ 此驗證本文先前之立論：所謂的完全共識，指的就是全體之意見一致的情況，此時共識度指數應該攀升到它的上限。顯而易見的是 s 滿足這項要求。

至於 s 的下限就不像它的上限一樣，那麼容易決定了。目前我們只知道此一極限值會隨著 K 與 α 而變動 (上限恆等於 0，不會改變)，

卻不清楚它的數值究竟應該等於多少。表 1 所列式的數據，將有助我們說明這項事實。

表 1

分配	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	0	0	0	7	6	5	4	3	2
	0	0	7	0	1	1	2	2	2
	0	7	0	0	0	1	1	2	2
	14	0	0	0	0	0	0	0	2
	0	7	0	0	0	1	1	2	2
	0	0	7	0	1	1	2	2	2
	0	0	0	7	6	5	4	3	2
$s_{.1}$	0	-0.18	-0.36	-0.54	-0.55	-0.56	-0.55	-0.52	-0.46
$s_{.5}$	0	-0.24	-0.47	-0.71	-0.91	-1.07	-1.08	-1.04	-1.01
$s_{.9}$	0	-0.31	-0.62	-0.93	-1.65	-2.26	-2.21	-2.11	-2.20
V	0	0.03	0.11	0.25	0.23	0.20	0.18	0.15	0.11
M	0	0.17	0.33	0.50	0.49	0.48	0.47	0.43	0.38
G	0	0.17	0.33	0.50	0.49	0.48	0.47	0.43	0.38
$P_{1.5}$	0	0.06	0.12	0.18	0.12	0.08	0.05	0.03	0.02
$s_{.1}^*$	1	0.67	0.33	0	-0.03	-0.04	-0.02	0.04	0.14
$s_{.5}^*$	1	0.67	0.33	0	-0.29	-0.51	-0.52	-0.47	-0.43

在表 1 中，總共有九項分配，每一個分配皆有十四個樣本點，分成七組資料 ($K = 7$)。分配 (1) 的樣本全數集中在中點；從分配 (2) 開始，樣本分成兩群，逐漸離散；至分配 (4) 時，兩群樣本的離散程度

達到極限，此時在兩個端點之上，各有七個樣本點，形成一個兩極分配。我們稱此一過程為「分散模式」。

相對於分散模式，從分配 (4) 的兩極分配變化至分配 (9)，則是一種持續的集中過程，簡稱「集中模式」。在此一模式下，樣本逐漸由兩端向中點滑動，至分配 (9) 時，形成一個齊一分配，此時各點的樣本數一律相等。

針對這九項分配，我們在 $\alpha \in (0,1)$ 的限制之下，設定了三個數值，分別為： $\alpha = 0.1$ ， $\alpha = 0.5$ ，以及 $\alpha = 0.9$ 。並且計算其所對應的共識度指數，結果列示在表 1 各分配之後的第一欄至第三欄之中。檢查這些數據，我們發現下述之事實：

- i) 當樣本全數集中的時候，共識度指數最高，此時 $s = 0$ ，並且不受參數 α 變動之影響。
- ii) 當分配按照「分散模式」變化之際，共識度指數隨之而下降（即負值愈變愈大）；但是當分配按照「集中模式」變化之時，共識度指數是先降後升，而非單調的 (monotonic) 遞增或遞減。
- iii) 在不同的 α 值之下，兩極分配與齊一分配，皆非共識程度最低的分配。
- iv) 當 $\alpha = 0.1$ 與 $\alpha = 0.9$ 時，型態為 (5, 1, 1, 0, 1, 1, 5) 的第 (6) 種分配，是共識程度最低的分配。但是當 $\alpha = 0.5$ 的時候，共識程度最低的分配由第 (7) 項分配取代之，其型態為 (4, 2, 1, 0, 1, 2, 4)。據此推測， s 的下限乃隨參數 α 之數值而變動。
- v) 兩極分配與齊一分配的共識程度，孰高孰低，並無定論。當 $\alpha = 0.1$ 時，兩極分配的共識程度較低；反之，當 $\alpha = 0.5$ ，以及 $\alpha = 0.9$ 的時候，兩極分配的共識程度較高。

根據上述之說明，顯然我們對於 s 的下限，所知極為有限。換言之，何謂「完全沒有共識」，乃是一個無法明確定義的概念。但是這並不表示，我們不能臆測，何種類型的分配會讓 s 降到臨界點。

事實上，從 s 指標的組成和原理，不難推論一個分配的 s 值會隨著該分配的兩極化與齊一化之程度而變化。質言之，一個分配的兩極化程度愈高，疏離感愈強，而疏離感增強則會降低共識程度；另一方面，如果一個分配的齊一化程度愈高，則認同感愈低，而認同感愈低，共識度愈小。因此，如果有一個分配既是兩極分配也是齊一分配，則該分配就是一個「完全沒有共識」的分配。

問題是只有當 $K = 2$ 的時候，才可能出現這種兩極兼齊一的分配。此時，樣本均分成兩群（所以是齊一分配），分佈在僅有的兩點之上，於兩極對峙（所以是兩極分配）。

經過計算以後發現，這種分配的 s 指數等於 $-2^{(\alpha-1)}$ 。因此當 α 趨近於其上界 1 的時候， s 指數趨近於 -1；而當 α 趨近於其下界 0 的時候， s 指數趨近於 -0.5。換言之， s 的下限乃隨參數 α 的數值而變動，而在 $K = 2$ 的情況下，最小的下限趨近於 -1。

但是當 $K > 2$ 的時候，兩極化與齊一化就不可能並存了。以 $K = 3$ 為例，如果是齊一分配，則各點所當配置的樣本，應為總樣本數的三分之一。反之，如果是兩極分配，則全部的樣本應該平均配置在兩個端點之上，而中點則不能有任何樣本。很顯然地，在此狀況下，齊一分配與兩極分配，並不相容。同理可證，在 $K = 4, 5, 6, \dots$ 的場合，也不可能出現這種兩極與齊一性共存的分配。

準此，我們的推測是：當 $K > 2$ 時，共識程度最低的分配，應該是一種介於兩極分配與齊一分配之間的一種分配。換言之，這種情況應該發生在「集中模式」的過程之中，例如表 1 中的第 (6) 種分配，或第 (7) 種分配。同時必須注意的是，當 $K > 2$ 時， s 的數值可以低於

-1。⁹所以 -1 雖然是 s 在 $K=2$ 時的最小下限，卻不是其他情況的下限。

4.3 s 指標與離散度指標、不均度指標之關係

在第 2 節中，我們曾經大略討論過共識度與離散程度的關係，現則進一步將分析之重點，聚集在 s 指標與變異數 V ，均互差 M ，基尼係數 G ，以及 Esteban 與 Ray (1994) 的兩極化指標 P 之關係上。

設以 θ 代表 z_i 的平均數， $\theta = \sum_{i=1}^K f_i z_i$ ，則各項指標的定義如下式所示：

$$V = \sum_{i=1}^K f_i (z_i - \theta)^2,$$

$$M = \sum_{i=1}^K \sum_{j=1}^K f_i f_j \delta_{i,j}, \quad (17)$$

$$G = \frac{M}{2\theta},$$

$$P = \sum_{i=1}^K \sum_{j=1}^K f_i^{1+a} f_j \delta_{i,j}, \quad a \in (0, 1.6).$$

比較 (16)、(17) 兩式，並且利用表 1 的數值，可以綜合整理出以下四點結論：

- i) [s 與 V 之關係]：從數學定義式的型態而言， s 與 V 並無直接之關連。此外，變異數最大的分配是兩極分配 (即第 (4) 項分配)，但是該分配並非共識程度最低的分配，除非 $K=2$ 。所以各種分配如果按照 V 的大小排序，則其順序不會與按照 s 由小到大所排出來

的順序相同。¹⁰ 由此可見， s 指標的功能與變異數有別，所以 V 無法取代 s 指標。

- ii) [s 與 M 之關係]：從這兩個指標的數學定義，可以發現，如果 α 趨近於 0，則 s 趨近於 $-M$ 。亦即， $\lim_{\alpha \rightarrow 0} s = -M$ ，所以 s 是一種廣義的均互差。更具體地說，由於認同感對共識度的影響反應在 α 的數值之上。所以當 $\alpha = 0$ 時，即表示認同感對共識度不發生影響，因此均互差實質上就是一種沒有將認同感納入考慮的共識度指標。由此引伸，不難理解 s 指標的 α 值為什麼必須限定在大於 0 的區間之內。¹¹
- iii) [s 與 G 之關係]：透過第 (17) 式中 G 與 M 之關係，¹² 立即得證： $\lim_{\alpha \rightarrow 0} s = -2\theta G$ 。所以 s 與 G 的關係與前述之 s 與 M 之關係相類似，於此不再贅述。¹³
- iv) [s 與 P 之關係]：兩極化指標與共識度指標的數學型態接近，兩者同時將認同感和疏離感的因素考慮在內，而兩者最主要差別則在於 α 係數與 a 係數的值域不同。所以兩極化程度最強的分配，就是兩極分配。¹⁴ 但是根據前文之說明，除非 $K=2$ ，否則最缺乏共識的分配不會是兩極分配。事實上，這兩項指標的異同，絕對是一項值得深入探索的問題，但是篇幅所限，進一步的研究有待另文繼續。惟需聲明的是，本研究受 Esteban 與 Ray (1994) 之影響甚為深遠，所以 s 指標無疑地是 P 指標的羽裔。

4.4 理想的指標 —— s^* 指標

前文曾經提到：「一般指標通常為正值，如果有一種指標是以負數表示，將會造成判讀上的不方便。」遺憾的是， s 指標正是一種以負數表示的指標。現在如果想要彌補這項缺陷，最理想的作法，莫過於進行指標的轉換。轉換的方法如下。

設若我們知道 s 指標的上限與下限，例如 $s \in [\underline{s}, \bar{s}]$ ，其中 $\underline{s} < \bar{s} \leq 0$ ，如此則不難衍生出另外一個指標，

$$s^\circ = 1 - \frac{\bar{s} - s}{s - \underline{s}}. \quad (18)$$

上項指標 s° 具備以下兩個優點：

- i) s° 是 s 的直線型轉換，所以 s° 能夠保留 s 的全部特性，不會發生扭曲的現象。
- ii) 經過轉換後，共識程度的數值會從負值，自動轉變成為一個介於 0 與 1 中間的實數，亦即 $s^\circ \in [0, 1]$ 。當一個分配的 s 到達它的上限 \bar{s} 時， s° 的數值也會升達其上限，此時 $s^\circ = 1$ 。反之，當 $s = \underline{s}$ 時， s° 則會降到其下限，此時 $s^\circ = 0$ 。因此 s° 是一個意義明確，界定清楚的指標。

新指標 s° 固然具備上述之優點，但是 (18) 式的轉換，係以 \underline{s} ，以及 \bar{s} 為已知作前提。唯所遺憾的是，我們知道 $\bar{s} = 0$ ，但是 \underline{s} 的數值，誠如第 4.2 節之分析，目前仍不清楚。所以 s° 在應用上，距離成熟仍有一段里程。

所幸的是，我們可以另闢蹊徑，建立下述之指標，

$$s^* = 1 - \frac{\bar{s} - s}{s - s_p}, \quad s_p \equiv -2^{(\alpha-1)}. \quad (19)$$

這項新指標係將 (18) 式的 \underline{s} 代之以 s_p 而成的， s_p 是一個兩極分配的 s 值，如上式所示。

s^* 指標的值域及其特性，如下文所述：

- i) 若為完全共識之分配，則 $s = \bar{s}$ ，所以 $s^* = 1$ ，此乃 s^* 的上限；
- ii) 一個分配的共識程度如果介於兩極與完全共識之間，則 $s_p < s < \bar{s} = 0$ ，因此 $0 < s^* < 1$ ；
- iii) 若為兩極分配，則 $s = s_p$ ，因此 $s^* = 0$ ；
- iv) 當一個分配的共識程度低於兩極分配時， $s < s_p < 0$ ，因此 $s^* < 0$ 。

綜合而論， s^* 乃是一個以兩極分配作為正負值分界點的指標，凡共識程度低於兩極分配者，其 s^* 值小於 0；凡共識度高於兩極分配者，其 s^* 值大於 0，而最大的 s^* 值等於 1。

s^* 是一個理想的指標，因為 s^* 乃是 s 指標的直線型轉換，故能保留原指標的性質，不會因轉換而失真。另一方面， s^* 雖然會變成負值，¹⁵ 但是正負值的分界點以兩極分配為準，意義既清楚又明確，所以我們建議社會學者在衡量共識程度時，應當採用 s^* 指標。

伍、結 論

結論這一節通常是一篇論文的闌尾，結論在正文的每一個角落，沒有人會將結論只寫在結論這一節裡。

如果一定要結論，我的結論是：社會調查是現代占星術，預言的根據；希望本研究所建立的共識度指標 s^* ，能夠幫助我們一眼看穿紛亂年代下的台灣社會意向。

註 釋

- 1 此一軟體係由臺大經研所博士班研究生林欣吾研發而成，特此致謝。
- 2 經濟學家對於以公設式分析法推演指標的研究並不陌生，例如：Encaoua 與 Jacquemin (1980) 的產業集中度指數 (concentration index) 研究，Fields 與 Fei (1978) 的所得不均度研究，以及 Esteban 與 Ray (1994) 的兩極化指標 (polarization index) 研究等，皆屬經典之作。
- 3 衡量所得不均度的指標很多，包括著名的吉尼係數在內。
- 4 均互差的性質參見 Kendall 與 Stuart (1963)，其數學定義見本文第4.3節的(17)式。
- 5 但是不論真實的問卷將「沒有意見」這個選項放在第幾位，在本研究之架構下，它必須被當作 K 級選項中的中位選項處理，因為如此方能滿足「選項係按照意見之強度排序」的要求。換言之，「沒有意見」就是不贊成也不反對，所以它的意見強度必須介於最低度的贊成與最低度的反對之間。
- 6 此乃模仿 Esteban 與 Ray (1994) 在發展兩極化指標時所應用的觀念。
- 7 這項證明的數學根據參見 Aczél (1966, p. 41) 的‘Theorem 3’。
- 8 檢查第(16)式，立即得證。
- 9 檢查表 1 第三欄中的 s 值，即可發現此一事實。
- 10 比較表 1 前三欄與第四欄的數值，即可驗證這項性質。
- 11 表 1 第五欄所列示的就是各項分配的 M 值。若與前三欄的 s 值相比，不難發現根據 M 以及根據 s 所做的排序並不相同。
- 12 吉尼係數與均互差之關係，參見曹添旺、賴東昇、陳昭南 (1979)，Theil (1967)，Kendall 與 Stuart (1963)，以及 Gastwirth (1972) 等文獻之討論。
- 13 G 值列示在表 1 的第六欄中，由於表中各項分配的平均數皆等於 0.5，所以 G 值等於 M 值。
- 14 參見 Esteban 與 Ray (1994, pp. 837-40) 的‘Theorem 2’所做之證明。
- 15 參見表 1 第八欄及第九欄的數字。

參考資料

曹添旺、賴東昇、陳昭南

- 1979 《廣義的吉尼係數：所得有負時的吉尼係數及其修正》。臺北：中研院三民所專題選刊(22)。

Aczél, J.

- 1966 *Lectures on Functional Equations and their Applications*. New York and London: Academic Press.

Encaoua, D. and A. Jacquemin

- 1980 "Degree of Monopoly, Indices of Concentration and Threat of Entry," *International Economic Review* 21: 87-105.

Esteban, J. and D. Ray

- 1994 "On the Measurement of Polarization," *Econometrica* 62: 819-851.

Fields, G. S. and J. C. H. Fei

- 1978 "On Inequality Comparisons," *Econometrica* 46: 303-316.

Foster, J.

- 1985 "Inequality Measurement," in *Fair Allocation*, ed. by H. P. Young. Providence: American Mathematical Society.

Gastwirth J. L.

- 1972 "The Estimation of the Lorenz Curve and Gini Index," *Review of Economics and Statistics* 54: 306-316.

Kendall M. G. and A. Stuart

- 1963 *The Advanced Theory of Statistics*. London: Charles Griffen and Company.

Theil H.

- 1967 *Economics and Information Theory*. Chicago: Rand McNally and Company.

模糊統計在社會調查分析的應用*

吳柏林** 楊文山***

壹、前言

社會科學中的實地調查方法，常以受訪者填答問卷，或是經由面對面、電話訪談，而由訪員填寫問卷，做為收集社會科學經驗資料的一種重要工具(Sudman and Bradburn, 1982)。但是，在社會科學理論中所提出的假設，定義的範圍(concepts)，及其發展成可以測量的經驗指標(empirical indicators)，由於遣辭用句、字眼順序以及問題型式迥異，問卷的設計可能會產生相當多的歧異，並進而影響到研究結果的信度與效度(Schuman and Presser, 1981；楊文山與蔡瑤玲，1995)。另

* 本論文所使用之部份資料係採自行政院國家科學委員會所贊助之「臺灣地區社會意向基本調查」(計劃編號 NSC81-0301-H-001-46-B1) 81年6月經濟組專題調查大臺北地區之預試問卷。該計劃係由中央研究院中山人文社會科學研究所執行。在此感謝參與此次研究計劃的諸位教授；中央研究院社科所林忠正教授對模糊問卷的設計所提出的寶貴意見，並提供測試此一問卷的機會，在此特別表示感謝之意。而二位匿名評審教授對於本文之數據以及文章所提供的寶貴修正意見，也使文章內容更加出色並精確。在此也感謝政治大學統計所研究生黃麗鳳同學在資料處理以及運算過程中所付出的時間與努力。社會意向研究助理對於收集資料所付出的辛勞，謹此一併致謝。本文之內容與解釋由作者負責。

** 政治大學應用數學系教授

*** 中央研究院中山人文社會科學研究所研究員

外，與自然科學研究對象不同的是，社會科學中的經驗指標，與經驗事實的相對關係並非一成不變，社會現象會因時、空、情境的不同，而產生很大的變化及差誤，若未能以更客觀的調查研究方法，很容易誤導研究結果。

社會調查過程中，社會科學家總是使用某些固定或清晰的回答模式測度喜好程度，並以整數所衍生的名義尺度記錄此評量和行為結果間的統計相關性 (Manton, Woodbury and Tolley, 1994)。然而，社會學中所涉及的概念常常是模糊 (fuzzy)，而非確定、清楚的。比方，在一個典型的調查中，受訪者被問及去年一年中的總收入 (total income) 是多少；又比方，受訪者被問及「理想子女數」，「支持政黨程度」，以及「老人的年齡標準和身份」等問題時，這些問題並沒有清楚的界定，受訪者在面對這些問題時可能在「贊成」與「不贊成」間產生許多不同程度的「贊成」與「不贊成」的答案。

但是，當社會學家從事社會經驗性研究時，以上的許多問題，常常以某些固定的認定做為回答問題的依據 (Li, 1989)。老人的年齡和身份認定就是一個相當有趣且值得探索的問題。65歲以上是否就可以算定是接受老人年金的年齡，或是更年輕一點，或是更年長一些，或是在60歲到70歲之間，但依老人身體及家庭收入狀況而定，就是一個相當有趣的經驗性政策問題。可是，當社會科學家設計問卷時，很容易把此一本質模糊的問題設計成「你贊不贊成政府開辦老年年金保險制度？」(臺灣地區社會意向調查，83年2月定期調查執行報告)，或「你是否贊成／不贊成65歲以上老人接受老人年金」或「不知道」做為可能的答項。所以，當問題本身的經驗指標可能是模糊，或不一定的狀況時，如果用結構性的問卷設計，並且延用清楚 (crisp) 的數學模式做為分析的依據時，就可能導致誤差，產生信度、效度的偏誤，影響研究的可能結果。

在這篇論文中，我們介紹一個模糊統計調查，與分析模糊資料的方法。並將此一模糊統計與分析方法實際地應用在社會調查中，從社會與經濟問題來分析人類思考與行為模式。用以驗證此一模式在測量人類模糊思維的可能性及可行性。這新的方法稱為模糊邏輯模式理念 (the fuzzy logic model of beliefs)。模糊邏輯的應用，主張個人喜好程度不需非常清晰或是有條理。所以，模糊邏輯與布林 (Boole) 模式理念大相逕庭，布林模式中 A 和 B 的比較只有三種結果： $A > B$ ， $A < B$ ， $A \sim B$ 。但人的思維是複雜的，具有不明確的喜好。因此能對真實狀況做全盤描繪的構想，早已成為某些邏輯學家的目標。換句話說，對人類而言，模糊模式比直接指定單一物體一個值，較合適於評估物體間的相關特性。由於其他特性有助於意見的評定，因此使用模糊邏輯時，必須對其他特性也加以說明，以便將人們的喜好程度轉換成便於計算的效用函數 (utility function)。

對那些人類在思考認知不易表達完善的問題，藉隸屬度函數與模糊統計分析似乎可提供更明確表達。本文亦架構幾個重要的模糊統計量的定義，如模糊期望值，模糊眾數，模糊秩階…等。最後我們做幾個實證調查，希望藉模糊問卷調查及模糊統計分析，對人類在思想認知不易表達之問題，能有更合理的測度與解釋。

模糊邏輯彌補了布林邏輯的不足，但並沒有取代布林邏輯，兩個方法可以互存。其關鍵在於每一理論各描述不同的理念。而模糊子集理論 (fuzzy set theory) 的主要特質，在於對於無界的可能偏好，能做出無限的解說。使用模糊邏輯來測度，必需要使用模糊統計代替傳統統計。而模糊邏輯測度主要根基於模糊集合的基本理論。由於模糊集合的基本理論，相當佔據篇幅，所以在下面的模糊測量中將做一簡要的說明，在此不做詳細介紹，有興趣的讀者可參考 Klir & Folger (1988), Li (1989), Zimrmann (1991) 或 Manton, Woodbury and Tolley (1994)。

貳、模糊邏輯與測度

首先，我們簡述有關模糊集合理論在社會科學研究上的應用。基本上，模糊邏輯模式理念的發展，源起於加州柏克來系統控制學家 L. A. Zadeh (1965) 所提出的模糊集合理論，對於測量模糊概念所做的重要貢獻 (McNeill and Fredberger, 1993)。Zadeh 以為過去數學理論中的集合論 (set theory) 和人思維的最大不同在於數學的精準性及排它性。在傳統集合理論中，若 x 為 A 集合中的任一元素，則 $x \in A$ ，而 x 具有某種性質，若 $x \notin A$ ，則 x 不為 A 集合中的元素，且不具有此一性質，其它性質則不可能同時存在。但集合論忽略了一項重要的事實，那就是 X 可能存在於 A 或 \bar{A} 的範圍中，而不固定於一點上，而是一項程度上的隸屬性。

模糊理論乃擴展傳統二值邏輯的觀念與運算方法，利用隸屬度函數 (membership function) 來表示事物的模糊現象，並對於不同的考量賦予特別的運算。對社會科學家而言，有時應用模糊數學方法與模式來表示可能較直接給定單一的特定值來得適切些，並可能較合於用來評估變數與變數之間的相關性。此外，由於其他因素往往也會影響人們對某一事物的評定，因此使用模糊邏輯時，必須對所謂的「其他因素」加以說明，以便將人們的喜好程度轉換成便於計算的隸屬度函數。

要對模糊的現象予以適當的集合描述，必須應用新的邏輯假設，即論域上的對象由屬於某一集合到不屬於某一集合是漸進而非突變的。因此，隸屬度函數的設定在模糊理論的測度與分類上，扮演相當重要的角色，其可定義如下：設 A 是 X 論域 (或定義域) 上的一個模糊子集 (論域指被討論之事物全體的變數，有時亦稱為空間)。 A 的隸屬度函數以 μ_A 表示，乃指滿足 $\mu_A: X \rightarrow [0,1]$ 的對應關係。此 μ_A 的值反映了 X 中的元素 x 對於 A 模糊子集的隸屬程度。 $\mu_A(x)$ 的值接近於 1，表示 x 隸屬於 A 的程度很高。 $\mu_A(x)$ 的值接近於 0，表示

x 隸屬於 A 的程度很低。利用這些介於 $[0,1]$ 的實數，即可標記人們對被討論之事物的看法。

在模糊集合中的元素，可能只有部份屬於某集合。例如，「年輕人」這一名詞，幾歲以上不再年輕，並無一個確切的界限來區分年輕和不再年輕。而模糊集合的定義，可以顯示出一個20歲的人是「90%的年輕」，當某一個60歲的人相對的只有「30%的年輕」。在模糊集合中，隸屬度的全距從 0到 1。因此，每個語言變項，比方「年輕」，代表一個可能性分布。此分布的均值表示人類真實評定是「年輕人」的數值。而關於分布的評定，可能因人而異，且不需為常態。

在社會學及經濟學的研究中，研究對象的答案很少是確切的真或偽。若要試著去分析人類理念，必會遭遇到行為的不確定性。而連續區間值具有能處理亦真亦偽狀況的能力，因此應用區間值的模糊特性所做的分析調查，可使研究者處理不確定性，確實是符合實際狀況的測量工具。

2.1 模糊隨機變數

在討論模糊分析調查之前，我們先給予模糊隨機變數的定義及基本觀念：

定義 2.1. 令 $Q(\mathfrak{R})$ 為實數軸 \mathfrak{R} 上所有模糊集合類的標準集合表象 (normal set representation)。若 \mathfrak{R} 的子集之一家族系統 $\{A_\alpha(\omega) \mid \omega \in \Omega, \alpha \in (0,1)\}$ 具有下列性質：(i) 對所有 $\omega \in \Omega, \{A_\alpha(\omega) \mid \alpha \in (0,1)\}$ 為 $X_\omega \stackrel{def}{=} X(\omega)$ 的標準集合表象。(ii) 對所有的 $\alpha \in (0,1)$ ，映射 A_α 和 \bar{A}_α 是 $u \rightarrow \bar{B}_1$ -可測的。(μ 是 σ -代數； \bar{B}_1 是 $\mathfrak{R} \cup \{-\infty, +\infty\}$ 上的柏瑞爾 σ -代數 (Borel σ - algebra)；此處我們定義對所有的 $\alpha \in (0,1)$ 和 $\omega \in \Omega$ ， $\underline{A}_\alpha \stackrel{def}{=} \inf A_\alpha(\omega)$ ； $\bar{A}_\alpha \stackrel{def}{=} \sup A_\alpha(\omega)$ 。則稱映射 $X: \Omega \rightarrow Q(\mathfrak{R})$ 稱為模糊隨機變數 (fuzzy random variable)。

欲應用傳統的數學邏輯觀念，明確定義一模糊隨機變數並不容易。此處僅提供一比較合理與符合經驗法則參考。目前對於模糊隨機變數、模糊統計量之定義，文獻上尚無一致的共識，有興趣之讀者可參考 Kwakernaak (1978)、Kruse & Meyer (1987)、或 Maton, Woodbury & Tolley (1994)。

2.2 模糊假設檢定

古典的統計檢定必須陳列明確的假設。比方，當我們想檢定兩母體平均數是否有差異時，虛無假設是「兩個平均數確實相等」。然而，有時我們想要知道的只是兩平均值是否非常逼近，此時傳統的檢定方法並不適用於這種包含不確定性的假設檢定。

Casals *et al.* (1986, 1989) 曾以模糊事件表徵的模式為基礎來討論統計假設檢定；Saade & Schwarzlander (1990) 針對混合資料 (hybrid data) 提出模糊假設檢定；Miyamoto (1990: 240) 也曾論及模糊假設檢定。但他們對模糊假設檢定方法的構思不但基於資料的隨機不確定性，也包含了資料中的非隨機不確定性。而「兩母體平均數幾乎相等」的假設檢定，只討論包含於假設中的非隨機不確定性，這樣的不確定性不但與資料本身無關，且無法假設於資料上。因此 Casals *et al.* (1986, 1989)，Saade & Schwarzlander (1990)，甚或貝氏近似所提供的架構都無法處理上述的假設檢定。

Watanabe & Imaizumi (1993) 將 Neyman & Pearson 設立的古典假設檢定擴展，提出了上述模糊假設的模糊統計檢定 (the fuzzy statistical test)，而用此方法檢定所得的推斷結論也是模糊的。在第三節中，我們試著利用模糊語言系統和模糊統計從事問卷調查，來分析和說明人類的理念模式。

參、模糊統計分析

將模糊統計分析運用於調查領域來測度理念，並不受限於特定理論。Sugeno (1977) 提議當客觀性伴隨著主觀性時，系統變得模糊。他並強調判斷一系統是否為模糊。也就是說，判斷系統是由一確定值或一模糊區間所構成，有其必要性。

通常利用確定數值所做的傳統抽樣調查分析的缺點包括：

- (1) 人類的思考與行為本來充滿著模糊過程，傳統問卷的數字常被過度解釋。
- (2) 為迎合數字的精確需求，實驗資料常有被過度使用之嫌。
- (3) 為了簡化或降低數學模式的複雜性，卻將實際狀況之相關與動態特質忽略。

而以模糊區間估計人類喜好，是一種描繪人們對字句的思維或感覺的方法。因此凡是牽涉到人的思慮的問題，用模糊觀念處理較合適。本論文即對抽樣調查所獲得的傳統和模糊統計樣本做進一步的統計分析與探討。

3.1 模糊統計

一般人在日常生活中常對可能發生的事進行推論。例如，一個決策的過程中，某些重要因素的考量，都可能是下決策的關鍵。所以當我們考慮的因素愈多愈詳密，愈能確信最後所做的決策。模糊統計資料就是考量了人們擁有多重喜好，而以模糊區間和加權模糊區間來進行討論。但因個人認知的喜好程度不同，而具有不定等長的期望區間。下面我們提出一些關於模糊統計量的定義：

定義3.1. 令 $F_i = (a_i, b_i)$ 是樣本模糊區間， $i = 1, 2, \dots, n$ ； n 為樣本數。則樣本 $\{F_i\}$ 的期望模糊區間 (the expected fuzzy interval) 上下

界分別是樣本模糊區間上下界值的個別加權和。換言之，期望模糊區間 (a, b) 中， $a = \frac{\sum a_i}{n}$ (期望最小值)， $b = \frac{\sum b_i}{n}$ (期望最大值)。

定義3.2. 樣本 $\{F_i\}$ 期望模糊區間的平均數 (average of the expected fuzzy interval) 定義為 $uF = \frac{a+b}{2}$

定義3.3. 樣本 $\{F_i\}$ 模糊區間的中位數 (median of the fuzzy interval) 定義為 $median_F = (m_l, m_u)$ ，其中 m_l 為 $\{a_i\}$ 的中位數， m_u 為 $\{b_i\}$ 的中位數。

定義3.4. 樣本 $\{F_i\}$ 模糊區間的眾數 (mode of the fuzzy interval) 定義為 $mod e_F = (m_L, m_U)$ ，其中 m_L 是 $\{a_i\}$ 的眾數， m_U 是 $\{b_i\}$ 的眾數。

定義3.5. (模糊排序) 樣本 $\{F_i\}$ 模糊區間的等級定義為 $Rank(F_i) \leq Rank(F_j)$ ，如果 $\frac{a_i+b_i}{2} \leq \frac{a_j+b_j}{2}$ 。當 $\frac{a_i+b_i}{2} = \frac{a_j+b_j}{2}$ 時，我們稱 $Rank(F_i) = Rank(F_j)$ ，即 F_i 與 F_j 形成相持 (tie)。

3.2 模糊問卷調查表

在社會調查中，若欲由喜好程度的資訊得知個人的理念，可以分別以傳統的調查表得到受訪者行為意圖的點估計，另由模糊問題方式得到模糊統計估計，以便兩者加以比較。例如，探討的問題為：

「在現代社會中(如臺灣)你預期男性和女性應受幾年教育?(請給一個確定的年數)」

要求受訪者回答一個確定的數字，此為傳統抽樣調查的問題形式。而調查表中另一問題詢問：

「在現代社會中(如臺灣)你預期男性和女性應受幾年教育?(請給一個範圍)」

這樣的問題要求受訪者給定一個區間來回答模糊形式的問題。

以上所介紹的是獲得模糊屬量資料以及傳統資料的問卷形式。至於屬性資料的搜集，傳統的問法是要要求受訪者回答一個確定的等距尺度，屬單一邏輯。例如評量「對生活是否有枯燥不快樂的感受」，問卷形式可以為

1.沒有感受 -----> 5.感受很強

「覺得生活枯燥不快樂 _____ . _____ . _____ . _____ . _____ .」

而應用模糊理論，語意與模糊統計所設計的問卷，運用可複選的形式符合實際狀況，可能可以得到較精確的結果。其問答方式為

「覺得生活枯燥不快樂 沒有感受 _____ . 感受很強」

要求受訪者在 _____ 線段上劃下感受範圍(區間)，比方：稍為感受可記為 ////// ；有時很強有時普通 ////// 。用如此的回答形式來獲得模糊感受程度。若對分配程度較具概念者，我們亦可引導他們以對事件感覺之隸屬度值，分別填入。例如：

等級	1	2	3	4	5	模糊感受程度
覺得生活枯燥不快樂		0.6	0.4			稍微感受

等級	1	2	3	4	5	模糊感受程度
覺得生活枯燥不快樂				0.2	0.8	感受很強

等級	1	2	3	4	5	模糊感受程度
覺得生活枯燥不快樂	0.5				0.5	有時感受很強有時沒有感受

以下，我們將以實際問卷調查所獲得的資料，比較傳統和模糊問卷形式所提供之資訊。其中實例研究一至三，乃截取中央研究院中山人文社會科學研究所1992年6月以隨機抽樣法，抽調100位大台北都會地區(包括台北市與台北縣)的民眾，所做的「臺灣社會意向調查」研究之預試問卷調查資料。其中有98份完整的回收調查表，成功率为98%；實例研究四，乃截取北部某商業專科學校學生輔導中心於1993年，以配額抽樣法對北部地區五所專科學校所做的寂寞感受抽樣調查。562位受訪者中，共回收497份完整的調查表，有效問卷回收率約88%。

3.3 實例分析

實例研究一

問題：在大台北地區，一個四口之家一個月的生活費用大概需要多少？(生活費用的定義：包括食、衣、住、行等)

表 3-1 為「台灣地區社會意向」預試調查搜集所得的資料。其中「上限」、「下限」是指將模糊區間資料的上、下界值各別加以彙總的次數統計資料；「確定值」資料，是指傳統要求受訪者回答確定數字，所彙總的次數統計資料。

表 3-1：人們對日常生活花費需求的看法

下限 (次數)	8 1	10 15	15 4	20 12	25 9	30 25	35 7	38 1	40 18	45 1	50 14	55 1	60 2	70 1	50 1
上限 (次數)	10 1	20 5	25 2	30 17	35 3	40 25	45 4	50 17	55 2	60 12	70 4	80 2	100 3	200 1	
確定值 (次數)	10 1	15 1	17 1	20 4	24 1	25 8	27 1	30 27	35 6	37 1	38 2	40 8	45 6	50 23	55 1
確定值 (次數)	60 2	70 1	75 1	90 1	100 1	150 1									

(單位千元)

其隸屬度函數為：

$$\begin{aligned}
 u(x) = & \frac{1}{98} I_{[8,10]}(x) + \frac{1}{98} I_{[10,15]}(x) + \frac{5}{98} I_{[15,20]}(x) + \frac{13}{98} I_{[20,25]}(x) + \frac{20}{98} I_{[25,30]}(x) \\
 & + \frac{28}{98} I_{[30,35]}(x) + \frac{32}{98} I_{[35,40]}(x) + \frac{25}{98} I_{[40,45]}(x) + \frac{23}{98} I_{[45,50]}(x) + \frac{20}{98} I_{[50,55]}(x) \\
 & + \frac{19}{98} I_{[55,60]}(x) + \frac{8}{98} I_{[60,65]}(x) + \frac{8}{98} I_{[65,70]}(x) + \frac{5}{98} I_{[70,75]}(x) + \frac{5}{98} I_{[75,80]}(x) \\
 & + \frac{3}{98} I_{[80,85]}(x) + \frac{3}{98} I_{[85,90]}(x) + \frac{3}{98} I_{[90,95]}(x) + \frac{3}{98} I_{[95,100]}(x) + \frac{1}{98} I_{[100,105]}(x)
 \end{aligned}$$

其中 $I_{[c,d]}(x)$ 為指標函數 (indicator function) ; 即 $I_{[c,d]}(x) =$

$$\begin{cases} 1 & \text{if } x \in [c, d]; \\ 0 & \text{if } x \notin [c, d]. \end{cases}$$

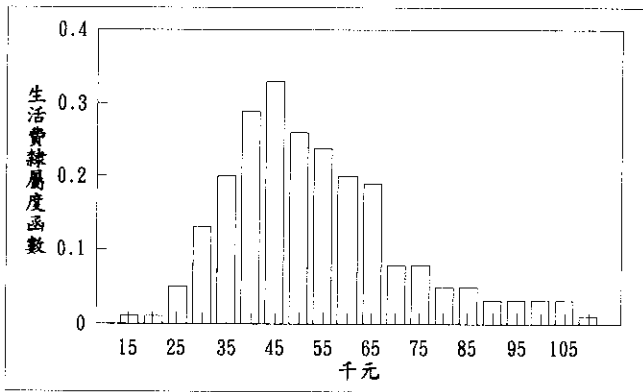


圖 3-1 生活費隸屬度函數

表 3-2 為根據模糊區間資料重新再統計彙總而得的相對頻數。

表 3-2：分組計算生活需求的相對頻數

需求分組(單位千元)	頻 數	相 對 頻 數
8--10	1	0.01
10--15	1	0.01
15--20	5	0.05
20--25	13	0.13
25--30	20	0.20
30--35	28	0.29
35--40	32	0.33
40--45	25	0.26
45--50	23	0.24
50--55	20	0.20
55--60	19	0.19
60--65	8	0.08
65--70	8	0.08
70--75	5	0.05
75--80	5	0.05
80--85	3	0.03
85--90	3	0.03
90--95	3	0.03
95--100	3	0.03
150--200	1	0.01

以「相對頻數」為各數值對於「生活費用」的隸屬度。

藉隸屬度函數，我們可以對本身是模糊的語言作定量的描述、劃分。35,000元對於「生活費用」模糊集而言，具有最高隸屬度，是最多民眾認可的生活所需成本。表 3-3將模糊調查和傳統調查的度量結果做一比較。

表 3-3：模糊統計與傳統統計分析結果之比較

模糊分析	$\mu_F = 40.9$	$F = (35.1, 46.7)$	$median_F = (30, 40)$	$mode_F = (30, 40)$
傳統分析	$\bar{x} = 39.9$	$S = 18.2$	$median = 35$	$mode = 30$

實例研究二

問題：在現代化的社會中，一個男生應該受多少年的教育才夠？

表 3-4 為「台灣地區社會意向」預試調查搜集所得的資料。其中「上限」、「下限」是指將模糊區間資料的上、下界值各別加以彙總的次數統計資料；「確定值」資料，是指傳統要求受訪者回答確定數字，所彙總的次數統計資料。為抽樣調查搜集的資料。

表 3-4：人們對男生教育年限的看法

下限 (次數)	9 5	12 36	14 18	15 3	16 34	18 1	21 1		
上限 (次數)	12 4	14 2	16 47	18 22	20 8	21 1	22 8	24 4	25 2
確定值 (次數)	12 18	14 8	15 4	16 56	18 11	21 1			

表 3-5 為根據模糊區間資料統計而得的相對頻數。

表 3-5：男生教育年限的相對頻數

教育年限	頻 數	相 對 頻 數
9	5	0.05
10	5	0.05
11	5	0.05
12	41	0.42
13	37	0.38
14	55	0.56
15	56	0.57
16	90	0.92
17	43	0.44
18	44	0.45
19	22	0.22
20	22	0.22
21	15	0.15
22	14	0.14
23	6	0.06
24	6	0.06
25	2	0.02

以「相對頻數」做為各數值對於「男生教育年限」的隸屬度。其隸屬度函數為

$$\begin{aligned}
 u(x) = & \frac{5}{98} I_9(x) + \frac{5}{98} I_{10}(x) + \frac{5}{98} I_{11}(x) + \frac{41}{98} I_{12}(x) + \frac{37}{98} I_{13}(x) + \frac{55}{98} I_{14}(x) + \frac{56}{98} I_{15}(x) \\
 & + \frac{90}{98} I_{16}(x) + \frac{43}{98} I_{17}(x) + \frac{44}{98} I_{18}(x) + \frac{22}{98} I_{19}(x) + \frac{15}{98} I_{20}(x) + \frac{15}{98} I_{21}(x) + \frac{14}{98} I_{22}(x) \\
 & + \frac{6}{98} I_{23}(x) + \frac{6}{98} I_{24}(x) + \frac{2}{98} I_{25}(x)
 \end{aligned}$$

其中 $I_c(x)$ 為指標函數；即 $I_c(x) = \begin{cases} 1 & \text{if } x = c; \\ 0 & \text{if } x \neq c. \end{cases}$

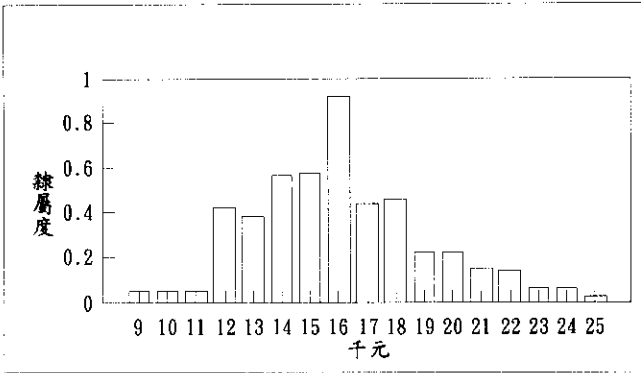


圖 3-2 男生教育年限隸屬度函數

16年對於「男生教育年限」模糊集而言，具有最高隸屬度，是最多民眾認為男生應受的教育年限。表 3-6將模糊調查和傳統調查的度量結果做一比較。

表 3-6：模糊統計與傳統統計分析結果之比較

模糊分析	$uF = 15.7$	$F = (13.8, 17.6)$	$median_F = (14, 16)$	$mode_F = (12, 16)$
傳統分析	$\bar{x} = 15.3$	$S = 1.9$	$median = 16$	$mode = 16$

實例研究三

問題：在現代化的社會中，一個女生應該受多少年的教育才夠？

表 3-7為「台灣地區社會意向」預試調查搜集所得的資料。其中「上限」、「下限」是指將模糊區間資料的上、下界值各別加以彙總的次數統計資料；「確定值」資料，是指傳統要求受訪者回答確定數字，所彙總的次數統計資料。為抽樣調查搜集的資料。

表 3-7：人們對女生教育年限的看法

下限 (次數)	6 1	9 5	12 37	14 19	15 3	16 31	18 1	21 1	
上限 (次數)	9 1	12 4	14 3	16 47	18 20	20 9	22 8	24 4	25 2
確定值 (次數)	9 1	12 19	14 10	15 5	16 56	18 6	21 1		

其隸屬度函數為

$$\begin{aligned}
 u(x) = & \frac{1}{98} I_6(x) + \frac{1}{98} I_7(x) + \frac{1}{98} I_8(x) + \frac{6}{98} I_9(x) + \frac{5}{98} I_{10}(x) + \frac{5}{98} I_{12}(x) + \frac{42}{98} I_{12}(x) \\
 & + \frac{38}{98} I_{13}(x) + \frac{57}{98} I_{14}(x) + \frac{57}{98} I_{15}(x) + \frac{88}{98} I_{16}(x) + \frac{41}{98} I_{17}(x) + \frac{42}{98} I_{18}(x) + \frac{22}{98} I_{19}(x) \\
 & + \frac{22}{98} I_{20}(x) + \frac{14}{98} I_{21}(x) + \frac{14}{98} I_{22}(x) + \frac{6}{98} I_{23}(x) + \frac{6}{98} I_{24}(x) + \frac{2}{98} I_{25}(x)
 \end{aligned}$$

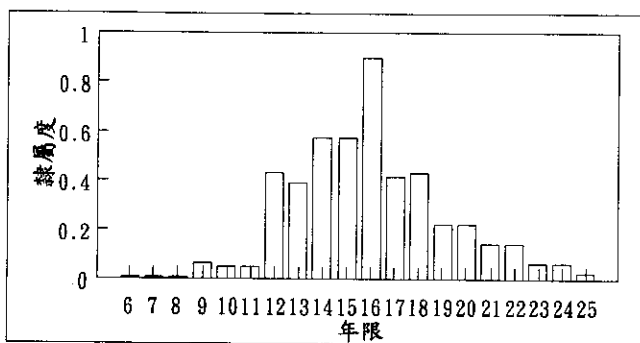


圖 3-3 女生教育年限隸屬度函數

表 3-8：女生教育年限的相對頻數

教育年限	頻 數	相 對 頻 數
6	1	0.01
7	1	0.01
8	1	0.01
9	6	0.6
10	5	0.5
11	5	0.5
12	42	0.43
13	38	0.39
14	57	0.58
15	57	0.58
16	88	0.90
17	41	0.42
18	42	0.43
19	22	0.22
20	22	0.22
21	14	0.14
22	14	0.14
23	6	0.06
24	6	0.06
25	2	0.02

表 3-8 為根據模糊區間資料統計而得的相對頻數。以「相對頻數」做為各數值對於「女生教育年限」的隸屬度。

16 年對於「女生教育年限」模糊集而言，具有最高隸屬度，是最多民眾認為女生應受的教育年限。表 3-9 將模糊調查和傳統調查的度量結果做一比較。

表 3-9：模糊統計與傳統統計分析結果之比較

模糊分析	$uF = 15.6$	$F = (13.7, 17.5)$	$median_F = (14, 16)$	$mode_F = (12, 16)$
傳統分析	$\bar{x} = 15.1$	$s = 1.9$	$median = 16$	$mode = 16$

實例研究四

由於快速的社會變遷，個人與個人間的關係逐漸由個人與團體間的關係所取代。因此，寂寞在現代社會中是一個相當普遍而惱人的問題。Russel et al. (1978) 爲了在寂寞的實證研究上有所突破，尋找建立一個可廣泛使用之單一而整體性的寂寞量表，經過一再地修訂，發展出 UCLA 寂寞量表。而後 Hays & DiMatteo (1987) 以因素分析等統計方法，由原 UCLA 量表中，抽出八題建立了 ULS-8 簡式寂寞量表。研究中針對寂寞感受程度的測量，此研究乃根據 UCLA 寂寞量表，配合臺灣的特殊背景，經過統計相關分析與模糊集合檢驗加以綜合、刪減後，發展而得精簡的8題寂寞量表。

本研究在探討如何應用模糊統計方法，對寂寞量表建立一套更有效率的測量準則。並希望藉這套測量準則，能事先預測出那些具有極端或雙重性格之學生，而施予事前輔導。首先我們針對北部地區五所專科學校作抽樣調查，表 3-10 爲傳統抽樣調查搜集的各題反應統計資料。

由表 3-10，我們發現對寂寞的綜合感受程度，呈平穩然後遞減的分配。但是此量表對於那些情緒起伏甚大者或過於內向者，可能無法得到一良好的度量。對寂寞所衍生出的各種副作用，如自殺，破壞性行爲亦無法提出較明確的預警作用。

表 3-10：傳統抽調寂寞反應頻率統計表

題目\感受程度	沒有感受 → 感受很深					合 計
	1	2	3	4	5	
1. 生活枯燥無趣味 (百分比)	96 (19.3)	132 (26.6)	152 (30.6)	76 (15.3)	41 (8.2)	497 (100)
2. 覺得生命失去目標	150 (30.2)	141 (28.4)	113 (22.7)	60 (12.1)	33 (6.6)	497 (100)
3. 無法肯定自我成就	67 (13.5)	118 (23.7)	142 (28.6)	102 (20.5)	68 (13.7)	497 (100)
4. 目前所學非所興趣	109 (21.9)	119 (23.9)	139 (28.0)	65 (13.1)	65 (13.1)	497 (100)
5. 在團體中被忽視	149 (30.0)	173 (34.8)	114 (22.9)	46 (9.3)	15 (3.0)	497 (100)
6. 沒有人真正瞭解我	92 (18.5)	127 (25.6)	153 (30.8)	76 (15.3)	49 (9.9)	497 (100)
7. 覺得缺少親近知己	143 (28.8)	132 (26.6)	109 (21.9)	65 (13.1)	48 (9.7)	497 (100)
8. 覺得需要異性朋友	132 (26.6)	107 (21.5)	158 (31.8)	55 (11.1)	45 (9.1)	497 (100)
合 計	938 (23.6)	1049 (26.4)	1080 (27.2)	545 (13.7)	364 (9.1)	3976 (100)

事實上寂寞是一個相當模糊的名詞，任何人都會有寂寞的感覺。而感覺強度的指標，傳統的測量工具卻很難訂定出來。主要困難有

- (1) 母體寂寞表的分配難以確立。
- (2) 寂寞強度的動態性、週期性無法測知。
- (3) 寂寞強度的兩極性或多極性無法區分。

基於以上原因，我們考慮用模糊問卷的方式，經由統計分析與模糊量度理論，建立一套模糊寂寞量度。希望能藉此一新的量表，更實際的測量出寂寞的感受與反應。表 3-11 為抽樣調查搜集的各題模糊反應統計資料。

表 3-11：模糊抽調資料寂寞反應頻率統計表

題目\感受程度	沒有感受 → 感受很深					合計
	2	3	4	5		
1. 生活枯燥無趣味 (百分比)	236.91 (47.7)	115.82 (23.3)	75.49 (15.2)	38.33 (7.7)	30.41 (6.1)	496.96 (100)
2. 覺得生命失去目標	260.98 (52.5)	117.72 (23.7)	65.72 (13.2)	31.40 (6.3)	21.15 (4.3)	496.96 (100)
3. 無法肯定自我成就	219.21 (44.1)	120.20 (24.2)	83.70 (16.8)	46.54 (9.4)	27.30 (5.5)	496.96 (100)
4. 目前所學非所興趣	54.36 (51.2)	115.52 (23.2)	68.35 (13.8)	29.20 (5.9)	29.53 (5.9)	496.96 (100)
5. 在團體中被忽視	261.76 (52.7)	120.68 (24.3)	76.34 (15.4)	24.68 (5.0)	13.52 (2.7)	496.96 (100)
6. 沒有人真正瞭解我	226.84 (45.6)	118.84 (23.9)	72.67 (14.6)	47.68 (9.6)	30.93 (6.2)	496.96 (100)
7. 覺得缺少親近知己	246.02 (49.5)	107.52 (21.6)	75.19 (15.1)	42.36 (8.5)	25.87 (5.2)	496.96 (100)
8. 覺得需要異性朋友	262.39 (52.8)	111.14 (22.4)	69.64 (14.0)	30.65 (6.2)	23.15 (4.7)	496.96 (100)
合計	1968.47 (50)	927.44 (23)	587.10 (15)	290.84 (7)	201.86 (5)	3975.71 (100)

根據表 3-11，得到二專生的模糊綜合寂寞隸屬度函數為

$$u(x) = 0.05 \times I_1(x) + 0.23 \times I_2(x) + 0.15 \times I_3(x) + 0.07 \times I_4(x) + 0.05 \times I_5(x)$$

我們訂定以下的寂寞量表標準：

定義 3.6. 單項模糊寂寞量表分配，SFDL (fuzzy distribution of loneliness for single item)，通常以抽樣結果得到。令 m 為感受程度之等級量表； P_{ij} 為個人在寂寞量表中，第 i 題感受程度第 j 等級的隸屬度。則單項模糊寂寞指數 I_{SFL} 定為

$$I_{SFL} = \frac{1}{m} \left(\sum_{j=1}^n jP_{ij} + \frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m jP_{ij} \right| \times P_{ij}}{k} \right) \tag{3.1}$$

其中 k 為感受程度等級中， P_{ij} 不為零的個數。

式 (3.1) 中，分母的第一項即為傳統的度量方法，第二則可視為模糊樣本的加權計分。單項模糊寂寞量表的分配函數，通常以抽樣結果得到。

以下我們以幾個特殊情況，說明定義 3.6 中對模糊感受的加權情形。以較常用之 5 等第量表為例 ($m = 5$)：

1. 當樣本選 5 時，此道題感受很強，個人寂寞量表感受百分比為 (0,0,0,0,1)，我們定此樣本之感受程度為清晰集合「5」， $I_{SFL} = 1$ 。
2. 當樣本選 3 時，此道題感受普通，個人寂寞量表感受百分比為 (0,0,1,0,0)，我們定此樣本之感受程度為清晰集合「3」， $I_{SFL} = 0.6$ 。
3. 當樣本為兩極化時，例如：個人寂寞量表感受百分比為 (0.5,0,0,0,0.5)，表示感受程度有時很強烈，有時沒有感受，此時樣本之模糊集合「 $0.5I_1 + 0.5I_5$ 」。感受程度為 $4 (= 3+1)$ ， $I_{SFL} = 0.8$ 。

4. 當樣本感受程度均等分佈時，例如：寂寞量表感受百分比為 (0.2,0.2,0.2,0.2,0.2)，此時樣本之模糊集合「 $0.2I_1+0.2I_2+0.2I_3+0.2I_4+0.2I_5$ 」。感受程度為 3.48 (3+0.45)， $I_{SFL} = 6.96$ 。

定義 3.7. 設 FDL 為模糊綜合寂寞量表分配。令 n 為模糊問卷題數，則綜合模糊寂寞指數 I_{SFL} 定為

$$I_{FL} = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{m} \left(\sum_{j=1}^m j p_{ij} + \frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m j p_{ij} \right| \times p_{ij}}{k} \right) \right] \quad (3.2)$$

式 (3.2) 亦可寫成

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{m} \sum_{j=1}^m j p_{ij} + \frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m j p_{ij} \right| \times p_{ij}}{m \times k} \right) \\ &= \frac{\sum_{j=1}^m j \left(\sum_{i=1}^n p_{ij} \right)}{n \times m} + \sum_{i=1}^n \left(\frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m j p_{ij} \right| \times p_{ij}}{n \times m \times k} \right) \\ &= \frac{\sum_{j=1}^m j s_j}{n \times m} + \sum_{i=1}^n \left(\frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m j p_{ij} \right| \times p_{ij}}{n \times m \times k} \right) \end{aligned}$$

其中最後一個等號第一項內的 s_j 表示為第 j 級感受程度的總合； $j=1, \dots, m$ 。

根據定義，我們得知一般二專生的單項模糊寂寞指數分別為：

表 3-12：二專生模糊抽調資料的單項模糊寂寞指數

題號	1	2	3	4	5	6	7	8
寂寞指數	0.5	0.47	0.51	0.48	0.46	0.51	0.49	0.47

綜合模糊寂寞指數為0.49

3.4 結果分析

通常，若有較多的樣本資料居於中間位置，在統計學上我們稱為集中趨勢 (central tendency)。一般以眾數或平均數用來代表此集中趨勢的量數。在實例一至三的研究，表 3-3、3-6、3-9 中，我們以統計上常用以代表群體的三種平均數——算數平均數、中位數、眾數，來比較傳統和模糊調查資料度量結果的異同。

在傳統的統計分析中，平均數都是用點來表示，所以又稱為點量數。但在模糊統計中，由於分析的是模糊區間資料，因此亦以區間形式來表示資料的集中趨勢，而其運算理念的構思，仍源於傳統統計的相關定義。另一方面，同一問題對於同一受訪者而言，其依傳統問卷形式所回答的確定數值答案，必包含於依模糊問題形式回答的模糊區間中，因此這兩種型式的資料，本身即具有某一程度的相關性，所以表 3-3、3-6、3-9，兩種資料分析的結果也有某種程度的相似性。

模糊邏輯考量了人類思維的複雜性及行為的不確定性，其允許模糊性現象的存在，較符合實際狀況，此一特點彌補了布林邏輯的不足，但並沒有取代布林邏輯。將模糊邏輯的觀念運用於問卷調查之分

析上，提供了一個新的搜集及分析資料的理念，和不同於往常的統計結果展示方法。

此外問卷調查時，對有多重感受的受訪者而言，傳統問法的單一邏輯，可能使其無所適從，且當尋問的問題涉及個人隱私時，受訪者很可能拒答。因此實例研究四中，允許人們擁有多重感受的模糊理論觀念，不但可減少廢卷率，還可幫助我們偵測出那些具有雙重性或危險傾向的學生，這是傳統的問卷型式所無法達成的目標。

肆、結 論

自從 Zadeh 在 1965 年提出了「模糊集合」的概念做為測試不明確隸屬度的工具以來，許多的模糊研究就嚐試以此方法設定理論架構，並嚐試以此一理論做為建構更為先進模糊理論的基礎。但模糊理論在社會科學中的應用迄今還是相當缺乏，而一些應用的範疇大部份均限定在計算機科學中的專家系統應用研究中 (Costas et al., 1994)。近年來社會科學中重要的研究，包括 Manton 和他的同仁所做有關模糊隸屬度和老人健康的一系列研究，這些研究的成果已受到相當廣泛的重視和應用 (Manton, Woodbury and Tolley, 1994)。雖然在心理計量學有一些模糊量表的設計，做為測量態度的依據 (Hesketh et al., 1988; Costas et al., 1994)，以及某些態度與偏好，如社會學中對於理想子女數以及男女教育年限的應用及測量 (Li, 1989)。但是大部份的研究對象均侷限於大學主修社會科學的學生的樣本中，這些研究幾乎不會將 Zadeh 所發展隸屬度的概念在一般人口的樣本中加以測試。本研究的重要貢獻在於利用「模糊集合」概念所發展出來的隸屬度概念應用於一般大型實地調查的樣本中，並且嚐試以模糊邏輯所發展出來的模糊測量量表，更進一步以模糊統計做為分析資料的依據。

在社會科學研究的測量理念上，模糊統計和模糊相關性的使用可說是數字模式的推廣。如果利用假性的「精確值」，可能誤導模型的

建構，也可能擴大研究結果和實際狀態之間的誤差。因此，具有模糊性的數值資料，在計量方法的演算過程中，是否符合人類的邏輯推論與歸納的原則，值得吾人深思。數值模糊的特性，在社會和經濟的實證研究過程中，應予以慎重考慮。

近年來，在人文社會科學的測度理念裡，模糊統計和模糊相關性日漸受到重視，這應是複雜的人文社會科學現象無法以傳統數值模型充分合理解釋的一種自然發展結果。Manski (1990) 曾指陳數字資料有需求過度及過度解釋的危險，如果應用模糊的數值資料，較能避免這樣的危險發生。有鑑於此，本文乃嚐試應用區間值的模糊特性所做的分析調查，可使研究者在處理不確定情境時能較符合實際狀況的測量工具。

本研究所獲得的結果，和 Manski (1990) 所發現的結果有些不同，即表現程度常發生在不是傾向於零，就是傾向於中立的狀態。因此我們建議對心理行為模式的建構與分析，可倚賴如語言法則的使用，而這些由許多語言造成的相關和傳統的積分、微分、代數和 ARMAX 方程間存在著直覺上的相似點，見 Sun & Wu (1993)。

最後，Kosko (1992) 建議神經網路和模糊語言系統共同應用做為函數估計量。希望此研究的方向將提供模糊統計一個有用的工具。然而，為了得到合適的信念準確度，期待神經計算學將成為值得花費時間的技巧，且將在社會科學研究方面，激勵更多未來模糊統計方法的實驗工作。

參考資料

楊文山與蔡瑤玲

- 1995 〈實地調查中複查資料的結構模型分析：以台灣地區社會意向調查為例〉，見章英華、傅仰止、瞿海源(編)，《社會調查與分析——社會科學研究方法檢討與前瞻之一》，頁83-110。臺北：中央研究院民族學研究所。

中央研究院中山人文社會科學研究所

- 1994.7 《台灣地區社會意向調查》，83年2月定期調查執行報告。

Casals, M. R., M. A. Gil, and P. Gil

- 1986 "On the Use of Zadeh's Probabilistic Definition for Testing Statistical Hypotheses from Fuzzy Information," *Fuzzy Sets and Systems*. 20: 175-190.

Casals, M. R. and M. A. Gil

- 1989 "A Note on the Operativeness of Neyman-Pearson Tests with Fuzzy Information," *Fuzzy Sets and Systems* 30: 215-210.

Costas, Concepcion San Luis and Pedro Prieto Maranon

- 1994 "Application of Diffuse Measurement to the Evaluation of Psychological Structures," *Quality and Quantity* 28: 305-313.

Hays, R. D. And M. R. DiMatteo

- 1987 "A Short-form Measurement of Loneliness," *Journal of Personality Assessment* 51 (1): 69-81.

Hesketh, B., R. Pryor, M. Gleitzman, and T. Hesketh

- 1988 "Practical Applications and Psychometric Evaluation of a Computerised Fuzzy Graphic Ration Scale," in Zeteny (ed.) *Fuzzy Sets in Psychology*. Elsevier Science Pub. North-Holland.

Klir, George, J., and Tina A. Folger

- 1988 *Fuzzy Sets, Uncertainty, and Information*. Englewood Cliffs, NJ: Prentice- Hall.

Kosko, B.

1992 *Neural Networks and Fuzzy System*. Englewood Cliffs, NJ: Prentice Hall.

Kruse R. and K. D. Meyer

1987 *Statistics with Vague Data*. D. Reidel Publishing Company.

Kwakernaak, H.

1978 "Fuzzy Random Variables. Part I: Definitions and Theorems," *Information Science*. 15: 1-15.

Li, Shaomin

1989 "Measuring the Fuzziness of Human Thoughts: An Application of Fuzzy Sets to Sociological Research," *Journal of Mathematical Sociology* 14: 67-84.

Manski, C.

1990 "The Use of Intention Data to Predict Behavior: A Best-case Analysis," *Journal of the American Statistical Associatio*. vol. 85, No 412: 934-940.

Manton, Kenneth G., Max A. Woodbury, H. Dennis Tolley

1994 *Statistical Applications Using Fuzzy Sets*. New York: John Wiley & Sons, Inc.

McNeill, Daniel and Paul Fredberger

1993 *Fuzzy Logic*. New York: Simon and Schuster.

Miyamoto, S.

1990 *Fuzzy Sets in Information Retrieval and Cluster Analysis*. Dordrecht: Kluwer Academic Publishers.

Russel, D., L. A. Peplau, and C. E. Cutrona

1978 "The Revised UCLA Loneliness Scale: Concurrent and Discriminant Validity Evidence," *Journal of Personality and Social Psychology* 39(3): 472-480.

Saade, J. J. and H. Schwarzlander

1990 "Fuzzy Hypotheses Testing with Hybrid data," *Fuzzy Sets and Systems* 35: 197-212.

Schuman, Howard, and Stanley Presser

1981 *Questions and Answers in Attitude Surveys*. New York: Academic Press.

Sudman, S. and N. M. Bradburn

1982 *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.

Sugeno, M.

1977 "On Structured Set of Systems. Report No. 5," *The Working Group on Fuzzy Systems*. Tokyo.

Sun, C. And B. Wu

1993 "On Fuzzy Analyzing and Modeling in Linguistic Semantics," *1993 Conference of Fuzzy Theory and Applications*. Hsinchu, Taiwan.

Watanabe N. and T. Imaizumi

1993 "A Fuzzy Statistical Test of Fuzzy Hypotheses," *Fuzzy Sets and Systems* 53: 167-178.

Zadeh, L.

1965 "Fuzzy Sets," *Information and Control* 8(3): 338-53.

Zimmmann, H. J.

1991 *Fuzzy Set Theory and Its Applications*. Boston: Kluwer Academic Publishers.

作者簡介

周雅容 國立中正大學心理系及研究所副教授。美國紐約大學心理學博士。研究及教學著重以社區生態及本土化取向探討社會及心理層面議題。領域包括老人及家庭研究，心理健康，兩性研究，健康心理學，評鑑研究，心理計量及質性研究。論文著作有〈不同來源的社會支持與老年人的心理健康〉、〈象徵互動論與語言的社會意涵〉、〈焦點團體法在調查研究的應用〉等篇。目前以兩性研究及心理健康研究為重點。

薛承泰 臺大社會學系副教授，1978年臺大社會學系畢業，1986年美國威斯康辛大學(麥迪遜)社會學碩士，1992年社會學博士。主修社會階層與研究方法，副修統計。近兩年之論文發表有〈臺灣的地位取得研究：回顧與前瞻〉、〈臺灣單親戶的分佈與特性〉、〈臺灣地區職業地位測量與性別差異〉、及〈影響國初中後教育分流的實證分析〉等篇。目前研究興趣在於婦女再就業之職業流動以及教育制度變遷之人口因素等議題。

蔡瑞明 東海大學社會學系副教授。國立臺灣大學社會學系及研究所畢業，美國康乃爾大學社會學博士。主要的研究興趣在社會階層、組織社會學、社會研究方法與統計。目前從事有關階級分類、組織內社會流動的研究。著有〈臺灣階級結構與社會流動〉、〈夫妻教育階級配對〉、〈美國工會的沒落〉等論文。

謝雨生 國立臺灣大學農業推廣學系及研究所教授。國立臺灣大學農業推廣學系及研究所畢業，美國賓州州立大學鄉村社會學博士。研究範圍在社會階級與流動、人力資源運用、人口學和社會研究法及統計。發表論文有〈農村人力資源、就業結構及改換工作意願之變遷研究〉、〈臺灣勞動力人口結構變遷對人力運用的影響〉，及

〈臺灣的社會階級及其流動〉等篇。目前研究偏重與臺灣的社會階級有關的議題。

鄭宜仲 國立臺灣大學農業推廣學研究所博士候選人。國立臺灣大學農業推廣學系及研究所畢業。專長學科包括鄉村社會學、農業推廣學、社會研究法及社會統計。發表論文有〈多元迴歸分析的假定與實例檢討〉、〈農業生產勞力婦化之再探討〉及〈農民友誼網絡資源之初探〉等篇。目前的研究興趣為農業婦女勞動力及農民友誼網絡之相關議題。

翁儷禎 國立臺灣大學心理學系副教授。美國加州大學洛杉磯分校心理學博士，主修心理計量與測量。研究範圍包括相依受試者之共變數結構分析、智力評估工具之心理計量特徵、推理思考能力之測量和評定量尺的心理意義等。

P. M. Bentler 美國加州大學洛杉磯分校教授，研究範圍包括心理計量與心理計量軟體之發展與運用，研究論文著作，見於主要心理學期刊與心理學計量期刊中。

林祖嘉 國立政治大學經濟系教授。大同工學院事業經營系學士，國立政治大學經濟研究所碩士，美國洛杉磯加州大學 (UCLA) 經濟學博士。研究範圍在住宅經濟、大陸經濟、及勞動經濟等。發表論文有〈價格分散與搜尋均衡在臺灣地區住宅市場之驗證〉、〈失業搜尋、在職搜尋、與工作轉換：巢式 Logit 模型的應用〉、及〈Production Function, Factor Substitutibility, and Direct Foreign Investment: A Case of Taiwan〉等篇。目前研究偏重於與臺灣地區有關的應用個體議題。

李紀珠 國立政治大學經濟系教授；國立臺灣大學直升經濟學博士，哈佛大學訪問學者。研究興趣主要在金融、貨幣理論與政策、總體經濟學、應用計量經濟學、大陸經濟與兩岸經貿。發表論文有

〈金融機構失敗預測模型——加速失敗時間模型之應用〉、〈美國景氣循環期間穩定性、相依性及政策影響力——另一種實證分析方法〉、〈釘住通貨匯率制度設計〉、〈兩岸三地金融合作與競爭〉、〈臺灣經濟發展中金融政策之角色〉、〈中央銀行獨立性之測量及對政策效果之影響〉等篇。

陳樹衡 國立政治大學經濟系副教授。國立政治大學經濟系畢業，國立臺灣大學經濟研究所碩士，美國加州大學洛杉磯分校經濟學博士。研究範圍在 genetic programming、stochastic complexity、Neural Network 等之應用。發表論文有〈From the Hayek Hypothesis to Animal Spirits: The Phase Transition based on Competitive Experimental Markets〉、〈Elements of Information Theory: A Book Review〉、〈Genetic Programming Learning and the Cobweb Model〉、〈Genetic Programming in the Coordination Game with a Chaotic Best-Response Function〉、〈Toward a Computable Approach to the Efficient Market Hypothesis: An Application of Genetic Programming〉、〈Measuring Randomness by Rissanen's Stochastic Complexity: Applications to the Financial Data〉等。

葉佳炫 國立政治大學經濟系博士班研究生。國立清華大學數學系畢業，國立政治大學經濟系碩士。研究範圍在以 genetic programming 架構與分析多決策者系的經濟環境。發表論文有〈Genetic Programming Learning and the Cobweb Model〉、〈Genetic Programming in the Coordination Game with a Chaotic Best-Response Function〉、〈Toward a Computable Approach to the Efficient Market Hypothesis: An Application of Genetic Programming〉等。

施俊吉 中央研究院社科所研究員，中央大學產經所合聘教授。臺灣大學經濟學博士。研究領域為：產業經濟學、管制經濟學與賽局理論。曾於國際及本土經濟學期刊發表四十餘篇學術論文。目前之研究重點為自由化與民營化等管制革新之課題。

吳柏林 國立政治大學應用數學系教授。1988年美國印地安那大學數學博士。研究領域包括時間數列分析，統計在社會科學中之應用及模糊理論與神經網路發展。近兩年之論文發表有〈時空數列在航空機位保留比率預測與應用〉、〈模糊時間數列與臺灣地區中學教師需求人數之預測〉、〈臺灣短期貨幣需求函數穩定性的檢定：模糊時間數列方法之應用〉、〈臺灣地區失業率的時空數列分析〉及〈Identification environment and robust forecasting for nonlinear time series; *Comp. Econ.*〉、〈Pattern recognition and classification in time series analysis; *Applied Math. and Comp.*〉、〈Model-free forecasting for nonlinear time series: with application in exchange rates; *Comp. Stat. & Data Anal.*〉等篇。目前研究興趣著重在時間數列模式之整合分析與預測之探討。

楊文山 中央研究院中山人文社會科學研究所研究員。輔仁大學社會系畢業，美國維州理工大學社會學碩士，德州奧斯汀大學社會學博士，匹茲堡卡內基米倫大學統計系博士後研究，並曾任職於美國福利衛生部亞特蘭大之國家疾病控制中心。主要研究興趣在於人口學、醫療社會學、統計與調查方法。目前則致力於臺灣地區死亡率長期變遷之趨勢、老人健康與照護、殘障之隨機過程模型、臺灣地區嬰兒猝死症之社會因素分析與日據時代以來人口普查與人口變遷之研究。論文著作見於國內外學術期刊，最近發表之論文有〈Gender Differences in Infant Mortality in Taiwan〉, *Soc. Sci. & Med.* 〈Risk Factors of Sudden Infant Death in Chinese〉 *American Journal of Epidemiology* 以及〈臺灣地區老人自殺長期趨勢之研究，1981~1991：人口學分析〉等多篇論文。合編有人口變遷、國民健康與社會安全。

索 引

中英對照

- 「二元波比模型」(Bivariate Probit Model) 40
- 「人類資源研究中心」(Center for Human Resources Research) 42
- 「干擾項」(disturbance) 36
- 「內衍變項」(endogenous variable) 36
- 「手足研究」(Sibling studies) 33
- 「手足配對」(sibling pairs) 33
- 「兄弟姊妹教育取得的相似性」(sibling resemblance in educational) 27
- 「外衍變項」(exogenous variables) 36
- 「因素分析」(factor analysis) 36
- 「地位取得研究」(status attainment research) 31
- 「多重群體」(multi-groups) 35
- 「完全適合」(perfect fitting) 41
- 「完全辨識」(just-identified) 38
- 「角色模範」(role-modeling) 35
- 「波比模型」(Probit Model) 40
- 「門檻」(thresholds) 39
- 「指標」(indicators) 36
- 「限控」(censoring) 39
- 「迷迷模型」(MIMIC Model) 27
- 「假設模型」(the hypothesized model) 56
- 「巢狀」(nested) 41
- 「測量誤差」(measurement errors) 37
- 「等級迷迷模型」(Ordered MIMIC Model) 27
- 「結構共變模型」(Structrual Covariance Model) 39
- 「順序層級」(Ordinal Level) 39
- 「概似比測試」(Likelihood Ratio Test) 41

- 「過度辨識」(over-identified) 38
 「飽和模型」(saturated model) 56
 「對數迴歸模型」 31
 「構念」(construct) 55
 「標準化」(standardization) 38
 「潛在變項」(latent variable) 36
 「餘值共變量」(residual covariance) 34
 「辨識」(identified) 38
 「選擇性偏差」(selection bias) 46
 「遺失變項」(omitted variables) 54
 「聯合機率函數」(joint probability function) 40

英文部份

A

- Amemiya 140
 Amemiya's 146

B

- Benin 27
 Bentler 146, 160

- Bentler Psychological Inventory (BPI) 160
 bic (Bayesian Information Criterion) 41
 Blau and Duncan 31
 Bollen 7
 Byrne 14

C

- Comparative Fit Index (CFI) 8
 confirmatory factor-analytic (CFA) approach 6
 covariance structure analysis 139

D

- degree of consensus 263
 dependent observations 140
 Donner 143

F

- Freedman 139

G

- generalized gamma distribution, GGD 189
 Goldberger 27

H

- Hauser 27, 28
 Holland (1975) 214

J

Jöreskog 56

Johnson 27

K

Koval 143

Koza (1989) 214

L

L. A. Zadeh 292

Lagrange Multivariate (LM) test
14

Likert 量表 263

LISCOMP (Analysis of Linear
Structural Relations Using a
COMprehensive Measurement
Model) 39LISP (List Programming) 語言
214

LISREL 27

M

MacCallum 8

matrix normal distribution 140

maximum likelihood estimators
140multiple-indicators and multiple-
causes model 27

Muthén 39

N

Newcomb 160

NLSY (National Longitudinal
Survey of the Youth) 28

R

Radloff, 1977 2

Raftery 41

root-mean-squared-error (RMSE)
151

S

Sörbom 56

Stadje 163

structural equation model 7

T

Tanaka 8

the Bentler-Bonett's Normed Fit
Index (NFI) 8the Bentler-Weeks representation
system (i.e. EQS notation,
Bentler 1992) 7the Center for Epidemiological
Study Depression (CES-D)
scale 2

U

UCLA 寂寞量表 307

ULS-8 簡式寂寞量表 307

W

Weibull 分配 189

Weng 146

Wong 28

中文部份

四劃

- 不平均度 (inequality) 267
公設式分析法 (axiomatic approach) 265
手足配對 (sibling pairs) 28
手足教育相似性 27

五劃

- 加速生命模型 (accelerated life model) 190
右方截斷 (right censored) 191
布林模式 291

六劃

- 交互作用效果 115
交換性流動 70
危險函數 (hazard function) 204
存活模型 185

八劃

- 社會共識程度 263

九劃

- 信度 290

相對流動 64

十劃

- 乘法模式 (multiplicative model) 113
效用函數 291
效度 290
時間相依性 187
真正差異 (true difference) 105
迴歸標準化 (regression standardization) 105
迴歸調整法 (regression adjustment method) 105

十一劃

- 淨組成效果 (partial composition effect) 113
淨團體效果 (partial group effect) 113
淨關係 (partial association) 113
率效果 (rate effect) 115
粗差異 (crude difference) 105
粗率 (crude rate) 104
粗率差異 115
粗率差異分解 115
組成分配 (distribution of composition) 104
組成特殊率 (composition-specific rate) 104
設計矩陣 74

十二劃

- 結構性流動 70
- 結構差異效果 106
- 絕對流動 64
- 集中趨勢 312

十三劃

- 經驗指標 289
- 路徑分析 (path analysis) 31

十四劃

- 團體結構性差異 105
- 團體結構性差異效果 104, 106
- 對數相乘模型 61
- 對數線性分析 61
- 對數線性消除法 (log-linear purging method) 105
- 對數線性模型 61
- 演化計算 (Evolutionary Computation) 213

十五劃

- 層級模型 61
- 標準團體 108
- 模糊子集理論 291
- 模糊統計 289
- 模糊統計檢定 294
- 模糊語言系統 294

- 模糊隨機變數 294
- 調整率 (adjusted rate) 110

十六劃

- 遺傳規畫 (Genetic Programming) 212
- 遺傳程式 (Genetic Algorithms) 214

十七劃

- 隸屬度函數 291

十九劃

- 關連模型 63
- 類別資料 62

二十三劃

- 邏輯模型 (logic model) 31

附 錄

吳齊殷：(中央研究院社會學籌備處)

潛藏成長曲線分析的邏輯與應用：重覆測量資料之動態結構方程模型

Yah-Jong Chou (周雅容)：(中正大學心理系和研究所)

The Structure of Depression: An Application of Confirmatory Factor Analysis to Higher-Order Factor Structure and Factorial Invariance

薛承泰：(國立臺灣大學社會學系)

「迷迷模型」(MIMIC Models) 的應用與延展：以「手足教育相似性」為例

蔡瑞明：(東海大學社會學系)

「對數線性模型」與「對數相乘模型」的發展與運用：以社會流動分析為例

謝雨生：(國立臺灣大學農業推廣學系及研究所)、鄭宜仲：(國立臺灣大學農業推廣學研究所博士候選人)

團體比較研究中結構差異的考慮與處理：論迴歸標準化及對數線性消除法的應用

李隆安：(中央研究院統計科學研究所)、陳君厚：(中央研究院統計科學研究所)、李克昭：(美國加州大學洛杉磯數學系)

反切迴歸在社會科學領域的應用

Li-Jen Weng (翁儷禎)：(國立臺灣大學心理學系)、P. M. Bentler：(美國加州大學洛杉磯分校)

Covariance Structure Analysis with Intraclass Dependent Observations

鄭讚源：(中正大學社會福利研究所)

多層級分析策略——從變異數分析 (ANOVA) 到多層級線性模型 (HLM)

陳信木：(國立政治大學社會系)、林佳瑩：(國立政治大學社會系)

勞工離職行為之動態過程——事件史分析模型之應用

田芳華：(中央研究院社會學籌備處和調查研究工作室)

我國大學教師升等制度對獎勵研究生產力功效之分析——事件歷史分析法應用實例

林祖嘉：(國立政治大學經濟系)、李紀珠：(國立政治大學經濟系)

大陸臺商工廠待獲利期間之分析：存活模型的應用

林建甫：(國立臺灣大學經濟系)、盧陽正：(銘傳管理學院金融所)

LSE 經濟計量學的分析方法：臺灣貨幣需求的實證研究

陳樹衡：(國立政治大學經濟系)、葉佳炫：(國立政治大學經濟系博士班研究生)

貨幣數量學說的競爭性：遺傳規畫在知識發掘上的應用

丘昌泰：(國立中興大學公共行政系)

當代政策研究中的匯合分析法

施俊吉：(中央研究院中山人文社會科學研究所)

社會共識程度之衡量

吳柏林：(國立政治大學應用數學系)、楊文山：(中央研究院中山人文
社會科學研究所)

模糊邏輯在社會調查上的應用

國家圖書館出版品預行編目資料

社會科學計量方法發展與應用 / 楊文山主編.

-- 臺北市：中研院社科所，民86

面；公分.-- (中央研究院中山人文社會科學
研究所專書：41)

含參考書目及索引

ISBN 957-671-541-5 (平裝)

ISBN 957-671-542-3 (精裝)

1. 社會科學 - 統計方法 - 論文, 講詞等

501.2807

86011820

中央研究院中山人文社會科學研究所專書(41)

社會科學計量方法發展與應用

主 編 楊 文 山
出 版 者 中央研究院中山人文社會科學研究所
發 行 者 中央研究院中山人文社會科學研究所
編 輯 者 林容如
校 對 者 彭玉屏
印 刷 者 天翼電腦排版印刷股份有限公司

中華民國八十六年九月出版

版權所有 不准翻印

SUN YAT-SEN INSTITUTE
FOR SOCIAL SCIENCES AND PHILOSOPHY
ACADEMIA SINICA

BOOK SERIES

(41)

RECENT DEVELOPMENT AND
APPLICATION OF QUANTITATIVE
METHODS IN SOCIAL SCIENCES

VOL. VI

Edited by

WEN-SHAN YANG

NANKANG, TAIPEI, TAIWAN
REPUBLIC OF CHINA
SEPTEMBER 1997