

# 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。

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

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。

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

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。<sup>1</sup> 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  $\tau_{ijk}^{CGD}$  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  $\tau_{ijk}^{CGD}$  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， $\tau^{CG}$  測量了組成與團體的淨關係， $\tau^{GD}$  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  $\tau^{CD}$  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  $\tau^{CG}$  和  $\tau^{CGD}$  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  $\tau_{ij}^{CG}$  和  $\tau_{ijk}^{CGD}$  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，<sup>2</sup>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<sup>+</sup>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由此實例，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 <sup>***</sup>
Ln(月工作所得)	10.2498 (.447)	9.7125 (.489)	.537 <sup>***</sup>
年齡	35.7052 (10.159)	32.9714 (9.590)	2.734 <sup>***</sup>
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 <sup>***</sup>
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 <sup>***</sup>
職業聲望	54.6191 (8.905)	53.2817 (7.300)	1.337 <sup>**</sup>

\*\* P &lt; .01, \*\*\* P &lt; .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 <sup>2</sup>	.2562		.2697	
N	563		489	

\* P &lt; .05, \*\* P &lt; .01, \*\*\* P &lt; .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 <sup>+</sup>	1.24	2.00	0.76	0.27	0.84	0.57
△ <sup>*</sup>			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 <sup>+</sup>	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 <sup>+</sup>	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 <sup>a</sup>	調整數	指數 <sup>b</sup>	結構差異效果	
					(5)=(1)-(3)	(6) <sup>c</sup> =(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.

Iams, 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：

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

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

$$+ \frac{\sum b_{if}(\bar{X}_{im} - \bar{X}_{if})}{(3)}$$

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

$$= \frac{(a_m - a_f) + \sum \bar{X}_{if}(b_{im} - b_{if})}{(1) \quad (2)}$$

$$+ \frac{\sum b_{im}(\bar{X}_{im} - \bar{X}_{if})}{(3) + (4)}$$