



# 建立工商業母體資料檔之研究

為銜接非工商普查年時間數列資料，即時陳示工商企業之動態趨勢，並提升企業面抽樣調查母體之時效及其統計品質，爰積極研究運用資訊及統計技術，協調各公務機關秉持互信互惠之原則，結合公務與統計調查檔案，建立定期更新及回饋機制，以期提升現行普查母體檔運用效益。

◎呂淑君、郭燕玲、吳聲和、林玉樹、謝致和(行政院主計處第四局編審、專員、科員、科員、專員)

## 壹、前言

本研究以90年工商及服務業普查母體資料檔為基礎，分為工商企業基礎資料及營運資料二部分，廣泛蒐集相關公務檔案及統計調查檔案，運用資料比對、檔案連結、資料插補及設算與資料挖掘之分類與迴歸樹分析法(CART分析法)以及指數平滑法等方法，更新廠商資料，以建置最新之工商業母體資料檔。

## 貳、名冊整編作業

運用相關公務檔案，以及大型抽樣調查檔案資料進行標準化及邏輯化比對，研判各檔案資料之正確性，並剔除倒閉、停歇業廠商，新增新開業廠商，進行各種異動狀況更新，以產生工商企業基礎資料。

### 一、蒐集相關檔案

(一) 公務檔案：中文稅籍主

檔、營業人銷售額年檔、公司登記基本資料檔、勞工保險投保事業單位資料檔、醫院診所機構基本資料檔。

(二) 調查檔案：工廠校正暨營運調查檔、受雇員工薪資調查檔。

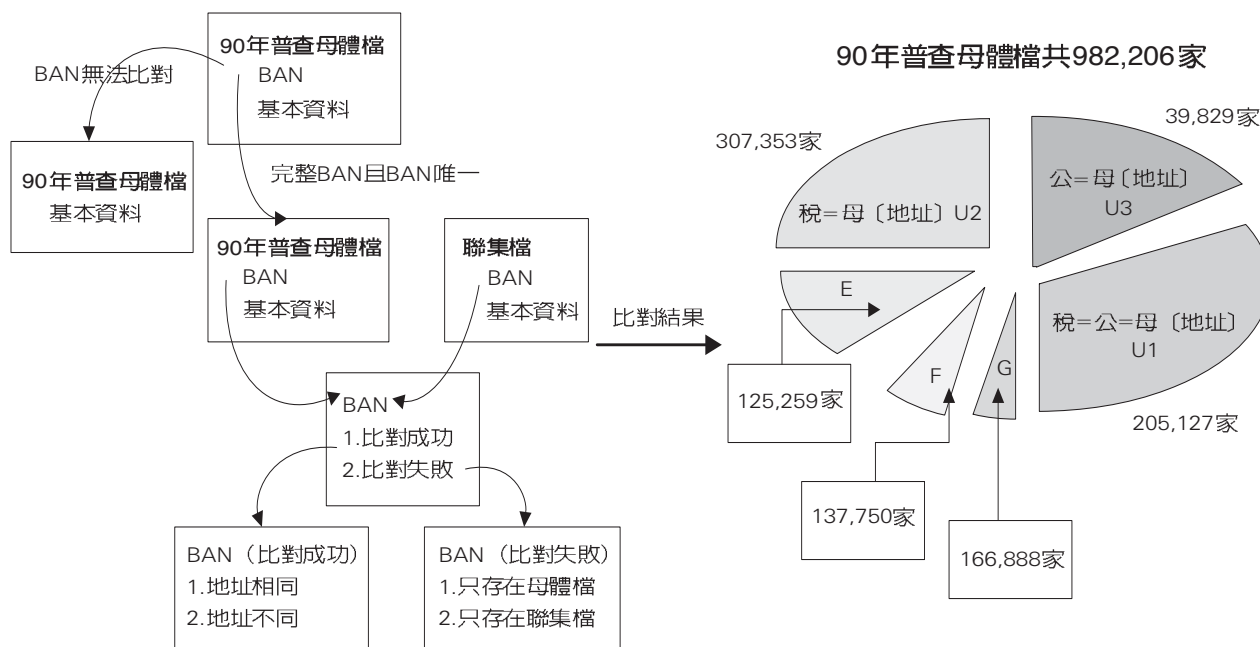
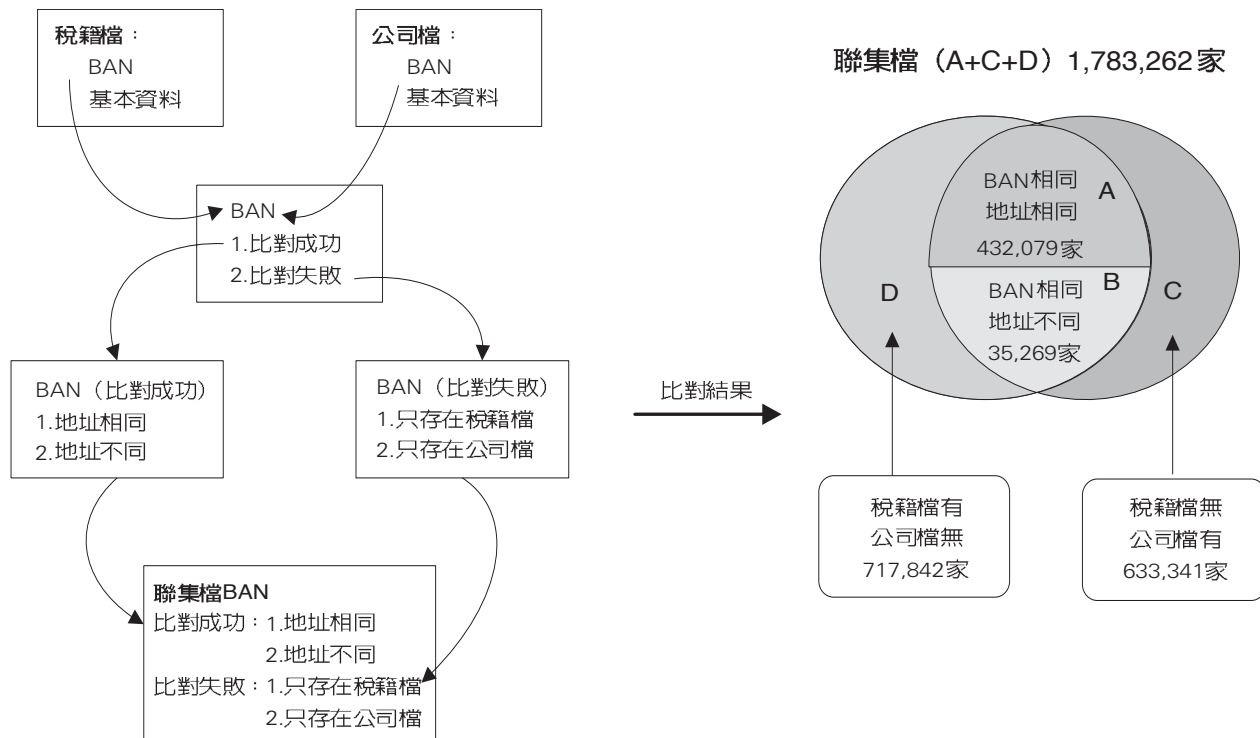
### 二、資料比對更新

(一) 母體檔地址比對結果

1. 最新公務檔案

(1) 處理流程：

# 論 述



(2) 比對結果：將BAN相同且地址相同及BAN不同者組成一聯集檔 (A + C + D)，將此聯集檔視為最新公務檔案之比對依據。

## 2.90年普查檔

(1) 處理流程：參見第54頁下圖。

(2) 比對結果：90年普查檔以BAN完整且唯一之部分 (U1 + U2 + U3 + E + F) 與前述聯集檔進行比對，其中地址相同者占56.2% (U1 + U2 + U3)；BAN相同但地址不同者占14.0% (F)；BAN比對失敗者占12.8%

(E)。

(3) 研判準則：

將90年普查檔樣本中BAN完整且唯一之部分 (U1、U2、U3、E、F) 與聯集檔依BAN比對後，分析比對結果，就各種比對狀況研訂40餘項判定準則，以判定廠商之正確地址，茲概述如下：

| 比對情形 / 判定準則   |
|---|
| <p>1. 聯集檔與普查檔比對成功且地址相同 (U1、U2、U3)</p> <p>準則：普查檔為90年廠商之實際地址，與93年稅籍檔及94年公司檔之地址相同，表示該廠商90年至今地址無異動，可認定普查檔即為目前正確地址。</p>  |
| <p>2. 聯集檔與普查檔比對成功但地址不同 (F)</p> <p>準則：A. 停歇業、他遷不明者，註記母體檔：分析稅籍檔91年至93年異動情形，有異動部分包含停歇業、他遷不明者，於母體檔註記。</p> <p>B. 與稅籍檔比對：若90年與93年稅籍檔之地址相同，卻與90年普查員實地踏查之地址不同，表示稅籍檔應為廠商之登記地址，可認定90年普查檔為廠商之實際營運地址。反之，表示廠商地址已經有所變更，故認定稅籍檔為廠商之實際營運地址。</p> <p>C. 與勞保檔比對：若90年與94年勞保檔之地址相同，卻與90年普查員實地踏查之地址不同，表示勞保檔應為廠商之登記地址，可認定90年普查檔為廠商之實際營運地址。反之，表示廠商地址已經有所變更，故認定勞保檔為廠商之實際營運地址。</p> <p>D. 無法比對廠商，若90年營業收入1千萬元以下者，以聯集檔為準，其餘以人工之方式列印檢查。</p> |
| <p>3. 聯集檔與普查檔BAN比對失敗者 (E)</p> <p>準則：加入勞保檔及稅籍檔與公司檔BAN比對成功但地址不同部分進行比對。</p> <p>A. 若普查檔地址與稅籍檔、公司檔或勞保檔任一檔案相同者，以普查檔為準，係因90年普查員實地踏查地址與最新之公務檔案相同，表示普查員可依此地址辦理調查工作，</p>  |

比對情形 / 判定準則

認定此地址為廠商實際營運地址。

B.若勞保檔與稅籍檔、公司檔任一檔案地址相同，但與普查檔不同者，以勞保檔為準，係因90年普查檔地址與2個最新之公務檔案不同，而這2個最新之公務檔案地址相同，表示廠商地址已經有所變更，認定94年勞保檔為廠商之實際營運地址。

C.BAN不同或地址皆不相同者，依BAN再與90年稅籍檔比對，比對成功且93年稅籍檔已不存在，則母體檔註記停歇業；比對失敗，再與醫療院所檔及執行業務檔進行比對。

以上第2種情形分析：137,750家廠商中可認定實際營運地址占94.8%；須以人工列印檢查占5.2%。第3種情形分析：125,259家廠商中可認定實際營運地址占18.5%；倒閉或停歇業占65.8%；與執行業務檔繼續比對占15.7%。

(二) 新增廠商處理

聯集檔存在而90年普查檔不存在者，計353,742家。因稅籍檔為實際納稅廠商資料，而公司檔為公司登記資料，可能發生只登記未實際營業狀況，故僅以稅籍檔部分比對92年銷售額年檔，經剔除非普查

對象、倒閉停歇業廠商後，新增廠商計289,496家。

(三) 未比對到之廠商地址處理

90年普查檔存在而聯集檔不存在者，計166,888家，按下表方式處理。

(四) 調查檔案比對

為有效掌握主管單位常川

| 未比對到之廠商情形    | 處理方式  |
|--------------|---|
| 1.BAN空白      | 包括小規模廠商、醫院、診所、計程車，將與執行業務檔及醫療院所檔再進行比對。                               |
| 2.企業之總管理單位   | 以企業之地址更新即可。   |
| 3.BAN重覆      | 大多為分公司，或公營事業之分支單位，將與其他相關公務檔案比對後予以更新。                                |
| 4.計程車重覆BAN   | 以相同BAN之計程車車行地址取代。   |
| 5.郵政儲金匯兌業    | 因中華郵政公司包含郵政儲金匯兌業與郵政業，同一家郵局之郵政業地址即為郵政儲金匯兌業之地址，故以郵政業之地址取代即可。          |
| 6.BAN完整但地址缺漏 | 將與新增廠商比對，比對成功即認定新增廠商為正確營業地址，比對失敗則再依序與公司檔、勞保檔比對，若仍比對失敗，則認定此為倒閉停歇業廠商。 |

蒐集更新之事業單位資料，以BAN再與經由上述比對過程所產生之工商業母體資料檔進行比對，醫院診所部分以醫療院所檔為準；製造業部分以工廠校正檔為準；受雇員工薪資調查之受查廠商以受雇員工薪資調查檔為準，完成工商業母體資料檔名冊整編工作。

## 參、資料連結作業

續就名冊整編完成之母體檔，蒐集運用相關公務及統計調查檔案進行連結，無法連結者則以統計方法進行資料預測，據以更新廠商之營運資料，產生最新之工商業母體資料檔。

### 一、運用統計調查連結測試

#### (一) 檔案來源

1. 公開資料：公開資訊公司之財務資料、公營事業決算書、行政院金融監督管理委

員會銀行、壽險公司資料。

2. 統計調查：工廠校正暨營運調查、工商企業經營概況調查、營造業經濟概況調查、運輸倉儲及通信業產值調查、會計師事務所服務業調查。
3. 勞工保險投保事業單位資料檔。

#### (二) 結果說明

以90年普查檔與91年統計調查、92年3月勞保檔進行連結測試，可知90年工商企業營業收入可與統計調查檔案連結成功者占7.8%，無法連結者占92.2%。其中可連結之7萬餘家廠商，其營業收入占全體62.6%，無法連結者則占37.4%。若按營業規模觀察，71.5%普查營業收入10億元以上之廠商，可連結成功，其所占營

業收入達86.3%，而營業收入未達1,000萬元之廠商，連結效果最低。

90年工商企業員工人數與勞保檔連結成功者占31.4%，無法連結勞保檔者占68.6%。而可連結到勞保檔之普查廠商員工人數占全體73.9%，無法連結者占26.1%。若按員工規模觀察，普查員工人數未滿5人者連結比率最低，僅19.3%，主要係因勞保檔5人以下得自由投保，致該規模層連結比率較差，若扣除5人以下規模別廠商，則71.9%之家數與86.3%之員工人數均可連結勞保檔資料，其中百人以上企業之連結成功家數及人數均逾93%，連結效果最佳。

| 項 目        | 90年普查     | 可連結 (%) | 無法連結 (%) |
|------------|-----------|---------|----------|
| 員工人數       |           |         |          |
| 企業家數 (家)   | 935,316   | 31.4    | 68.6     |
| 員工人數 (人)   | 6,663,350 | 73.9    | 26.1     |
| 營業收入       |           |         |          |
| 企業家數 (家)   | 935,316   | 7.8     | 92.2     |
| 營業收入 (十億元) | 26,764.7  | 62.6    | 37.4     |

## 二、建立工商業母體資料檔之處理方式

觀察90年普查結果，其中製造業生產總額最高，批發及零售業總家數居冠，且該二業別通常亦為相關產經研究之重要範疇，爰以該二業別為例，按以下方式建置資料檔。

- (一) 運用CART分析法，將90年普查同質性之廠商，歸入相同之群組。
- (二) 同組廠商中，可連結92年統計調查資料者代入連結資料，並計算平均值。
- (三) 按指數平滑法預測續存廠商92年資料平均值結構變動之趨勢。
- (四) 新增廠商中，可連結統計調查資料者代入連結



資料，無法連結者，則按同業同層廠商90年新增廠商與續存廠商之相關性，並參考92年續存廠商之同業同層資料，預測該業該層新增廠商資料。

(五) 資料結果及評估

### 1. 員工人數

觀察92年工商業母體資料檔企業單位員工人數預測值之整體變動趨勢，與受雇員工薪資調查結果相較，發現製造業兩年來分別減少1.2%與0.1%，兩者之減幅相近，而批發及零售業則分別增加5.1%與

| 行業別    | 工商業母體資料檔      |              |             | 受雇員工薪資調查   |            |             |
|--------|---------------|--------------|-------------|------------|------------|-------------|
|        | 92年預測值<br>(人) | 90年普查<br>(人) | 增減比較<br>(%) | 92年<br>(人) | 90年<br>(人) | 增減比較<br>(%) |
| 製造業    | 2,389,505     | 2,419,712    | -1.2        | 2,345,600  | 2,348,301  | -0.1        |
| 批發及零售業 | 1,895,428     | 1,803,945    | 5.1         | 1,378,750  | 1,362,523  | 1.2         |

1.2%，工商業母體資料檔雖較受雇員工薪資調查結果高出3.9個百分點，惟兩者結果均顯示，批發及零售業之員工人數，應呈增加趨勢，且增加幅度不大。

## 2.營業收入

觀察92年工商業母體資料檔企業單位營業收入預測值之整體變動趨勢，製造業與營利事業銷售額資料結果相較，發現製造業兩年來分別增加22.9%與20.9%，兩者之增加幅度相若，而批發及零售業與工商企業經營概況調查結果相較，

則分別增加11.4%與9.7%，增長趨勢相近。

## 3.自有資產

觀察92年工商業母體資料檔企業單位自有資產預測值之整體變動趨勢，與公民營企業資金狀況調查結果相較，發現製造業兩年來分別增加12.9%與10.3%，兩者之增幅相近，而批發及零售業則分別增加45.2%與53.8%，工商業母體資料檔雖較公民營企業資金狀況調查結果減少8.6個百分點，惟兩者結果均顯示，批發及零售業之自有資產，應呈增

加趨勢，且增加幅度較大。

## 肆、研究發現

(一) 90年工商及服務業普查資料與相關公務及調查檔案進行比對，經研判後更新地址變更及停歇業廠商計16萬1,393家（或占16.4%），新增廠商28萬9,496家，有助掌握母體名冊及提升地址確度。

(二) 本研究運用40餘項判定準則，採多檔比對及同

| 行業別    | 工商業母體資料檔        |                |             | 92年與90年增減比較 (%) |                |
|--------|-----------------|----------------|-------------|-----------------|----------------|
|        | 92年預測值<br>(十億元) | 90年普查<br>(十億元) | 增減比較<br>(%) | 營利事業<br>銷售額     | 工商企業經<br>營概況調查 |
| 製造業    | 12,408.5        | 10,100.5       | 22.9        | 20.9            | -              |
| 批發及零售業 | 8,774.6         | 7,874.5        | 11.4        | -               | 9.7            |

註：因稅務行業標準分類修訂，無法單獨觀察批發及零售業之變動趨勢。

| 行業別    | 工商業母體資料檔        |                |             | 公民營企業資金狀況調查  |              |             |
|--------|-----------------|----------------|-------------|--------------|--------------|-------------|
|        | 92年預測值<br>(十億元) | 90年普查<br>(十億元) | 增減比較<br>(%) | 92年<br>(十億元) | 90年<br>(十億元) | 增減比較<br>(%) |
| 製造業    | 17,255.8        | 15,278.5       | 12.9        | 17,067.2     | 15,466.8     | 10.3        |
| 批發及零售業 | 7,991.6         | 5,504.7        | 45.2        | 5,840.7      | 3,797.6      | 53.8        |

檔變遷比對，可有效減少以往因單一公務檔案未及更新或企業登記與實際營業地址不同等造成之錯漏情形。另有7千餘家廠商（或占0.7%）需輔以人工檢核，鑒於本系統之判定準則可依各調查應用效果續予修訂，未來判定確度將續提升，人工成本可更為簡省。

（三）運用90年工商及服務業普查資料進行連結，若以統計調查檔案及公開資訊測試，可連結企業僅占8%，惟其營業收入占全體63%；而以營業規模10億元以上企業，連結成功率72%最佳，其營收可達同規模企業之86%。

（四）運用勞保檔進行連結，百人以上企業之連結成功家數及人數均逾93%。惟受加保對象之法

令限制影響，未滿5人之事業單位未強制納保，其連結率僅2成左右，致全體企業家數連結成功率31%，員工人數涵蓋率74%。

（五）連結製造業及批發零售業廠商，再運用分類與迴歸樹分析法及指數平滑法進行插補後，與相關調查比較顯示，其變動趨勢大抵一致，其中92年員工增減幅度與受雇員工薪資調查差異在4個百分點以內，製造業營業收入與營利事業銷售額資料、批發零售業與工商企業經營概況調查相差均約2個百分點，自有資產與公民營企業資金狀況調查差異則在3至9個百分點間。

### 參考文獻

1. 中央銀行經濟研究處（2003），90年中華民國臺灣地區公民營企業資金狀況調查結果報告。

2. 中央銀行經濟研究處（2005），92年中華民國臺灣地區公民營企業資金狀況調查結果報告。
3. 行政院主計處（2003），90年臺灣地區工商及服務業普查報告。
4. 行政院主計處（2004），92年中華民國臺灣地區薪資與生產力統計年報。
5. 行政院主計處第四局（2004），運用公（業）務檔案連結及輔助統計調查之經驗》第一輯《〈〈。
6. 財政部統計處（2005），中華民國臺灣地區財政統計月報。
7. 黃登源（2005），市場調查與分析暨市場調查資料分析實作講義。
8. 張淑珍（2005），「利用一次性的SQL改良決策樹建立信用卡審核之信用評等」，東吳大學商學院資訊科學系碩士論文。
9. 經濟部統計處（2004），92年中華民國臺灣地區工商企業經營概況調查年報。
10. 經濟部統計處（2005），國內外經濟統計指標速報。
11. 經濟部統計處（2005），92年中華民國臺灣地區工業統計調查報告。
12. 韓歆儀（2004），「應用兩階段分類法提昇SVM法之分類準確率」，國立成功大學工業與資訊管理研究所碩士論文。
13. 羅國華、劉惠玲、周元暉（2003），「資訊技術在戶口及住宅普查上之應用與發展」，主計月刊，No.576，頁7~18。❖