

第四章 研究設計

在開始進行實證研究前，本章將先詳細說明本研究使用之研究設計。首先，在第一節中將闡明使用追蹤資料進行研究的原因，並比較其下的固定效果模型和隨機效果模型（random-effect model）之異同和取捨方法。³⁴接著，第二節中將建立本研究的實證模型和說明實證資料之來源。至於依據本研究實證模型設立之實證變數假設，將陳述於第三節中。

第一節 研究方法說明

如前所述，本研究的目的是在於瞭解外商直接投資中國的決定因素，並決定以特定國家的追蹤資料為觀察對象進行研究，藉此得到最佳的估計。然而，擇定追蹤資料並非隨意挑選，在本節中便將對該擇定理由加以說明，並更深入地探討其下的固定效果模型和隨機效果模型。

一、追蹤資料的使用

追蹤資料是指針對某一特定調查對象組群，鎖定這些組群持續一段時間所得到的各種資料。由於是時間序列資料（time-series data）和橫斷面資料（cross-section data）的合併使用，因此不但擁有時間序列的動態性質，而且又能兼顧橫斷面資料可以表達不同樣本間特性的優點。和橫斷面資料相較之下，不但得觀察到變數間原來無法觀察到的動態變化，甚者得使用固定效果模型來分析觀察個體的固定效果，瞭解個體的特性，藉此降低參數估計上的偏誤。因此，若想要分析某觀察群體長期性的決定因素，使用追蹤資料較能分析出正確且嚴謹的結果，而這正符合本研究的要求。

³⁴ 本文研究方法主要參閱Hsiao（1986）及Greene（2000）。

Hsiao (1986) 認為追蹤資料與橫斷面或時間序列資料相較之下，除了提供更多的樣本數，改進估計參數時的效率外，尚可歸納出下列優點。

1. 降低估計上的偏誤

如果設定模型時忽略某一顯著變數，則在最小平方法 (OLS) 的估計下，將產生參數偏誤的情形。由於設定橫斷面資料的模型時，未含潛在重要的質化資料 (qualitative data) 因此可能產生偏誤。但若用追蹤資料中的固定效果模型進行估計，則可在模型中以虛擬變數 (dummy variable) 的形式代表一些難以獲得的質化變數，因而得以避免因模型不完整而產生的估計偏誤問題。

2. 減少共線性的問題

多數時間序列資料存在缺乏自由度及高度共線性 (multicollinearity) 的問題，這是源於資料本身期間太短及現有資訊無法滿足模型設定所導致。此時，若使用含有橫斷面性質的追蹤資料，除了能提供比較多的樣本數滿足模型的資訊及自由度外，不同觀察個體間的特質差異也會比使用單一觀察個體的時間序列資料來得大，因此會減少共線性的產生。

3. 提供更完整的訊息

一般而言，時間序列資料比較能夠解釋觀察個體本身 (intra-individual) 的動態變化，而橫斷面資料所估計的參數比較能夠解釋不同觀察個體間 (inter-individual) 的差異。追蹤資料由於包含以上兩者的特點，因而得以涵蓋更完整的訊息，藉此降低模型設定的錯誤及估計參數時的誤差。

由於追蹤資料擁有上述眾多的優點，因此本研究採用其資料型態，來進行外商直接投資中國的研究，希望藉以得到更佳之估計結果。但追蹤資料又可區分為固定效果模型和隨機效果模型，以下將繼續對上述兩者進行

更深入的探討。

二、固定效果模型

固定效果模型又稱為虛擬變數模型 (Least Square Dummy Variable Model; LSDV) 或共變異數模型 (Covariance Model)。其模型特點除了可同時考慮橫斷面及時間序列並存的資料之外，並允許觀察個體間有差異性的存在，即具固定效果。每個觀察個體在該模型中的截距項皆固定，展現其各自獨具的特質。截距項隨著不同個體及不同期間而變動，並假設其中觀察到的個體效果 (W_i) 以及時間效果 (Z_t) 和其他自變數 (X_k) 相關。

³⁵固定效果模型的迴歸式可表示為：

$$Y_{it} = \beta_{1it} + \sum_{k=2}^K \beta_k X_{kit} + e_{it}, \quad i=2, \dots, N, t=2, \dots, T \quad (1)$$

其中， Y 為應變數， X 為自變數， e 為隨機殘差項， β_{1it} 為截距項變數， β_k 為斜率項變數， i 代表第 i 個個體， t 代表第 t 期時間，至於 k 則代表第 k 個外生變數 (exogenous variables)， $\beta_{1it} = \beta_1 + \mu_i w_i + \lambda_t z_t$ 。³⁶對於 i 個體而言， $i=2, \dots, N$ 時 $w_i=1$ ，反之則 $w_i=0$ ；對於 t 個體而言， $t=2, \dots, T$ 時 $z_t=1$ ，反之則 $z_t=0$ 。同時，並假設 $E(e_{it})=0$ 且 $E(e_{it}^2)=\sigma_e^2$ 。若此模型的參數以最小平方法來估計，將可得到不偏且一致的參數估計，自由度為 $NT-K-N-T+2$ 。假設只考慮個體效果，則對於 i 個體，模型可改寫為：

$$y_i = (\beta_1 + \mu_i) j_T + X_{si} \beta_s + e_i, \quad (2)$$

其中， $y_i = (Y_{i1}, Y_{i2}, \dots, Y_{iT})'$ ， $e_i = (e_{i1}, e_{i2}, \dots, e_{iT})'$ ， $j_T = (1, 1, \dots, 1)'$ ， X_{si} 維度 (dimension) 為 $T \times K'$ ， $K' = K-1$ ， X_{si} 為不包括常數項之所有解釋變數，而 $\beta_s = (\beta_2, \beta_3, \dots, \beta_k)'$ ，為需要被估計之斜率項係數。

³⁵ $W_i = \beta_1 + \mu_i w_i$ ； $Z_t = \beta_1 + \lambda_t z_t$ 。

³⁶ 令 $\mu_1 = \lambda_1 = 0$ 。

由於固定效果模型中， β_{1i} 已被設定為固定參數且令 $\mu_i=0$ ，此時若再令 $\beta_{1i}=\beta_1+\mu_i$ ，則完整的模型應該表示為：

$$y = [I_N \otimes j_T \quad X_s] \begin{pmatrix} \beta_r \\ \beta_s \end{pmatrix} + e, \quad (3)$$

其中， $y'=(y_1', y_2', \dots, y_N')$ ， $X_s'=(X_{s1}', X_{s2}', \dots, X_{sN}')$ ，

$e'=(e_1', e_2', \dots, e_N')$ ， $\beta_r'=(\beta_{11}, \beta_{12}, \dots, \beta_{1N})'$ ，

$$I_N \otimes j_T = \begin{bmatrix} 11\dots 1 & & & \\ & 11\dots 1 & & \\ & & \ddots & \\ & & & 11\dots 1 \end{bmatrix},$$

最後，關於模型是否存在固定效果之問題，可用 F 統計量來檢定。

令虛無假設 $H_0: \beta_{11}=\beta_{12}=\dots=\beta_{1N}=\beta_1$

對立假設 H_1 : 各項不全相等

$$F_{N-1, NT-K-N+1} = \frac{(ESS_1 - ESS_2)/(N-1)}{ESS_2/(NT-K-N+1)} \quad (4)$$

其中， ESS_1 、 ESS_2 分別為只考慮個體效果之簡單組合模型與固定效果模型誤差平方和。當F統計量大於其對應的臨界值時將拒絕虛無假設，表示有固定效果的存在。

最後，值得注意的是，固定效果模型的使用將產生一些附帶問題，例如自由度的消耗以及並未明確指出造成迴歸估計式平移之原因。

三、隨機效果模型

隨機效果模型又稱為誤差成分模型 (error component model)。之前由於對模型的認知不足，懷疑有遺漏變數，因此在模型中加入虛擬變數而成為固定效果模型。但當同樣從干擾項著手，卻假設未觀察到的個體效果和時間效果與其他自變數無關時，個體間的固定差異將隨不同個體而隨機分布。此即為隨機效果模型，其迴歸式可表示為：

$$Y_{it} = \sum_{k=2}^K \beta_k X_{kit} + \varepsilon_{it}, \quad i=2, \dots, N, t=2, \dots, T \quad (5)$$

其中， $\varepsilon_{it} = \beta_1 + \mu_i + \lambda_t + e_{it}$ ，並假設 $E(\mu_i)=0$ 、 $E(\lambda_t)=0$ 、 $E(\mu_i^2)=0$ 、 $E(\lambda_t^2)=\sigma_\lambda^2$ 、 $E(\mu_i\mu_j)=0 \quad i \neq j$ 、 $E(\lambda_i\lambda_j)=0 \quad i \neq j$ 。與固定效果模型相似的是， μ_i 以及 λ_t 分別代表了個體效果和時間效果。至於其相異處，則隨機效果模型假設 μ_i 和 λ_t 為隨機變數，可視為 N 個個體和 T 個期間是由一龐大母體所抽出的隨機樣本。假設只考慮個體效果，於是對於 i 個體可將模型改寫為：

$$y_i = X_i\beta + \mu_i j_T + e_i, \quad (6)$$

其中， $\beta=(\beta_1, \beta_2, \dots, \beta_k)$ ， X_i 維度為 $T \times K$ ，含常數項。若考慮 N 個個體， T 時間長度下，完整的模型應該表示為：

$$y = X\beta + \mu \otimes j_T + e, \quad (7)$$

其中， $y'=(y'_1, y'_2, \dots, y'_N)$ ， $X'=(X'_1, X'_2, \dots, X'_N)$ ， $\mu=(\mu_1, \mu_2, \dots, \mu_N)'$ ， $e'=(e'_1, e'_2, \dots, e'_N)$ 。此外，隨機效果模型的估計，採一般化最小平方法 (generalized least-squares, GLS) 的估計方式。為了簡化起見，此處只考慮個體效果，並令 $w_{it}=e_{it}+\mu_{it}$ ，於是可以得到 $E(w_{it}^2)=\sigma_e^2+\sigma_\mu^2$ ， $E(w_{it}w_{is})=\sigma_\mu^2$ ， $t \neq s$ 。對於 i 個體而言，令 $\Omega=E[(w_{i1}, w_{i2}, \dots, w_{iT})(w_{i1}, w_{i2}, \dots, w_{iT})']$ ，即：

$$\Omega = \begin{bmatrix} \sigma_e^2 + \sigma_\mu^2 & \sigma_\mu^2 & \cdots & \sigma_\mu^2 \\ \sigma_\mu^2 & \sigma_e^2 + \sigma_\mu^2 & \cdots & \sigma_\mu^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\mu^2 & \sigma_\mu^2 & \cdots & \sigma_e^2 + \sigma_\mu^2 \end{bmatrix} = \sigma_e^2 I_T + \sigma_\mu^2 i_T i_T', \quad (8)$$

其中， i 為 $T \times 1$ 之向量，內容全為 1。而由於任兩個體 i 、 j 間互為獨立，因此，整個模型干擾項之共變異矩陣（covariance matrix）為如下之形式：

$$V = \begin{bmatrix} \Omega & 0 & \cdots & 0 \\ 0 & \Omega & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Omega \end{bmatrix} = \Omega \otimes I_N,$$

經由 (8) 式，可以得到：

$$\Omega^{-1/2} = \frac{1}{\sigma_e} \left[I - \frac{\theta}{T} i i' \right], \quad \text{且 } \theta = 1 - \frac{\sigma_e}{\sqrt{T\sigma_\mu^2 + \sigma_e^2}}.$$

於是，為了進行 GLS 估計，我們將 (5) 式中所有自變數、應變數加以轉換並重新進行迴歸與估計後，可得以下結果：

$$\Omega^{-1/2} Y_i = \frac{1}{\sigma_e} \begin{bmatrix} Y_{i1} - \theta \bar{Y}_i \\ Y_{i2} - \theta \bar{Y}_i \\ \vdots \\ Y_{iT} - \theta \bar{Y}_i \end{bmatrix}, \quad \text{且 } \Omega^{-1/2} X_{ki} = \frac{1}{\sigma_e} \begin{bmatrix} X_{ki1} - \theta \bar{X}_{ki} \\ X_{ki2} - \theta \bar{X}_{ki} \\ \vdots \\ X_{kiT} - \theta \bar{X}_{ki} \end{bmatrix}$$

由此並可得參數估計如下：

$$\hat{\beta} = \hat{F}^w b^w + (1 - \hat{F}^w) b^b, \quad (9)$$

$$\text{其中， } \hat{F}^w = [S_{xx}^w + \lambda S_{xx}^b]^{-1} S_{xx}^w, \quad \lambda = \frac{\sigma_e^2}{\sigma_e^2 + T\sigma_\mu^2} = (1 - \theta)^2.$$

然而實際上 Ω 往往是無法事先得知的，需要先對 σ_e^2 和 σ_μ^2 進行估計，在

針對模型的參數進行估計與檢定，這樣的過程便是合理一般化最小平方法（feasible generalized least squares, FGLS）。經由推導，對於 σ_e^2 可使用一不偏之估計式如下：

$$\hat{\sigma}_e^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T (e_{it} - \bar{e}_i)^2}{NT - N - K}, \quad (10)$$

其中， e_{it} 是經由估計固定效果模型之過程而產生。而後將（5）式加以變化如下：

$$\bar{Y}_i = \beta_1 + \sum_{k=2}^K \beta_k \bar{X}_{ki} + \bar{e}_i + \mu_i, \quad (11)$$

$$\text{可推得： } e_{**i} = \bar{Y}_i - \beta_1 - \sum_{k=2}^K \beta_k \bar{X}_{ki} = \bar{e}_i + \mu_i, \quad \text{Var}(e_{**}^2) = \sigma_{**}^2 = \frac{\sigma_e^2}{T} + \sigma_\mu^2, \quad (12)$$

在對於（11）式的估計過程中，可以得到 $\text{Var}(e_{**i})$ 的不偏估計值 $\hat{\sigma}_{**}^2$ ，因此便得到了 $\hat{\sigma}_\mu^2$ 的估計值，如（13）式。如此一來，便可以回到GLS的程序，估計模型的參數。

$$\hat{\sigma}_\mu^2 = \hat{\sigma}_{**}^2 - \frac{\hat{\sigma}_e^2}{T}, \quad (13)$$

至於隨機效果模型的檢定，除了可利用之前所敘述過的F統計量檢定外，另可採LM統計量檢定方法。³⁷其檢定方式如下：

令虛無假設 $H_0 : \sigma_\mu^2 = 0$

對立假設 $H_1 : \sigma_\mu^2 \neq 0$

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^N \left[\sum_{t=1}^T e_{it} \right]^2}{\sum_{i=1}^N \sum_{t=1}^T e_{it}^2} - 1 \right]^2 \quad (14)$$

³⁷ LM統計量檢定方法可參閱Breusch and Pagan (1980)。

其中， e_{it} 為OLS模型的殘差值。當LM統計量大於其對應的臨界值時將拒絕虛無假設，表示有隨機效果的存在。

四、Hausman 檢定

前述之固定效果模型和隨機效果模型的選擇標準，目前為止看法迥異最簡單的方法是以「樣本有無透過抽樣過程」來分辨，若樣本為母體或無透過抽樣過程，則採用固定效果模型較佳。至於計量方法上，則通常透過Hausman (1978) 的方法檢定。由於固定效果模型和隨機效果模型最大的差異，在於隨機效果模型中的 u_i 與解釋變數間是否具有相關性。若有，則固定效果模型之估計將具一致性和有效性，而隨機效果模型之估計將不具一致性，故應採用固定效果模型；若無，隨機效果模型之估計將具一致性和有效性，而固定效果模型之估計式將具一致性但不具有有效性，故應採用隨機效果模型。其檢定方式如下：

令虛無假設 $H_0: E(u_i X_{it})=0$ ，即 u_i 與解釋變數間不具相關性；

對立假設 $H_A: E(u_i X_{it})\neq 0$ ，即 u_i 與解釋變數間具相關性。

Hausman 的檢定統計量可表示如第 (15) 式。

$$H = (b_{fix} - b_{ran})'(M_{fix} - M_{ran})^{-1}(b_{fix} - b_{ran}) \sim X^2(K) \quad (15)$$

第 (15) 式中， H 統計量呈 χ^2 分配， K 為其自由度。而 b_{fix} 與 b_{ran} 分為固定係數與隨機係數的參數估計值， M_{fix} 與 M_{ran} 分別為其共變異矩陣。當 $H < X^2(K)$ 時無法拒絕虛無假設，即兩模型之估計無差異，或 u_i 與解釋變數間不具相關性，則選擇隨機效果模型；反之，當 $H > X^2(K)$ 時拒絕虛無假設，即兩模型之估計有差異，或 u_i 與解釋變數間具相關性，則應選擇固定效果模型為佳。

第二節 實證模型設定與資料來源

一、本研究之實證模型設定

如前所述，本文使用結合橫斷面與時間序列的追蹤資料，其優點是相較於橫斷面與時間序列兩種資料，追蹤資料包含更多的資訊。如此，不但有助於提高樣本數與自由度，使估計結果較為準確；而實證模型的採用也將更具有更多的選擇。而在回顧眾多探討外商直接投資決定因素的文獻後，歸納出相對國內生產毛額、相對每人國內生產毛額、相對工資率、對中國貿易依賴程度、相對匯率、相對借貸成本、相對國家風險和地理距離等，都可能是影響跨國企業對外直接投資的重要決定因素。據此，本文設出下列的關係式：

$$FDIR = f(GDP, GDPPC, WAGE, EXIMRA, EXR, RBC, RCR) \quad (16)$$

第(16)式中， $FDIR$ 為每年各國在中國的實際直接投資額， GDP 為每年中國和各國實質國內生產毛額之比率， $GDPPC$ 為每年中國和各國實質每人國內生產毛額之比率， $WAGE$ 為每年中國和各國實質工資率之比率， $EXIMRA$ 為每年各國對中國的貿易依賴程度， EXR 為每年中國和各國實質匯率之比率， RBC 為每年中國和各國實質借款利率之比率，而 RCR 則為每年中國和各國國家風險之比率。第(16)式可以進一步表示為：

$$\begin{aligned} \log(FDIR)_{it} = & \beta_{0i} + \beta_1 GDP_{i(t-1)} + \beta_2 GDPPC_{i(t-1)} + \beta_3 WAGE_{i(t-1)} \\ & + \beta_4 EXIMRA_{i(t-1)} + \beta_5 EXR_{i(t-1)} + \beta_6 RBC_{i(t-1)} + \beta_7 RCR_{i(t-1)} \\ & + \beta_8 T + \varepsilon_{it} \end{aligned} \quad (17)$$

第(17)式中， i 表示國家或地區， t 代表時期， ε_{it} 則為殘差項。而為了探討外商直接投資在研究期間內，是否存在某種時間趨勢，本研究加入時間趨勢變數(T)加以研究。此外，由於各投資國可能存在各別特質，

而使各投資國的跨國企業在中國進行直接投資時，產生不同的總直接投資額。為了衡量這些個別特質的型態，本文在估計的實証模型中，加入各投資國的特質效果 (national-specific effect) β_{oi} 。因此，第 (17) 式的實証估計模型為固定效果模型。若假設這些個別特質效果，是以隨機變動形式來影響各投資國在中國的總直接投資額。也就是說， $\beta_{oi} = \bar{\beta}_0 + u_i + \lambda_i$ ， $u_i \sim \text{iid} (0, \sigma_u^2)$ 且 $\lambda_i \sim \text{iid} (0, \sigma_\lambda^2)$ ，則實証估計模型將改為隨機效果模型。至於使用哪種模型來估計較為適當，本文將以前述 Hausman (1978) 的方法檢定。

最後，為了瞭解第 (17) 式中的截距項，是否與地理距離有關聯，本文另設以下實証模型：

$$b_{oi} = \alpha_0 + \alpha_1 \log(\text{GEOD})_i + \gamma_i \quad (18)$$

在第 (18) 式中， b_{oi} 為 β_{oi} 之估計值， γ 代表該式的殘差項， GEOD_i 表示第 i 個投資國和中國的地理距離。之所以將第 (17) 式中的截距項與各投資國和中國的地理距離單獨提出作迴歸分析，主要是因為如此將可進一步瞭解，各投資國對中國的自發性直接投資 (autonomous investment) 狀況與地理距離是否有關。即可研究即便是在誘發性投資 (induced investment) 為零的情況下，為何有些國家對中國的自發性直接投資仍特別多。

二、本研究之資料來源

至於實証資料來源方面，由於本文將焦點置於 1992 年鄧小平南巡後，中國外商直接投資開始蓬勃發展的時期。因此，涵蓋年份從 1993 年到 2003 年，共 11 年的資料。³⁸ 至於追蹤對象方面，過去的研究多僅挑選數個國家

³⁸ 本文假設政策及各項因素改變所造成的影響具有延遲性，故 1992 年的政策 1993 年才開始發生作用。

為研究標的。因此，到目前為止，僅Wei and Liu (2001) 在研究中國 1983 至 1998 間的合作外商直接投資因素時，涵蓋 87% 的所有直接投資額，以及在研究 1984 年至 1998 年間的實際外商直接投資因素時，涵蓋 88% 的所有直接投資額，此文為涵蓋範圍比率最高的文獻。本文為求研究資料的最大涵蓋範圍比率，在尋找全球所有國家（地區）之變數資料後，選擇 24 個資料齊全的國家或地區做為研究對象，並涵蓋了近 96% 的實際直接投資額。儘管這樣的作法極耗人力、物力與時間，但若能將外商直接投資的範圍比率，盡量提高趨近到 100%，則作出來的研究也將為最接近事實的估計值。

在整合上述的外商直接投資理論與相關實證文獻，並逐一檢討各項可能的外商直接投資中國的決定因素後，本文由國家特性比較的觀點，歸納出前述可能影響世界各國到中國進行直接投資的主要原因，並加以驗證。由於所有的變數，包含世界各國的資料，因此本文的變數資料來源較為複雜，故每一變數的預計的資料來源，將分述如下。

各投資國對中國的實際外商直接投資額，以及貿易總額來自於《中國統計年鑑》；各國的國內生產毛額、每人國內生產毛額、匯率以及借款利率來自於中國國家統計局出版的《國際統計年鑑》以及國際貨幣基金會出版的 *International Financial Statistics Yearbook*；各國工資率來自於國際勞動組織出版的 *Yearbook of Labour Statistics* 以及其網站上的資料庫。³⁹ 由於中國外商直接投資的行業以製造業為主，因此工資率的選定，主要是著重在各國的製造業；國家風險的資料來自於 *Institutional Investor* 期刊每年對各國風險的評分；⁴⁰ 至於各投資國到中國的地理距離定義為，由中國首都北

³⁹ 國際勞動組織網站為 <http://laborsta.ilo.org/>。

⁴⁰ 其對各國國家風險的評定由 0 分到 100 分，分數越高，發生金融危機的機率便越小。其評估所需的資訊來自於 Senior Economists and Sovereign Risk Analysts 此一涉略全球銀行、金融管理與安全之機構。

京市到各國首都的距離，該資料可在BALI & INDONESIA ON THE NET網站上獲得。⁴¹

在所有國家或地區之中，以台灣的資料最為特別。因為政治因素，使其許多項資料在以上提到的各資料來源中從缺。因此，有關台灣的資料，主要是根據台灣政府所提供的官方資料為主。以上各變數，都以 1993 年為基期的國內生產毛額平減指數，或消費者物價指數加以平減，以消除物價波動的影響。

⁴¹ BALI & INDONESIA ON THE NET的網站為<http://www.indo.com/distance/>。

第三節 實證變數假設

由前所述，投資國和地主國間的相對國內生產毛額、相對每人國內生產毛額、相對工資率、對地主國貿易依賴程度、相對匯率、相對借貸成本、相對國家風險和地理距離，都是影響外商直接投資規模的重要指標客觀因素。以下針對上述幾項的變數，加以詳細說明其對於投資國對地主國進行投資的影響。

投資國和地主國間的相對國內生產毛額，具有有兩層的意義。首先，它代表了相對的市場大小。一國的國內生產毛額越高，代表了其潛在市場越大。由於投資國的企業赴地主國進行直接投資時，也會著眼於地主國市場的大小，以利產品的銷售。因此，本文假設地主國國內相對的市場越大，對投資國尋求市場的廠商吸引力也越大，地主國所吸引到的外商直接投資規模也會越大。Wang and Swain (1995)、Liu et al. (1997)、Wei and Liu (2001) 等對以中國為研究對象的實證研究，皆支持了該觀點。而 Pitelis (1996) 則認為，投資國國內需求的不足，導致其廠商不得不對外進行直接投資以擴張市場規模。因此，投資國國內市場的大小與地主國的外商直接投資規模具有負向關係。

而從另一方面來看，投資國的國內生產毛額越高，也代表了該國具有較多的大型企業，這些大企業具有較多的所有權優勢 (ownership-specific advantage) 可以從事對外直接投資。故投資國的國內生產毛額規模可視為擁有大型企業多寡的代理變數。Ajami and BarNiv (1984) 對美國外商直接投資的研究，也證實了此項關係。因此，投資國國內生產毛額規模越大，對外直接投資的規模也會越大。綜合以上兩個觀點，投資國和地主國間的相對國內生產毛額，對投資國對地主國進行投資的影響並不確定，端視兩正負影響力間的強弱而定。

此外，投資國對地主國的相對人均國內生產毛額，代表了投資國和地主國間的相對經濟發展程度。通常經濟發展程度較高的國家，其市場化程度也會較高，越趨近於開放型經濟。從各國經驗看，越是開放的國家（地區），技術進步越快，國際競爭力越強，因此可將人均國內生產毛額視為國際競爭力的代理變數。故本文假設，投資國人均國內生產毛額越大，對外直接投資也會越多。

而在所有生產要素成本中，相對工資的高低是極重要的一項考量因素。特別是對勞力密集的製造業而言，更是如此。Lucas（1993）指出，一國工資率的提高會促使該國生產成本提高，進而降低該國所吸收到的外商直接投資額，反之亦然。這種工資率和外來直接投資間的負向關係，也在許多研究中被證實。例如，Moore（1993）、Kumar（1994）、Wei and Liu（2001）、Wheeler and Mody（1992）皆得到類似的結論。因此，本文假設地主國相對的工資率越低，其吸引到的外商直接投資的金額也就越大。

另一項重要的可能決定因素為投資國對地主國的貿易依賴程度。當貿易依賴程度高時，意味著投資國與地主國的關係越親密。Wei and Liu（2001）指出，投資國和地主國間的關係越密切，而使得投資國的企業在地主國的市場上能獲得更充分的訊息，進而擁有更多的投資機會。另外，地主國的進口額有部分為外商從其母國的基地進口機械或材料等；相反的，出口額也有部分代表了外商在地主國的新投資廠，回銷成品或半成品回投資國。這樣的關係，在高度開發國家向開發中國家進行直接投資時更是常見。⁴²如果由上述的理由推論，可得到投資國對地主國的貿易依賴程度越高，則投資國流入地主國的外商直接投資規模也會越大的推論。

相對匯率的變動，藉由許多方式影響外商直接投資的規模。Froot and

⁴²例如，台灣廠商到中國進行直接投資時，運送生產財過去進行生產，再把半成品運回台灣加工，或回銷消費財，便是一個顯而易見的例子。

Stein (1991) 認為，國際間各主要貨幣價值的變化，在一定程度上導致了各國金融資產和實物資產價值的短期波動。這種波動一方面會導致企業在國內外融資成本的變化，另一方面會導致企業財富儲量的短期變化。換言之，投資國匯率的相對攀升（即貶值），會使得投資國廠商的資產隨之降低，因而降低對外直接投資的能力。即投資國流入地主國的外商直接投資規模也會跟著減少，反之亦然。Cushman (1985, 1987) 和 Culem (1988) 則基於另一個觀點指出，地主國貨幣貶值使投資國廠商在每單位的本國幣下能雇用更多的勞動力。因此，會促成地主國外商直接投資的增加。綜合以上所述，本文假設，投資國的貨幣相對於地主國升值時，將會增加對地主國的外商直接投資。

相對借貸成本，也是討論外商直接投資理論中常被探討的變數。Aliber (1970) 認為，由於外商直接投資的資金通常來自於投資國，若投資國的借貸成本相對小於地主國，則將降低投資國廠商所須付出的成本，使得他們更樂於以外商直接投資的形式向地主國投資。同樣的觀點在實證研究中，也為 Pain (1993, 1997)、Barrel and Pain (1996)、Grosse and Trevino (1996) 以及 Wei and Liu (2001) 等人所證實。因此，本文假設地主國相對於投資國的借貸成本越高，所吸引的外商直接投資也越多。

就一般所觀察到的現象而言，投資者都會趨向於避免不確定性或風險。因此，地主國與投資國相較之下的政治、經濟、社會狀況越穩定，外商在地主國市場上所能得到的資訊就越多，其所承受的風險也就越小，外商也越敢放心對地主國進行投資，投資的規模也趨越大。Wei and Liu (2001) 使用 1983 年至 1998 年銀行破產風險衡量中國的國家風險時發現，外商直接投資顯著地與國家風險呈負相關。然而，在 Korbrin (1979) 和 Tallman (1988) 中，此項關係並未被證實。而 Grosse and Trevino (1996) 則發現，外商直接投資和國家風險間僅具有微弱的負相關。

投資國與地主國間的地理距離，也是影響外商直接投資的重要因素。地理距離的遠近不僅影響運輸成本，也影響了管理成本，所以地理距離越遠，運輸成本和管理成本越高。Davidson (1980) 指出地理距離的接近能夠減少訊息和管理上的不確定性，減少監視成本並降低廠商的風險。另外，投資國與地主國間的地理距離，亦可以用來作為兩國文化差異的替代變數。距離越遠，國與國之間文化的外溢與趨同的力量可能越薄弱，致使文化差異越大。而文化差異越大，所代表的亦是訊息和管理上的越不確定性。因此，本文假設投資國和地主國地理距離越遠，地主國所吸引的外商直接投資也越少。以上所述各項實證變數的說明與基本統計量，均彙整列於表 6 之中。

表 6：實證變數之基本統計量

變數	變數說明	平均值	標準差	預期影響方向
實際外商直接投資額	每年各國在中國的實質直接投資額（單位：百萬美元）	1102.09	2781.13	
相對國內生產毛額	每年中國實質國內生產毛額/各國實質國內生產毛額	51.74	211.44	?
相對每人國內生產毛額	每年中國實質每人國內生產毛額/各國實質每人國內生產毛額	0.18	0.33	-
相對工資率	每年中國實質工資率/各國實質工資率	0.16	0.29	-
對中國貿易依賴程度	每年中國與各國的貿易額/每年各國總貿易額	0.03	0.38	+
相對匯率	每年中國實質匯率/各國實質匯率	20.41	47.55	+
相對借貸成本	每年中國實質借款利率/各國實質借款利率	-6.05	14.19	+
相對國家風險	每年中國國家風險評分/各國國家風險評分；其對各國國家風險的評定由 0 分到 100 分，分數越高，發生金融危機的機率便越小。	0.98	0.51	+
地理距離	各國家（地區）首府到中國首都北京市之距離（單位：公里）	7580.00	4154.07	-

資料來源：實際外商直接投資額，以及進出口總額來自於《中國統計年鑑》；各國的國內生產毛額、每人國內生產毛額、匯率以及借款利率來自於中國國家統計局出版的《國際統計年鑑》以及國際貨幣基金會（International Monetary Fund, IMF）出版的 *International Financial Statistics Yearbook*；各國工資率來自於國際勞動組織（International Labour Organization, ILO）出版的 *Yearbook of Labour Statistics* 以及其網站上的資料庫。由於中國外商直接投資的行業以製造業為主，因此找各國工資率資料時，選定的是製造業的工資率；國家風險的資料來自於 *Institutional Investor* 期刊每年對各國風險的評分，其評估所需的資訊來自於 senior economists and sovereign risk analysts 此一涉略全球銀行、金融管理與安全之機構；至於各投資國到中國的地理距離定義為，由中國首都北京市到各國首都的距離，該資料可在 BALI & INDONESIA ON THE NET 網站上獲得。至於台灣的相關資料，則來自於其政府出版的官方資料。

第四節 本章小結

本研究的目的是在於瞭解 1993 至 2003 年間，外商直接投資中國的決定因素，並決定以特定國家的追蹤資料為觀察對象進行研究，藉此得到最佳的估計。本章先詳細說明本研究使用之研究設計。首先，在第一節中將闡明使用追蹤資料進行研究的原因，並比較其下的固定效果模型和隨機效果模型之異同和取捨方法。接著，第二節中建立本研究的實證模型和說明實證資料之來源。至於依據本研究實證模型設立之實證變數假設，則陳述於第三節中。本章經整理重點如下：

一、研究方法說明

本研究之設計採用結合橫斷面資料和時間序列資料的追蹤資料，因此不但擁有時間序列的動態性質，而且又能兼顧橫斷面資料可以表達不同樣本間特性的優點。和橫斷面資料相較之下，能觀察到變數間原來無法觀察到的動態變化，甚者得使用固定效果模型來分析觀察個體的固定效果，瞭解個體的特性。此外，尚有降低估計上的偏誤、減少共線性的問題以及提供更完整的訊息等好處。因此，符合本研究觀察群體長期性的要求。其中，追蹤資料模型又可分為固定效果模型和隨機效果模型，至於該使用何種模型較佳，則通常透過 Hausman (1978) 的方法檢定。

二、實證模型設定與資料來源

(1) 在實證模型設定方面，首先，分別以每年各國在中國的實際直接投資額為應變數，而以每年中國和各國實質國內生產毛額之比率、每年中國和各國實質每人國內生產毛額之比率、每年中國和各國實質工資率之比率、每年各國對中國的貿易依賴程度、每年中國和各國實質匯率之比率、每年中國和各國實質借款利率之比率，和每年中國和各國國家風險之比率為自變數，來建構出本研究主要的實證模型。接著，經由 Hausman (1978)

的方法檢定後，判斷模型應採用固定效果模型或是隨機效果模型。如果經檢定後應採用固定效果的模型，則為了瞭解主要實證模型的固定效果是否與地理距離有關聯，則另設迴歸式加以分析之。

(2) 在實證資料來源方面，上述實證模型中的各變數資料由於牽涉世界各國，因此資料來源非常複雜，主要來自於《中國統計年鑑》、《國際統計年鑑》、*International Financial Statistics Yearbook*、*Yearbook of Labour Statistics*、*Institutional Investor*，以及 BALI & INDONESIA ON THE NET 網站等。此外，由於本研究將焦點置於 1992 年鄧小平南巡後，中國外商直接投資開始蓬勃發展的時期。因此，涵蓋年份從 1993 年到 2002 年，共 10 年的資料。並且，在尋找全球所有國家之變數資料後，選擇 24 個資料齊全的國家或地區做為研究對象，並涵蓋了近 96% 的實際直接投資額。

三、實證變數假設

實證模型設定後，本研究對實證變數進行假設。經由過去相關文獻的探討後，本文初步的假設為：負相關的變項有相對每人國內生產毛額、相對工資率和地理距離；正相關的變項有貿易密切程度、相對匯率、相對借貸成本以及相對國家風險。至於相對國內生產毛額變項則無法預知，端視其內兩力量的強弱而定。