

第四章、實證模型

第一節 模型介紹

(一) PROBIT MODEL

(1) 概念簡介

在許多的情形下，我們尋求的模型建立並不是連續的，而是離散的。例如：考慮勞動力的參與與否或是在一次的總統大選中人們會選擇投票給哪一位候選人等議題的研究。對於勞動力的參與原因，直覺的原因可能有年齡、教育、婚姻狀況、受扶養子女的數目以及其他的一些經濟因素來解釋人們為何選擇加入勞動市場。諸如此類的選擇性問題，如果我們採用一般的線性迴歸模型來作實證分析，顯然會出現一些問題。

在離散選擇模型 (discrete choice model) 中，包括了兩項和多項的形式，例如：前文中所舉例的勞動力參與與否，其被解釋變數 (dependent variable) 僅僅只有選擇工作和選擇不工作兩個結果，這就是所謂的兩項形式；另一方面，前文中舉的另一個總統大選的例子，假設總統候選人有兩人以上，則其被解釋變數就是總統候選人的人數，這就是所謂的多項形式。經濟學中較常研究的議題是兩項模型，本文中所使用的 PROBIT MODEL 就是屬於二元選擇模型 (Binary Choice Model)。

(2) 模型方法簡介

考慮一個納稅義務人捐贈與否的模型，在調查納稅義務人捐贈 ($Y=1$) 和不捐贈 ($Y=0$) 的反應情形，我們可以大致上的推估影響納稅義務人捐贈

與否的因素的 x 向量可能包括婚姻狀況、扶養子女數目、本人薪資所得、可支配所得、配偶的薪資所得等等的因素，所以模型可以寫成：

$$\text{Pr ob}(Y = 1) = F (\mathbf{b} \cdot x)$$

$$\text{Pr ob}(Y = 0) = 1 - F (\mathbf{b} \cdot x)$$

參數集合 \mathbf{b} 反映 x 中的變化對於機率的影響，例如在各種的因素中我們比較感興趣的是婚姻狀況對於捐贈與否的邊際影響（marginal effect）。此時的問題是設計方程式右邊的一個適當模型。

一個可能是線性迴歸

$$F(x, \mathbf{b}) = \mathbf{b} \cdot x$$

由於 $E[y] = F(x, \mathbf{b})$ ，所以我們可以得到迴歸模型

$$y = E[y] + (y - E[y]) = \mathbf{b} \cdot x + \mathbf{e}$$

但是利用線性迴歸模型存在許多的問題，其中最大的問題就是因為機率值是介於 0 1 之間，所以我們必須令 $\mathbf{b} \cdot x$ 也只能介於 0 1 之間。若要達成這個限制只能在 x 值是有界限的或是 \mathbf{b} 值滿足某些限制，但是這個情況通常很難滿足，所以經濟學家採用二元選擇模型來解決這樣的問題。

給定一個迴歸向量，我們預期會產生：

$$\lim_{b \cdot x \rightarrow \infty} \text{Pr ob}(Y = 1) = 1$$

$$\lim_{bx \rightarrow -\infty} Prob(Y = 1) = 0$$

原則上，任何適當的、連續的機率分配都適用，常態分配在許多的分析中都被使用，所以產生了 Probit 模型

$$Prob(Y=1) = \int_{-\infty}^{bx} \Phi(t) dt = \Phi(\mathbf{b}'x)$$

二元選擇模型的估計通常使用最大概似法 (Maximum likelihood estimation)，設定模型如下：

$$\text{Max}_b \sum_{i=1}^n \ln f(y_i; x_i, \mathbf{b})$$

y_i 為白努利隨機變數 (Bernoulli random variable)

其機率密度函數 (Probability density function) 為

$$L = \prod_i [F(\mathbf{b}'x_i)]^{y_i} [1 - F(\mathbf{b}'x_i)]^{1-y_i}$$

取對數得到

$$\ln L = \sum_i [y_i \ln F(\mathbf{b}'x_i) + (1 - y_i) \ln(1 - F(\mathbf{b}'x_i))]$$

求得極大值的一階條件為

$$\frac{\partial \ln L}{\partial \mathbf{b}} = \sum_i \left[\frac{y_i f_i}{F_i} + (1 - y_i) \frac{-f_i}{(1 - F_i)} \right] x_i = 0$$

(二) TOBIT MODEL

(1) 概念簡介

所謂的 TOBIT MODEL 意指在考慮樣本具有截斷 (truncated) 和審查 (censored) 的性質時，近代的計量經濟學者所發展出來的一套適用的模型。假設我們考慮的樣本是屬於一個國家中高所得族群的經濟行為，如此一來，整個實證的樣本僅為所得高於某一特定的水準的個人而非所有的人民，這樣的情形我們就稱樣本具有截斷 (truncated) 的性質。至於被解釋變數的審查 (censored) 則是個體經濟學中常常出現的情況，當被解釋變數被審查時，某一個特定範圍內的值全部會被替換成單一的數值。舉例說明：研究人們對於工作小時數的選擇時，被解釋變數即為實際工作的小時數，假設實際工作的小時數為大於零的一個數值，則被解釋變數即為實際工作的小時數；反之，若實際工作的小時數為小於或等於零的數值，被解釋變數則一律以零為代表，由此可知，這個被解釋變數已經被審查 (censored) 了。

當樣本出現上述所提到的情形，但我們仍然使用 PROBIT MODEL 甚至是線性迴歸模型來進行實證分析時，將會出現估計上不可避免的錯誤，造成結果的誤判。所以為了解決這樣的問題，我們必須利用 TOBIT MODEL (= Truncated + probit model) 來進行實證研究。在本文的個人所得稅資料中，關於捐贈的部分也具有截斷和審查的性質，故本文部分的實證分析採用 TOBIT MODEL。

(2) 模型方法簡介

一般的 TOBIT MODEL 設定通常以一個指標函數 (latent variable) 的形式出現：

$$y_i^* = \mathbf{b}'x_i + e_i$$

$$y_i = 0 \quad \text{若 } y_i^* \leq 0$$

$$y_i = y_i^* \quad \text{若 } y_i^* > 0$$

當 $y_i = 0$, 若 $y_i^* \leq 0$ 時 ,

$$\begin{aligned} P_r(y_i = 0) &= P_r(y_i^* \leq 0) = P_r(x_i' \mathbf{b} + \mathbf{e}_i \leq 0) = P_r(\mathbf{e}_i \leq -x_i' \mathbf{b}) \\ &= P_r\left(\frac{\mathbf{e}_i - 0}{\mathbf{s}_e} \leq \frac{-x_i' - 0}{\mathbf{s}_e}\right) = P_r\left(Z \leq \frac{-x_i' \mathbf{b}}{\mathbf{s}_e}\right) \\ &= \Phi\left(\frac{-x_i' \mathbf{b}}{\mathbf{s}_e}\right) \end{aligned}$$

當 $y_i = y_i^*$, 若 $y_i^* > 0$ 時 ,

$$\begin{aligned} P_r(y_i = y_i^*) &= P_r(y_i^* > 0) = P_r(x_i' \mathbf{b} + \mathbf{e}_i > 0) = P_r(\mathbf{e}_i > -x_i' \mathbf{b}) \\ &= P_r\left(\frac{\mathbf{e}_i - 0}{\mathbf{s}_e} > \frac{-x_i' - 0}{\mathbf{s}_e}\right) = P_r\left(Z > \frac{-x_i' \mathbf{b}}{\mathbf{s}_e}\right) \\ &= 1 - \Phi\left(\frac{-x_i' \mathbf{b}}{\mathbf{s}_e}\right) \end{aligned}$$

假設在常態分配下 , 其機率密度函數為 :

$$f(y | y > 0) = \frac{f(y)}{P_r(y > 0)} = \frac{\frac{1}{\mathbf{s}} \mathbf{f}\left(\frac{y_i - x_i' \mathbf{b}}{\mathbf{s}_e}\right)}{1 - \Phi\left(\frac{-x_i' \mathbf{b}}{\mathbf{s}_e}\right)}$$

TOBIT MODEL 通常是使用最大概似法 (maximum likelihood estimation) 來估計，其取對數過後的最大概似函數為：

$$L(\mathbf{b}, \mathbf{s}_e^2) = \sum_{y_i=0} (1 - y_i) \ln P_r(y_i = 0) + \sum_{y_i>0} y_i \ln [f(y_i | y_i > 0) P_r(y_i > 0)]$$
$$= \sum_{y_i=0} (1 - y_i) \ln(1 - \Phi(\frac{x_i' \mathbf{b}}{\mathbf{s}_e})) + \sum_{y_i>0} y_i \ln(\frac{1}{\mathbf{s}_e} \mathbf{f}(\frac{y_i - x_i' \mathbf{b}}{\mathbf{s}_e}))$$

利用以上的式子即可以求得估計值。

第二節 模型架構

本文考慮四種不同的的實證模型，分別利用 Probit 和 Tobit 模型來操作各種不同的迴歸式並對中華民國台灣地區個人所得稅申報資料進行橫斷面(cross section) 的分析，分別敘述如下：

(一)捐贈的所得彈性和價格彈性的的衡量

依據過去國外的文獻回顧，我們可以得知價格彈性和所得彈性的高低是決定政府的對於個人捐贈的所得稅抵減是否有其政策效果的一個主要的衡量指標。一直以來，國內的文獻較沒有利用個人所得稅申報資料來進行衡量，所以本文在此仍先對於國內的個人捐贈所得稅抵減的價格效果和所得效果加以衡量。設定的迴歸式如下：

$$\ln C_i = a_0 + a_1 \ln Y + a_2 \ln P + e \quad i=1\sim 5$$

C_1 ：慈善捐贈實際金額

Y：可支配所得（從事各種捐贈之前）

$$\text{當 } 175,000 < Y \quad (1)$$

$$175,000 < Y \leq 400,000 \quad (2)$$

$$400,000 < Y \leq 900,000 \quad (3)$$

$$900,000 < Y \leq 1,800,000 \quad (4)$$

$$1,800,000 < Y \leq 3,600,000 \quad (5)$$

P：捐贈價格=1-所得稅邊際稅率（捐贈前）

(二) 各種捐贈的誘因分析

根據所得稅資料中顯示，台灣的所得稅資料將捐贈分成四種不同的類型，包括慈善捐贈、政黨捐贈、候選人捐贈和私立學校捐贈等四種捐贈類型¹。但由於從事候選人捐贈及政黨捐贈的納稅義務人人數過於稀少，是以本文將候選人捐贈及政黨捐贈排除。根據這兩類型的資料，本文將試著探討納稅義務人在從事這兩種不同類型的捐贈時，影響納稅義務人從事此種捐贈的誘因為何。設定的迴歸式如下：

$$C_1 = a_0 + a_1 \ln AGE + a_2 \ln wage 1 + a_3 DEP + a_4 \ln wage + e \quad (6)$$

C_1 ：慈善捐贈實際金額

AGE：從事捐贈的納稅義務人的年齡

DEP：從事捐贈的納稅義務人的扶養子女數

Wage：納稅義務人的配偶薪資所得

Wage1：本人（納稅義務人）薪資所得

¹ 詳細的規定本文的研究範圍中已經說明。

（二） 慈善捐贈和其他扣除額的相互關係

根據我國的所得稅法規定，關於納稅義務人可以採列舉扣除額共有以下十種：

1. 捐贈
2. 人身保險費
3. 醫藥及生育費
4. 災害損失
5. 自用住宅購屋借款利息
6. 房屋租金支出
7. 依公職人員選罷法規定對候選人之捐贈
8. 依公職人員選罷法規定對政黨之捐贈
9. 依公職人員選罷法規定對候選人之捐贈
10. 依私立學校法第 51 條規定之捐贈

本文認為納稅義務人在申報稅額的時候，決定採取列舉扣除額或是標準扣除額的因素，也許並不是受到捐贈金額可以扣抵的影響。嚴格來說，本文認為如果納稅義務人可列舉的扣除額只有捐贈一項，並不足以使其採用列舉扣除額來申報稅額；納稅義務人之所以採取列舉扣除額來申報稅額應該受到其他可列舉扣除額的影響，所以本文試圖利用各列舉扣除額，探討捐贈比例和各個可列舉扣除額比例之間的相互關係，但因為受到資料限

制的影響²，本文只能採用人身保險費、醫藥及生育費這兩項列舉扣除額來和慈善捐贈的列舉扣除額來互相比較。設定的迴歸式如下：

$$C = a_0 + a_1 wage1 + a_2 wage + a_3 INS + a_4 AGE + a_5 DEP + e \quad (7)$$

$$C = a_0 + a_1 wage1 + a_2 wage + a_3 MED + a_4 AGE + a_5 DEP + e \quad (8)$$

INS：人身保險費

MED：醫藥及生育費

其餘變數的定義與上文相同。

(四) 影響個人是否捐贈的二元選擇模型 (Probit Model)

有別於上述的三種實證設定，第四個實證設定是採用二元選擇模型 (Binary Choice Model)，這個模型主要是在探討個人對於某一個選擇只有兩種不同的結果 (是或不是；有或沒有)，應用在本項迴歸設定中則是檢測納稅義務人在於捐贈與不捐贈兩種不同的結果之間的選擇，本文因為資料限制的關係，只採用慈善捐贈的樣本來作分析。設定的迴歸式如下：

如果捐贈的效用 (U_1) 大於不捐贈的效用 (U_2) 時，個人會選擇捐贈 (U_1)

² 因為其他的列舉扣除額，諸如災害損失，自用住宅購屋借款利息，政黨捐贈，私立學校捐贈等列舉扣除額的樣本數在本資料中數量過少，是以在實證模型中不得不將其排除。

在以上的行為假設下，我們設定 y_i^* 為一指標變數 (latent variable)，
則

$$y_i^* = x_i' \mathbf{b} + \mathbf{e}_i \quad (9) \sim (12)$$

在樣本中發現任何的 y_i 為兩種不同的選擇 (捐贈或不捐贈)

$$y_i = 1 \quad \text{if } y_i^* > 0 \quad (\text{捐贈})$$

$$0 \quad \text{if } y_i^* < 0 \quad (\text{不捐贈})$$

x_i' 包括：

AGE：從事捐贈的納稅義務人的年齡

DEP：從事捐贈的納稅義務人的扶養子女數

Wage：納稅義務人的配偶薪資所得

MAR：婚姻狀況 (為虛擬變數，已婚：1 未婚：0)

Wage1：納稅義務人的薪資所得

Y：納稅義務人的可支配所得

P：捐贈價格

第三節 變數說明

(一) 捐贈價格 (P)

因為在所得稅法中的規定，捐贈金額可以採用列舉扣除額的方式扣抵所得稅，因此我們可以得知納稅義務人的捐贈所負擔的比例大小隨著他們所適用的所得稅邊際稅率而遞減。根據關於捐贈方面的國外文獻，一律將納稅義務人所負擔的比例稱為捐贈的租稅價格，是以，如果納稅義務人所適用的所得稅邊際稅率為 T ，則捐贈的租稅價格即為 $1-T$ 。

根據以上的定義，我們可以得知捐贈金額的大小會對於納稅義務人所適用的邊際稅率有所影響，納稅義務人的捐贈越高，所適用的邊際稅率越低，也就是說，納稅義務人的捐贈金額和捐贈價格之間會呈現一定的相互變動關係。但是在本文所設定的迴歸式中，捐贈的租稅價格是捐贈金額的解釋變數，如此一來，這種雙向的影響關係在計量的方法上應該用聯立方程式估計，如果採用單一的方程式估計顯然會有聯立的偏誤 (simultaneity bias)。本文為了解決這個問題，採用先計算沒有捐贈時的納稅義務人所適用的邊際稅率，再用 1 減去邊際稅率來計算納稅義務人的捐贈價格。以這種替代方案，就可以免除聯立偏誤的問題，排除了捐贈金額和捐贈價格之間的關連性。

(二) 可支配所得 (Y)

關於可支配所得的定義，國外文獻在使用的定義上因為所使用的資料來源不同而有不同的定義，因為本文是採用個人的所得稅申報資料，所以為了實證操作上的方便，將可支配所得定義為綜合所得總額減去應納稅額，應納稅額即為綜合所得淨額乘上應納稅額再減去累進差額。但是根據前文的討論，捐贈的金額也會對於可支配所得產生影響，所以本文在決定應納稅額時，是指在尚未

捐贈前的應納稅額，如此一來，才可以避免產生可支配所得和捐贈金額存在内生性（endogeneity）的問題。

（三）人身保險費（INS）

根據所得稅法的規定，人身保險費的定義如下：納稅義務人本人、配偶及直系親屬之人身保險、勞工保險及軍、公、教保險之保險費。但每人每年扣除數額以不超過 24,000 元為限，申報時，應檢附保險費收據正本。

（四）醫藥及生育費（MED）

根據所得稅法的規定，醫藥及生育費的定義如下：納稅義務人及其配偶或受扶養親屬之醫藥費及生育費，以付與公立醫院、公務人員保險特約醫院，勞工保險特約醫療院所，或經財政部認定其會計記錄完備正確之醫院者為限。但受有保險給付部分，不得扣除。此項扣除額，應檢附醫療院所開立之單據。

（五）薪資所得（wage & wage1）

根據所得稅法的規定，薪資所得的定義如下：即職務上或工作上所取得的各種收入，包括薪金、俸給、工資、津貼、歲費、獎金、紅利、各種補助費和其他給與（如車馬費）