مجلة البحوث المالية والتجارية

المجلد (24) – العدد الأول – يناير 2023

المتغيرات التابعة المقيدة و نماذج انحدارها

(دراسة محاكاة)

Limited Dependent Variables and Their Regression Models (Simulation Study)

الباحث/ رامي محمد طايع محمد

مدرس مساعد الاحصاء بالمعهد العالي للعلوم الادارية و التكنولوجيا بالمنزلة

مرشح للدكتوراة

كلية التجارة . جامعة بورسعيد . قسم الاحصاء والرياضيات والتأمين

إشــراف

الأستاذ الدكتور

محمد المهدي محمد علي

أستاذ الرياضيات و الاحصاء الاكتواري

كلية التجارة . جامعة بورسعيد

الدكتورة

سمر أحمد حلمي

مدرس الاحصاء التطبيقي

كلية التجارة . جامعة بورسعيد

تاريخ الإرسال	2023-01-23
تاريخ القبول	2023-03-01
رابط المجلة:	https://jsst.journals.ekb.eg/

الملخص :

تتسم المتغيرات الاقتصادية القياسية في بعض الدراسات بسمات خاصة وتحديدا في البيانات التي تتعلق بمتغيرات لها طبيعة سلوكية حيث يأخذ المتغير قيما محددة أو مقيدة أو مراقبة في ظروف معينة , حيث يستهدف البحث التعريف بالمتغيرات التابعة المقيدة و أنواعها و الحالات التي تتواجد فيها في الدراسات الاقتصادية القياسية و خاصة النماذج السلوكية , و التي يأخذ فيها المتغير التابع عادة قيما محددة أو مبتورة عند حد معين و خاصة عند الصفر و ذلك عندما لا يقوم الشخص بنشاط معين في وقت ما و يسجل فيه قيمة صفرية و يصبح المتغير التابع في هذه الحالة يحتوي على قيما صفرية للعديد من مشاهداته , وكيفية التعامل مع هذا النوع من البيانات و مقارنة النماذج و الطرق المستخدمة في مثل هذا النوع من الدراسات و أفضلية استخدام نموذج توبيت عن باقي النماذج و الطرق الأخرى و تسليط الضوء على مميزات و عيوب الطرق المستخدمة في هذا المجال من الدراسات الاقتصادية القياسية , و كذلك النماذج المستخدمة في حالة البيانات الطولية المراقبة و المقارنة بين الطرق المستخدمة في تقدير معلمات النماذج و مناقشة كفاءة و أفضلية تلك الطرق , ثم اجراء دراسة محاكاة مونت كارلو على النماذج و الطرق المستخدمة في تلك الحالة لتحديد أفضلية كل طريقة حسب حجم العينة و عدد المشاهدات الخاصة بالسلسلة الزمنية لكل مفردة من مفردات الدراسة و ذلك تحت افتراضات معينة لمعالم العلاقة بين متغيرات الدراسة للوصول للطريقة المناسبة في تقدير معالم العلاقات الاقتصادية القياسية في مثل هذا النوع من البيانات و الدراسات .

الكلمات المفتاحية : المتغير التابع المقيد – نموذج الحل الركني – المتغير التابع المراقب – المتغير التابع المبتور – نموذج توبيت – نموذج توبيت للتأثيرات العشوائية – نموذج توبيت الديناميكي .

Abstract :

Econometric variables in some studies are characterized by special features, specifically in the data that relate to variables of a behavioral nature, where the variable takes specific, restricted or censored values in certain cases , Especially behavioral models, in which the dependent variable usually takes specific or truncated values at a certain limit, especially at zero, when a person does not perform a specific activity at some time and records a zero value, and the dependent variable in this case contains zero values for many observations, and how to deal with this type of data, and compare the models and methods used in this type of studies, and the preference of using the Tobit model over other models and other methods, and highlighting the advantages and disadvantages of the methods used in this field of econometric studies, As well as the models used in the case of censored Panel data and comparison between the methods used in estimating the parameters of the models and discussing the efficiency and preference of those methods, then conducting a Monte Carlo simulation study on the models and methods used in that case to determine the preference of each method according to the size of the sample and the number of observations of the time series for each item of the study, under certain assumptions of the parameters of the relationship between the variables of the study to reach the appropriate method in estimating the parameters of standard economic relations in such a type of data and studies.

Keywords : Limited Dependent Variable - The Corner Solution Model - Censored Dependent Variable - Truncated Dependent Variable - Tobit Model – RE Tobit Model – Dynamic Tobit Model .

مقدمــة :

هناك الكثير من المشاكل التي يمكن أن تواجه النماذج الاقتصادية القياسية ، والتي تجعل الباحث قد يلجأ إلى تطبيق طرق أخرى غير الطرق الأكثر شيوعاً والمستخدمة في تقدير معالم النماذج الاقتصادية القياسية ، تلك الطرق الأخرى تستخدم للوصول إلى تقديرات لمعالم النماذج الاقتصادية تتسم ببعض الصفات المطلوبة ، كعدم التحيز والاتساق والأفضلية عن التقديرات التي يتم التوصل إليها باستخدام الطرق الأكثر شيوعا في تلك النماذج القياسية ، لتقدير معالم نماذج الانحدار الخطية .

فعلى سبيل المثال تعتبر طريقة المربعات الصغرى العادية (OLS) هي الطريقة الأكثر استخداماً في تقدير معالم نماذج الانحدار الخطية ، لما لها من خصائص تتميز بها مقدراتها عن باقي المقدرات التي يتم الحصول عليها بالطرق الأخرى ، ولكن ظهور بعض المشاكل في النماذج الاقتصادية يجعل استخدام تلك الطريقة غير مناسب ، حيث يتم الحصول على مقدرات لمعالم النماذج الاقتصادية تفتقد للدقة والكفاءة ، نظرا لأنها متحيزة وغير متسقة .

فإذا كان تقدير معالم الانحدار الخطي باستخدام المتغيرات موضع البحث هو الهدف المنشود لكثير من البحوث الاقتصادية ، فإنه يجب تقدير هذه المعالم بالطرق التي تتلاءم مع طبيعة المتغيرات التي تدخل في تكوين العلاقات الاقتصادية ، ولأن الهدف الأساسي من تقدير هذه المعالم هو الاستفادة منها في إجراء عملية التنبؤ ، فإنه من الواجب علينا تقدير هذه المعالم ، مع الأخذ في الاعتبار خضوع المتغير التابع لقيد من القيود.

هدف البحث :

و يهدف هذا البحث إلى إلقاء الضوء على إحدى المشاكل التي تواجه النماذج الاقتصادية القياسية ، وهي مشكلة المتغير التابع المقيد ، حيث يتناول هذا الفصل التعريف بالمتغير التابع المقيد ، مع شرح بعض الحالات والأمثلة التي يكون فيها المتغير التابع مقيدا ، مما يستوجب استخدام طرق التقدير غير الخطية التي تتناسب مع خضوع المتغير لهذه القيود , و ذلك للحصول على مقدرات متسقة و غير متحيزة لمعالم علاقات الانحدار التي يحتوي فيها المتغير التابع على الكثير من القيم الصفرية و يعتبر مقيدا بطريقة ما و ذلك باستخدام طرق التقدير الملائمة .

أهميـة البحـث :

عرض المشكلات التي تواجه باحث الاقتصاد القياسي في بعض النماذج القياسية، وطرق اكتشافها ، والآثار المترتبة عليها ، وكذلك طرق علاجها يسهم في تقديم الحلول والبدائل المختلفة والتي تفيد في تصميم النماذج الاقتصادية القياسية الهامة بشكل صحيح، وكذلك الحصول على تقديرات لمعالم العلاقات الاقتصادية بشكل صحيح ، مما يسهم في دقة التحليل والتنبؤ وهما من الأغراض الأساسية لعلم الاقتصاد القياسي .

المتغير التابع المقيد Limited Dependent variable :

عند دراسة بعض الظواهر الاقتصادية ، فإننا نجد ومن وجهة نظر معينة أن كل متغير يعتبر مقيدا في المدى الذي يأخذه (مدى دراسته) ، ففي بعض الأحيان تأخذ بعض المتغيرات التابعة المستمرة نطاق محدود من القيم فقط ، حيث يتم مراقبتها أو اقتطاعها بطريقة ما ، وهو ما يقال عنه متغير تابع محدود (مقيد) ، والمتغير التابع المحدود (المقيد) له طرق خاصة للتعامل معه ، لأنه لو تم استخدام الطرق الاقتصادية القياسية الشائعة مثل طريقة المربعات الصغرى العادية (OLS) فإن النتائج تكون مضللة (متحيزة وغير متسقة) .

و تنشأ مشكلة المتغير التابع المقيد عندما يكون هناك مدى معين ومحدود لمتغير الاستجابة (التابع) والذي ينشأ عن مشاهدة جزئية أو مشاكل في الاختيار بسبب سلوك تفضيل شخصي أو عائلي يقود لمثل هذه المشاكل التي تؤدى لوجود متغير تابع مقيد ، وبالرغم من أن النموذج الخطى يعتبر ملائما في هذه الحالة ، إلا أنه من المحتمل أن يؤدى إلى تنبؤ غير دقيق لمعظم المشاهدات .

نماذج الانحدار المراقبة (المحددة): Censored Regression Models

بفرض أن y_i متغير تابع يتوزع توزيعا معتدلا بمتوسط μ وتباين σ^2 ، وباعتبار وجود عينه حجمها n وقد تم تسجيل قيم y^* التي تزيد عن قيمه ثابتة c كما هي ، أما بالنسبة للقيم y^* التي تقل عن أو تساوى القيمة c فإننا نقوم بتسجيلها بالقيمة c ، هنا يقال بأنها عينة مراقبة وتكون المشاهدات كالتالي :

$$Y_i = y_i^* \quad \text{if} \quad y_i^* > c$$
$$Y_i = c \quad \text{if} \quad y_i^* \leq c$$

والعينة المراقبة هي التي لم يستثنى فيها أي مشاهدات بطريقه متماثلة أي بقاعدة معينة ، ولكن بعض المعلومات قد تم تقييدها ، وهو نوع من العينات غالبا ما يتم في الممارسة العملية

، وهذه الحالة تعنى أن العينة لا تزال عشوائية لكل الأفراد ، لكن القيم المستخدمة لبعض المشاهدات ليست حقيقية .

و الاسم التقليدي لهذه النماذج هو نماذج الانحدار المراقبة ، وهي تُطبق بوجه عام عندما يكون المتغير المراد وصفه (التابع) مستمر بطريقة جزئية أي في مدى محدد من القيم ، ولكن له كتلة احتمال (قيمة مجموع احتمالات) عند نقطة واحدة أو أكثر .

ويعتبر نموذج الانحدار مراقباَ عندما تقطع البيانات المسجلة للمتغير التابع مدى محدد من المشاهدات المتعددة عند نقطة النهاية لهذا المدى ، وعندها فإن التغير الذي يحدث في المتغير المراقب سيحدد مدى تأثير المتغيرات المفسرة على المتغير التابع الحقيقي، لذا سيؤدى انحدار (OLS) في البيانات إلى ظهور تقديرات تتجه نحو الصفر، ويستخدم بدلا منها دالة الإمكان الأعظم للتعامل مع تلك المشكلة ، إلا أن تلك الطريقة تتطلب تحديدا صحيحا لتوزيع الخطأ والذي يمثل مشكله عند الدراسة ، فإذا كان توزيع حدود الخطأ المعطاة في المتغيرات المفسرة لها شكل بارامتري معروف ، كأن تكون موزعه طبيعيا ومتجانسة التباين ، فمن الدقة أن نصل للحد الأقصى من الدالة الاحتمالية (دالة الإمكان) ، مما يوفر مقدر طبيعي ثابت لمعامل الانحدار β ، أما في حالة عدم معرفة توزيع الأخطاء أو أن تكون عرضة لاختلاف التباين فبالتالي لن يوفر مقدر الإمكان الأعظم تقديراَ ثابتاَ للمعاملات ، وهذا ما سوف يتم تناوله لاحقا . (Kenneth, 2001)

ويمكن تصنيف تطبيقات الانحدار المراقب إلى فئتين (Wooldridge, 2002) :

1 – يوجد متغير كمي *y ونكون مهتمين بانحدار المجتمع $E(Y^*|X)$ فلو أن *Y ، X تمت مشاهدتهما لكل مفردة في المجتمع فلا يوجد شيء جديد ، ونستخدم طرق الانحدار الخطي العادية (OLS أو اللاخطية) ، لكن المشكلة تنشأ لأن *Y مراقب أعلى أو أسفل قيمة معينة ، أي أن قيمة المتغير التابع قابلة للمشاهدة لجزء من المجتمع و لكنها لا تأخذ قيمتها الحقيقية ، وكمثال على ذلك تكويد بيانات الثروة لبعض العائلات المسحوبة بطريقة عشوائية بحد معين للثروة التي تزيد عليه ، حيث تسجل قيمة *Y ثروة العائلة ، والتي تسجل بقيمتها الحقيقية عندما تكون أقل من أو تساوى 200000 جنيه ، لكن عندما تكون أكثر من 200000 جنيه فإننا نعرف فقط هذا ولكن لا نعرف قيمتها الحقيقية ، ويتم تسجيلها بالقيمة 200000 جنيه ، ويتم التعبير عن الثروة المشاهدة كما يلي :

Y = min (y* ، 200000)

ولتقدير β يجب افتراض أن الثروة *y المقدرة بواسطة المتغير المفسر x لها توزيع طبيعي متجانس التباين في شكل خطأ كما يلي :

$$Y^* = x\beta + u \qquad (u|X) \sim Normal\,(0, \sigma^2) \qquad (1)$$

2 – النوع الثاني هو النوع ذو الأثر الأكبر في الاقتصاد القياسي ، حيث المتغير التابع y يعبر عن متغير كامن يمكن مشاهدته ويصف عنصر اقتصادي معين بخصائص معينة ، حيث y يمكن أن تأخذ القيمة صفر باحتمال موجب لكنه متغير عشوائي مستمر على قيم موجبه بشكل كامل ، وكأمثلة على ذلك مبلغ تغطية التأمين على حياة الفرد ، وإنفاق الشركات على البحث والتنمية ، في هذه الأمثلة وغيرها يمكن أن نجد أن y=0 لبعض المشاهدات في المجتمع أو كثيرا منها ، وقد يسمى هذا النموذج "نموذج الحل الركني" ، ولكن اسم "نموذج الانحدار المراقب أو المحدد" هو الاسم الأكثر شيوعا لدى الباحثين ، وفى هذه الحالات نهتم بخصائص توزيع y المقدرة من x ، مثل $E(Y|X)$ و $P(Y=0|X)$ ، و إذا كنا مهتمين فقط بتأثير X_i على متوسط الاستجابة $E(Y|X)$ فلماذا لا نفترض أن : $E(Y|X) = \beta_0 + \beta X$ ونطبق المربعات الصغرى العادية (OLS) على عينه عشوائية ؟

الإجابة على هذا السؤال هي أنه عندما $y \geq 0$ فإن $E(Y|X)$ لا يمكن أن تكون خطية في x مع العلم بأنها تأخذ قيماً معينة لـ (x) ، و في هذه الحالة فإن القيم التنبؤية لـ (y) يمكن أن تكون سالبه لمجموعات كثيرة من x ، β ، وهي مشابهة لعيوب نموذج الاحتمال الخطى للاستجابات الثنائية **Linear Probability Model For Binary Responses** .

و في تلك الحالات لا نستطيع استخدام لوغاريتم المتغير التابع y كمتغير تابع Log (y) في انحدار خطى ، لأن قيمة Log(0) غير محددة ، كما أنه لا يتم استخدام المربعات الصغرى غير الخطية (NLS) والتي تضمن أن القيم التنبؤية موجبة والمؤشرات سهلة التفسير لأن لها أيضا حدود ، حيث من الممكن أن يكون التوزيع غير متجانس التباين ، وهنا طريقة (NLS) تكون غير كافية , هذا النوع من المشكلات يمكن صياغته في هذا النموذج :

$$Y_i^* = X_i\beta + U_i \qquad (U_i|X) \sim Normal\,(0, \sigma^2) \qquad (2)$$

$$Y_i = max(0, Y_i^*) \qquad (3)$$

وتمثل هذه المعادلات ما يعرف بنموذج توبيت المعياري المراقب (توبين عام 1958) ، أو نموذج توبيت من النوع الأول ، وهو الشكل القانوني من النموذج والذي عادة ما يدرس في

البحوث المنهجية ، ويمكن تطبيقه على الحزم الاقتصادية القياسية على الحاسبات الآلية التي تدعم العديد من أنواع حذف أو مراقبة البيانات .

ولكي يكون للنموذج السابق معنى ، فإن Y^* يجب أن تكون لها خصائص المتغير العشوائي الطبيعي ، حيث يجب أن تتسم بالطبيعية وتجانس التباين ، كما يجب أن يكون المتغير Y (تقريبا) مستمر عندما تكون $Y>0$ ، ولا يجب تطبيقه على المتغيرات المتقطعة أو العددية ، و خاصة التي تأخذ أرقما ضئيلة (عدد براءات الاختراع الممنوحة سنويا لشركة معينة – عدد مرات اعتقال شخص خلال سنة) ، عندئذ يناسب البيانات نموذج انحدار بواسون .

وفيما سبق يمكن بيان أن استخدام (OLS) على العينة التي بها $y_i > 0$ يعطى مقدراً متناقصاً لـ (β) ، حيث أنه عند استخدام المربعات الصغرى العادية (OLS) في العينة التي بها $y_i > 0$ تكون المعادلة كما يلي (Wooldridge 2002) :

$$Y_i = X_i\beta + \sigma\frac{\phi\left(\frac{X_i\beta}{\sigma}\right)}{\Phi\left(\frac{X_i\beta}{\sigma}\right)} + U_i \qquad (6)$$

$$E(U_i|X_i ،Y_i > 0) = 0 \qquad (7)$$

ويترتب على ذلك أن $\widehat{\beta}$ مقدر غير متسق وهي حالة مقصورة على ذلك الجزء من العينة عندما $y_i > 0$ ، أي حالة مراقبة البيانات الحقيقية .

وفي تطبيقات النماذج المراقبة فإن الكميات ذات الأهمية هي β_i وأخطاؤها المعيارية ، فيجب تقدير قيمة دالة لوغاريتم الإمكان الأعظم لأي نموذج مقدر لدورها في الحصول على إحصاء نسبة الإمكان الأعظم ، ونستطيع اختبار المتغيرات المحذوفة باستخدام اختبارات t أو اختبار نسبة الإمكان LR ، وذلك اعتمادا على افتراضات الطبيعية وتجانس التباين في المجتمع الأساسي للعينة .

نماذج الانحدار المستخدمة في المتغيرات المراقبة :

1 – نموذج الانحدار المراقب الطبيعي (Wooldridge 2003):

حيث y يتبع نموذج خطي كلاسيكي :

$$Y_i^* = \beta_0 + X_i\beta + U_i \qquad (U_i|X_i) \sim N(0،\sigma^2) \qquad (8)$$

$$Y_i = min(Y_i^*، C) \qquad (9)$$

وهو ما يعني أننا نشاهد Y_i عندما تكون Y_i^* أقل من قيمة مراقبة هي C ، كما أن U_i مستقلة عن C ، وهذه الحالة تسمى مراقبة يمنى ، كما أن مشكلة المراقبة اليسرى من الأسفل تعالج بطريقة مشابهة و كمثال على ذلك التكويد بحد معين للثروات التي تزيد عليه .

ومن الواضح في المثال السابق أن استخدام (OLS) يسبب مشكلة عند المراقبة ، حيث تستخدم تلك الطريقة السابقة البيانات غير المراقبة فقط (المتاحة في العينة) حيث $Y_i < C_i$ ، وتلك الطريقة تنتج مقدرات غير متسقة كما أن استخدام (OLS) على جميع البيانات أيضا ينتج تقديرات غير متسقة للمعلمة β لأنها تصبح بيانات غير مراقبة .

وتحت الظروف الخاصة بالنموذج نستطيع تقدير σ^2 ، β باستخدام دالة الإمكان الأعظم ، ويمكن صياغة هذا النموذج كما يلي :

$$Y_i = \begin{cases} y_i^* & IF \quad y_i^* < C \\ C & IF \quad y_i^* \geq C \end{cases} \qquad (10)$$

حيث دالة كثافة Y_i المعطاة من X_i ، C كما يلي :

$$f(Y|X,C) = 1 - \Phi\left(\frac{C-\beta X_i}{\sigma}\right) \qquad IF \qquad Y \geq C \qquad (11)$$

$$f(Y|X,C) = \frac{1}{\sigma}\,\phi\left(\frac{Y-\beta X_i}{\sigma}\right) \qquad IF \qquad Y < C \qquad (12)$$

دالة لوغاريتم الإمكان الأعظم للمشاهدات نحصل عليها بأخذ اللوغاريتم الطبيعي لدالة كثافة الاحتمال للمشاهدات ، ويمكننا تعظيم مجموع هذا الجانب فيما يتعلق بـ σ ، β للحصول على دالة الإمكان الأعظم لهما .

2 – طريقة توبيت : Tobit Model

هي الأشهر بين نماذج الانحدار المراقبة في حالة متغير استجابة الحل الركني ، أو بمعنى آخر المتغيرات التابعة المقيدة والتي تكون بها قيم صفرية لجزء غير بسيط من المجتمع ويكون هناك استمرارية تقريبا على القيم الموجبة ، فمثلا يعتبر الإنفاق على الكحول في شهر معين في مجتمع الأشخاص أعلى من 21 سنة في بلد معين متغير يأخذ مدى كبير من البيانات ولكن يوجد في المجتمع إنفاق صفري على الكحول لعدد معين من مشاهداته .

ويمكن صياغة هذا النموذج كما يلي (Davidson 1999) :

$$Y_i^* = \beta X_i + U_i \qquad U_i \sim N(0,\sigma^2) \qquad (13)$$

$$Y_i^* = \begin{cases} Y_i^* & IF \quad Y_i^* > 0 \\ 0 & IF \quad Y_i^* \leq 0 \end{cases} \qquad (14)$$

حيث المتغير Y_i^* يحقق فروض الانحدار الخطي الكلاسيكي (طبيعي ومتجانس التباين) ، وتأخذ توزيع شرطي للقيم الموجبة على نحو كامل , وهو نموذج يمكن تعديله بسهولة للسماح

بالحذف من أعلى بدلا من الحذف من أسفل أو بالحذف من كلا الطرفين أو تعديله للسماح لنقطة الحذف بالتنوع عبر المشاهدات بطريقة محددة .

وحيث أن:

$$P(Y_i = 0) = P(Y_i \leq 0) = 1 - \Phi\left(\frac{\beta X_i}{\sigma}\right) \qquad (15)$$

$$f(Y_i) = \frac{1}{\sigma}\,\phi\left(\frac{Y - \beta X_i}{\sigma}\right) \qquad (16)$$

ويكون لوغاريتم دالة الإمكان الأعظم للنموذج كما يلي :

$$log\, L = \sum_{Y_i=0} log\left[1 - \Phi\left(\frac{\beta X_i}{\sigma}\right)\right] + \sum_{Y_i>0} log\left[\frac{1}{\sigma}\,\phi\left(\frac{Y - \beta X_i}{\sigma}\right)\right] \qquad (17)$$

ويمكن تعظيم لوغاريتم الإمكان الأعظم لـ σ ، β للحصول على مقدرات الإمكان لهما ، ويمكن اختبارهما باستخدام اختبارات (t) ، أو اختبار (Wald) ، أو اختبار(نسبة الإمكان LR) ، وذلك بافتراض أن النموذج محدد بدقة وهوما يعطي مقدرات متسقة وذات كفاءة مقاربة لكل من σ ، β تحت شروط قياسية معتدلة .

ويصبح (Wooldridge 2002):

$$E(Y|Y > 0) = \beta X_i + \sigma\frac{\phi\left(\frac{X_i\beta}{\sigma}\right)}{\Phi\left(\frac{X_i\beta}{\sigma}\right)} \qquad (18)$$

$$E(Y|X) = \Phi\left(\frac{X_i\beta}{\sigma}\right)\left[\beta X_i + \sigma\frac{\phi\left(\frac{X_i\beta}{\sigma}\right)}{\Phi\left(\frac{X_i\beta}{\sigma}\right)}\right] \qquad (19)$$

$$= \Phi\left(\frac{X_i\beta}{\sigma}\right)\beta X_i + \sigma\phi\left(\frac{X_i\beta}{\sigma}\right)$$

لماذا لا يتم استخدام المربعات الصغرى العادية (OLS) ؟

افترض وجود نموذج الانحدار التالي :

$$Y_i^* = X_i\beta + U_i \qquad (29)$$

$$Y_i = \begin{cases} Y_i^* & if \quad Y_i^* > 0 \\ o & if \quad Y_i^* \leq 0 \end{cases} \qquad (30)$$

حيث Y^* متغير كامن (غير مشاهد) ، وفي الواقع نحن نشاهد Y_i التي تختلف عن Y_i^* لأنها مراقبة ، وحيث أن المراقبة تحدث عندما $Y_i^* \leq 0$ فكلما كان حد الخطأ أكبر كلما كانت Y_i^* أكبر وبالتالي يزداد الاحتمال بأن $Y_i^* > 0$ وهو احتمال يعتمد على X_i ، وهكذا في العينة المشاهدة بالفعل فإن قيمة U_i بالتالي لن يعد لها متوسط شرطي يساوي الصفر وسوف تكون مرتبطة بالمتغير المفسر X_i .

وبالتالي لم تعد U_i تفي بالافتراضات الرئيسية ، وهنا نجد أن تقدير المربعات الصغرى العادية (OLS) باستخدام العينات المراقبة يولد مقدرات متحيزة وغير متسقة وغير ثابتة .

المتغير المراقب في البيانات الطولية :

البيانات الطولية هي مزيج بين البيانات المقطعية و السلاسل الزمنية حيث يتم مشاهدة مفردات الدراسة في فترة زمنية لكل مفردة و يتم تجميعها طوليا لتكون سلاسل زمنية لعينة مقطعية من مفردات ظاهرة معينة , و عندما يحتوي المتغير التابع في البيانات الطولية على العديد من البيانات الصفرية نتيجة عدم القيام بنشاط معين في فترة ما , فإن طريقة OLS و تجميع البيانات في تجميعة واحدة تعتبر غير مناسبة لأنها تتجاهل الاختلاف غير المشاهد بين مفردات الدراسة وعدم تجانس التباين و تؤدي للحصول على مقدرات متحيزة و غير متسقة (Kenneth , 2001) .

يستخدم عادة في هذه الحالة نماذج انحدار تناسب طبيعة البيانات الطولية كنموذج توبيت للتأثيرات العشوائية , و لكن عادة في المتغيرات الاقتصادية القياسية تظهر مشكلة الانحدار الذاتي على القيم المتأخرة للمتغير التابع و كذلك الارتباط التسلسلي لحد الخطأ , و هو ما يلزم معه استخدام نماذج ديناميكية مثل نموذج توبيت الديناميكي للتأثيرات العشوائية لمراعاة الافتراضات السابقة في بيانات الدراسة .

1-نموذج توبيت للتأثيرات العشوائية : (Random Effects Tobit Model)

$$y_{it}^{*} = \beta' x_{it} + u_{it} \qquad i = 1,2,\ldots,N \qquad t = 1,2,\ldots T$$

$$u_{it} = v_i + \varepsilon_{it} \qquad v_i \sim N(0, \sigma_v^2) \qquad \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$$

عندما :

$$y_{it} = \begin{cases} y_{it}^{*} \ldots\ldots if \ldots\ldots y_{it}^{*} > 0 \\ 0 \ldots\ldots\ldots otherwise \end{cases}$$

حيث :

v_i : تأثيرات فردية غير مشاهدة .

في حالة :

افتراض الثبات لكل مفردة تسمى v_i مقدرات التأثيرات الثابتة .

افتراض أنها مأخوذة من توزيع عشوائي فإن v_i تسمى مقدرات التأثيرات العشوائية .

ε_{it} : تأثيرات فردية عشوائية غير مشاهدة .

u_{it} : يسمى نموذج مكونات الخطأ و يرتبط عبر الزمن و يقسم الخطأ إلى :

- **تأثير عشوائي فردي ثابت زمنيا (Random Effects) RE و هو v_i**
- **خطأ عشوائي متقلب و متباين زمنيا و هو ε_{it}**

و حيث :

$$E(u_{it}) = 0 \quad , \quad Var(u_{it}) = \sigma_v^2 + \sigma_\varepsilon^2 \quad ,$$

$$corr(u_{it}, u_{is}) = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\varepsilon^2} \quad , \quad E(v_i x_i) = 0$$

<u>فروض التوزيع :</u>

1- لا يوجد ارتباط تسلسلي في حد الخطأ .

2- التأثيرات الفردية غير مرتبطة ذاتيا .

3- $v_i / x_i \sim N(0, \sigma_v^2)$.

و يلاحظ على هذا النموذج ما يلي : (Bruno ,2004)

- في نموذج توبيت غير الخطي عند استخدام التأثيرات الثابتة سوف تزداد قيمة α_i بزيادة N و تظهر حالة المعلمات العرضية و يحدث التحيز في النتائج .

- التقريب الأفضل يحدث باستخدام تقدير التأثيرات العشوائية لنموذج توبيت للبيانات الطولية .

في حالة وجود ارتباط ذاتي في النموذج فمن الضروري التعامل معه وعدم تجاهله ، كما أن الانحدار الذاتي علي القيم (Autoregressive Process) قد يشير إلى الارتباط الذاتي , وهنا تظهر الحاجة إلى استخدام تحليل نماذج بيانات طولية ديناميكية حتى لا نواجه مشكلة التحيز في العينات الصغيرة ، حيث يمكن استخدام اسلوب العزوم المعممة (GMM) مع

الأدوات التي توفرها المتغيرات المبطأة (Lagged Variables) والعزوم للحصول علي تقديرات ذات كفاءة عالية .

2- نموذج توبيت الديناميكي للتأثيرات العشوائية (Dynamic Tobit Model For (Panel Data (Baltagi , 2005)

$$y^{*}_{it} = x_{it}'\beta + \lambda y_{i,t-1} + u_{it}$$

$$y_{it} = \begin{cases} y^{*}_{it} \ldots\ldots if \ldots\ldots y^{*}_{it} > 0 \\ 0 \ldots\ldots\ldots otherwise \end{cases}$$

for i= 1,2,….N , t=1,2,…….T

$$u_{it} = v_i + \varepsilon_{it} \qquad v_i \sim N(0, \sigma_v^2) \qquad \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$$

$$\varepsilon_{it} = \rho \varepsilon_{it-1} + c_{it} \qquad c_{it} \sim N(0, \sigma_c^2) \text{ orthogonal to } v_i$$

و يلاحظ ما يلي :

- يستخدم هذا النموذج في حالة الارتباط الذاتي في النموذج بين البواقي ويستخدم في حالة وجود متغيرات تابعة مبطأه أو انحدار ذاتي .
- حتي لو لم يكن هناك ارتباط ذاتي لحد الخطأ فإن المشكلة الناتجة من ارتباط المتغير المبطأ الداخلي و حد الخطأ قد تصيب التحليل ويحدث التحيز خاصة في العينات الصغيرة .
- يفترض الاستقلال الشرطي بين التأثيرات الفردية المحددة غير المشاهدة والمتغيرات التفسيرية , و قد يستخدم في هذه الحالة متغيرات مقيدة اكثر للنموذج للتغلب علي تلك المشكلة , أو تستخدم طريقة العزوم المعممة (GMM) للتقدير والحصول علي مقدرات أكثر كفاءه .
- هو نموذج يعالج حالة التبعية في النتائج (Stat – Dependence) وكذلك يعالج حالة عدم التجانس الفردي غير الملاحظ والارتباط التسلسلي للأخطاء , ويستخدم دالة الامكان وذلك مع استخدام التأثيرات العشوائية , لأنه من غير المنطقي افتراض ثبات حد الخطأ ، كما أنه يأخذ في الاعتبار تأثير الزمن علي كفاءة الدالة .
- في نموذج توبيت للتأثيرات العشوائية و كلك في نموذج توبيت الديناميكي للتأثيرات العشوائية تستخدم طريق دالة الامكان ML لتقدير معالم كلا النموذجان .

دراسة المحاكاة :

تشتمل الدراسة العملية علي تحليل و عرض نتائج المحاكاة لدراسة سلوك العوامل المختلفة للتقدير مثل عدد القطاعات N و عدد الفترات الزمنية لكل قطاع T و معامل الانحدار β للمتغير المستقل x و معامل الانحدار الذاتي λ للمتغير التابع y هذا بالإضافة إلي معلمات حد الخطأ العشوائي في النموذج (σ_ϵ,ρ,σ_v,σ_c).

و نهدف من دراسة المحاكاة تحديد العوامل و سلوك المعالم المقدرة لنماذج توبيت المفترضة و المتغيرات المؤثرة فيها و ذك في حالات العينات الصغيرة و المتوسطة و الكبيرة و كذلك في حالة اختلاف طول السلسلة الزمنية بالنسبة للمفردات . وسوف ينقسم المبحث إلي أربعة إجزاء رئيسة و يشتمل كل جزء على نتائج المحاكاة لنماذج توبيت المفترضة في هذه الدراسة، وتتم كالتالي:

الجزء الأول: دراسة المحاكاة لنموذج توبيت العادي

الجزء الثاني: دراسة المحاكاة لنموذج توبيت الديناميكي

المعلمات المفترضة لأجراء عملية المحاكاة :

يتم عمل المحاكاة باستخدام برنامج R(10) بعدد مرات 1000 (N.sim) مره , لعدد قطاعات N=(10,50,100) , و طول السلسلة الزمنية في كل قطاع T=(20,30) , بإجمالي عدد 6 محاكاة , كل محاكاة تتم 1000 مرة , و ذلك بافتراض المعلمات كما يلي :

معامل المتغير المستقل β يفترض أنه يساوي 1 .

معامل الانحدار الذاتي λ يفترض انه يساوي 0.5 .

بالنسبة لمعلمات أو مركبات حد الخطأ العشوائي (σ_ϵ,ρ,σ_v,σ_c) تفترض كالأتي:

σ_ϵ=1 & ρ=0.5 & σ_v=1 & σ_c=1

بعد الحصول علي نتائج المحاكاة نقوم بحساب المؤشرات التالية:

MB(β^),RMSE(β^),MedB(β^),MAD(β^)

MB(λ^),RMSE(λ^),MedB(λ^),MAD(λ^)

نتائج المحاكاة :

بالنسبة لمؤشرات الخطأ للمعلمات المقدرة لكل من النموذجين كما هو موضح بالجداول التالية :

$\beta = 1$		Tobit Model (الطريقة الأولى)			
N	T	MB	RMSE	MedB	MAD
10	20	0.76448	0.76782	0.76645	0.58745
	30	0.76448	0.76782	0.74993	0.56239
50	20	0.75743	0.76639	0.75242	0.56613
	30	0.75930	0.76605	0.75660	0.57244
100	20	0.76587	0.77029	0.76300	0.58217
	30	0.75802	0.76105	0.75918	0.57635

$\beta = 1$		Dynamic Tobit Model (الطريقة الثانية)			
N	T	MB	RMSE	MedB	MAD
10	20	0.74792	0.75617	0.77672	0.60329
	30	0.70917	0.76600	0.75919	0.57637
50	20	0.74914	0.76287	0.75948	0.57681
	30	0.74906	0.75913	0.76142	0.57976
100	20	0.75294	0.75887	0.75723	0.57340
	30	0.76727	0.77174	0.76954	0.59219
λ=0.5					
10	20	0.51715	0.53416	0.48387	0.23413
	30	0.52903	0.66961	0.47232	0.22309
50	20	0.49054	0.52217	0.48299	0.23328
	30	0.48636	0.51709	0.45935	0.21100
100	20	0.49678	0.51367	0.49011	0.24021
	30	0.46865	0.48283	0.46068	0.21223

الاستنتاجات :

1- يتم استخدام الصيغة الخطية في النماذج المقيدة نظرا لسهولتها بدلا من الصيغة اللوغاريتمية علاوة على عدم إمكانية استخدام طريقة توبيت لتقدير المعالم في الصيغة اللوغاريتمية نظرا لوجود قيم صفرية في مشاهدات المتغير التابع .

2- المقدرات المتحصل عليها باستخدام طريقة "توبيت" تعتبر متسقةً و تتوزع توزيعاً معتدلاً تقريباً في حين تعتبر المقدرات المتحصل عليها باستخدام طريقة المربعات الصغرى العادية OLS متحيزة في اتجاه الصفر و يزداد هذا التحيز بزيادة σ_u^2 .

3- يلاحظ أن معامل التحديد في بعض التطبيقات صغير و أقل من 50% ، فإذا كان التنبؤ ليس هو الهدف من النموذج ، فإنه يمكن التغاضي عن صغر معامل التحديد و الذي يلزم أن يكون أكبر من ذلك في حالة استخدام النموذج في التنبؤ .

4- قيمة دالة الإمكان Log-Likelihood المتحصل عليها من استخدام طريقة توبيت تزيد عن تلك المتحصل عليها من استخدام طريقة المربعات الصغرى العادية OLS في حالة المتغير التابع المقيد .

5- في البيانات الطولية للظواهر الاقتصادية القياسية عندما يحتوي المتغير التابع على بيانات صفرية و تحت فروض معينة , أظهرت نتائج المحاكاة أفضلية استخدام نموذج ديناميكي لتوبيت للتأثيرات العشوائية عن استخدام نموذج توبيت العادي للتأثيرات العشوائية و ذلك من خلال جميع مؤشرات أخطاء التقدير المستخدمة في مقارنة معلمات النموذج المقدرة , حيث كانت مؤشرات الخطأ أقل دائما في حالة استخدام نموذج توبيت الديناميكي .

قائمة المراجع

المراجع الأجنبية :

Baltagi , Badi H. , "Econometric Analysis Of Panel Data" ,John Wiley & Sons, LTD , Third Edition ,England ,2005 .

Bruno , Giuseppe ,"Limited Dependent Panel Data Models :A Comparative Analysis Of Classical and Bayesian Inference Among Econometric Packages", Bank Of Italy Research Department , Italy ,2004 .

Chay, Kenneth Y. and Powell, James L., "Semiparametric Censored Regression Models", Journal of Economic Perspectives, Vol.15, No. 4, 2001.

Davidson, R., and Mackinnon, James G., "Econometric Theory and Methods", 1999.

Tobin, J., "Estimation of Relationships for Limited Dependent Variables", Econometrica, 1958.

Wooldridge, Jeffrey M., "Econometric Analysis of Cross Section and Panel Data", The MIT Press, Cambridge, London, 2002.

Wooldridge, Jeffrey M., "Introductory Econometrics", Michigan State University, 2003.