المجلة العراقية للعلوم الإحصائية (21) 2012

ص ص [266-285]

اختيار النموذج في نماذج البيانات الطولية الثابتة والعشوائية

زكريا يحيى الجمال[1]

المستخلص

تم استخدام نماذج البيانات الطولية وكيفية الاختيار في توفيق نماذج البيانات الطولية الثلاثة، نموذج الانحدار التجميعي ونموذج التأثيرات الثابتة ونموذج التاثيرات العـشوائية. تم الاعتماد على معيار معامل التحديد المصحح ومعيار اكاكي للمعلومـات فـي اختيـار افضل نموذج جزئي من النموذج العام من خلال دراسة اربعة متغيرات توضيحية تـؤثر في الموازنة العامة لأية دولة، حيث شملت الدراسة ثماني عشرة دولـة عربيـة وللفتـرة الزمنية من 1995 الى 2009.

الكلمات الدالة: نموذج الانحدار التجميعي، ونموذج التاثيرات الثابتة، ونموذج التاثيرات العشوائية، ومعيار اكاكي للمعلومات، ومعامل التحديد المصحح

Selecting Model in Fixed and Random Panel Data Models

Zakariya Y. Algamal

Abstract

In this paper we used panel data models and how to choose the fitting three panel data models, the pooled regression model, fixed effects model, and random effects model, depending on the adjusted coefficient of determination and Akaike information criterion in

[1] مدرس/ قسم الإحصاء والمعلوماتية / كلية علوم الحاسوب والرياضيات/ جامعة الموصل.

تاريخ القبول 2012/2/7ـــــــــــــــــتاريخ التسليم 2012/3/20

selecting the best sub-model form the general model through studying the four affected variables on the public budget. This study included of eighteen Arab countries for the period from 1995 to 2009.

1- المقدمة

تعتمد العديد من الدراسات والبحوث على أساليب متطورة من أجل الحصول على نتائج تتصف بالفعالية والدقة العاليتين، وقد كان لعلم الإحصاء وفروعه المرتبطة به الأثر الكبير في بناء النماذج الرصينة وتحليل البيانات من خلالها وصولاً الى قرارات السليمة. يهتم تحليل الانحدار ببناء العلاقة الرياضية بين متغير الاستجابة والمتغيرات التوضيحية وتتمثل هذه العلاقة كتركيبة خطية تدعى معادلة الانحدار. لقد اكتسبت نماذج البيانات الطولية في العقد الحالي اهتماما بالغا وخصوصا في الدراسات الاقتصادية والطبية. تعرف هذه البيانات على انها بيانات مقطعية مقاسة في فترات زمنية معينة.

ان الفائدة الرئيسية من استخدام البيانات الطولية هي زيادة الدقة في التنبؤ من خلال زيادة عدد المشاهدات عن طريق ربط عدد المشاهدات المقطعية بعدد الفترات الزمنية. ذهب العديد من الباحثين في دراسة نماذج البيانات الطولية فمنهم من اهتم بدراسة خصائص نماذج البيانات الطولية رياضيا (Bramati & Croux, 2007) ، (Dustmann & Engarcia, 2007) ، (Sun, 2010) ، (Lee & Yu, 2010) ، (Baltagi et al, 2010) ومنهم من اهتم بتطبيق هذه النماذج في دراساتهم (El-Gamal & Inanoglu, 2005)، **(Chuang & Wang,2099)**، (Mikhed & Zemcik, 2009) ، (Lukas & Jan, 2011) ، (Kai & Qin, 2011).

لقد سعى هذا البحث الى تقديم مفهوم نموذج البيانات الطولية وكيفية الاختيار بين نماذج هذه البيانات ثم اختيار أفضل نموذج جزئي.

2- نماذج البيانات الطولية

لقد اكتسبت هذه النماذج التي تدعى ايضا بنماذج البانل (Panel Model) في العقد الحالي اهتماما بالغا وخصوصا في الدراسات الاقتصادية والطبية لانها تاخذ في الاعتبـار اثر التغير في الزمن وكذلك اثر التغير في المشاهدات المقطعية. تعرف البيانات الطوليـة على انها مشاهدات مقطعية مقاسة في فترات زمنية معينـة. حيـث تـشمل المـشاهدات المقطعية الدول، المحافظـات، المؤسـسات،..... الـخ (Gujarati, 2003). ولغـرض توضيح مفهوم البيانات الطولية سوف نفرض بان لدينا دراسة حول مقـدار الاسـتهلاك كمتغير اسـتجابة **(y)** ومقـدار الـدخل كمتغيـر توضـيحي **(x)** ولخمـس عوائـل (A, B, C, D, and E) ولثلاث سنوات (2009، 2010، و 2011)، يوضح الجـدول (1) كيفية عرض البيانات الطولية في مثالنا هذا.

الجدول(1): اسلوب عرض البيانات الطولية

العوائل	السنوات	مقدار الاستهلاك (y)	مقدار الدخل (x)
A	2009	$y_{A,2009}$	$x_{A,2009}$
	2010	$y_{A,2010}$	$x_{A,2010}$
	2011	$y_{A,2011}$	$x_{A,2011}$
B	2009	$y_{B,2009}$	$x_{B,2009}$
	2010	$y_{B,2010}$	$x_{B,2010}$
	2011	$y_{B,2011}$	$x_{B,2011}$
:	:	:	:

:	:	:	:
$x_{E,2009}$	$y_{E,2009}$	2009	E
$x_{E,2010}$	$y_{E,2010}$	2010	
$x_{E,2011}$	$y_{E,2011}$	2011	

من الجدول (1) نلاحظ بانه تم دمج البيانات المقطعية مع الفترات الزمنية. تتمتع نمـاذج البيانات الطولية في استخدامها مقارنة عند استخدام نماذج البيانات المقطعية بمفردهـا او نماذج بيانات السلسلة الزمنية بمفردها بالعديد من المزايا التي وضحها الباحـث Blatagi (Blatagi, 2005) منها:

1- التحكم في عدم تجانس التباين الخاص الذي قد يظهر في حالة البيانات المقطعيـة او حالة البيانات الزمنية.

2- تعطي البيانات الطولية كفاءة افضل وزيادة في درجات الحرية وكذلك اقل تعددية خطية بين المتغيرات، ومحتوى معلوماتي اكثر اذا مـا تـم اسـتخدام البيانـات المقطعية أو الزمنية.

عندما تكون المشاهدات المقطعية مقاسة لنفس الفترات الزمنية عندئذ يطلق على البيانات الطولية بانها بيانات طولية متزنة (Balanced Panel Data) كما في مثالنا السابق، ولو فرضنا بان واحدة من هذه العوائل قد تم تسجيل مشاهداتها المقطعية لسنتين فقط وباقي العوائل لثلاث سنوات عندئذ يطلق على هذه البيانات بالبيانات الطولية غير المتزنة (Unbalanced Panel Data). من خلال مثالنا السابق يمكننا استخدام اسلوب الانحدار في السلاسل الزمنية ولخمس مرات (لكل عائلة نموذج) او يمكننا استخدام الانحدار الاعتيادي عندما تكون بياناتنا مقطعية ولثلاث مرات (لكل سنة نموذج).

من هنا تاتي نماذج البيانات الطولية في ثلاثة اشكال رئيسية هي: نموذج الانحدار التجميعي (Pooled Regression Model(PM)) ، نموذج التأثيرات الثابتة (Fixed Effects Model(FEM)) ونموذج الثاثيرات العشوائية (Random Effects Model(REM)) . ليكن لدينا N من المشاهدات المقطعية مقاسة في T من الفترات الزمنية فان نموذج البيانات الطولية يعرف بالصيغة الاتية:

$$y_{it} = \beta_{\circ(i)} + \sum_{j=1}^{k} \beta_j x_{j(it)} + \varepsilon_{it} \qquad ,i = 1,2,.....,N \quad t = 1,2,....,T \quad \ldots\ldots(1)$$

حيث ان y_{it} تمثل قيمة متغير الاستجابة في المشاهدة i عند الفترة الزمنية t، $\beta_{\circ(i)}$ تمثل قيمة نقطة التقاطع في المشاهدة i، β_j تمثل قيمة ميل خط الانحدار، $x_{j(it)}$ تمثل قيمة المتغير التفسيري j في المشاهدة i عند الفترة الزمنية t، وان ε_{it} تمثل قيمة الخطأ في المشاهدة i عند الفترة الزمنية t . تجدر الاشارة هنا الى ان المقصود بالمشاهدة i هو عدد العوائل في مثالنا السابق. تعتمد تقدير المعلمات للنموذج في المعادلة (1) على نوع نموذج البيانات الطولية المستخدم.

2-1 نموذج الانحدار التجميعي

يعتبر هذا النموذج من ابسط نماذج البيانات الطولية حيث تكون فيه جميع المعاملات $\beta_{\circ(i)}$ و β_j ثابتة لجميع الفترات الزمنية (يهمل اي تاثير للزمن). باعادة كتابة النموذج في المعادلة (1) نحصل على نموذج الانحدار التجميعي وبالصيغة الاتية:

$$y_{it} = \beta_{\circ} + \sum_{j=1}^{k} \beta_j x_{j(it)} + \varepsilon_{it} \qquad ,i = 1,2,.....,N \quad t = 1,2,....,T \quad \ldots\ldots(2)$$

حيث ان $E(\varepsilon_{it}) = 0$ و $var(\varepsilon_{it}) = \sigma_{\varepsilon}^{2}$. تستخدم طريقة المربعات الصغرى الاعتيادية في تقدير معلمات النموذج في المعادلة (2) (Greene, 2012) بعد ان ترتب القيم الخاصة بمتغير الاستجابة والمتغير التوضيحي بدءا من اول مجموعة بيانات مقطعية وهكذا وبحجم مشاهدات مقداره $(N * T)$.

2-2 نموذج التاثيرات الثابتة

في نموذج التاثيرات الثابتة يكون الهدف هو معرفة سلوك كل مجموعة بيانات مقطعية على حدة من خلال جعل معلمة القطع $\beta_{\circ}$ تتفاوت من مجموعة الى اخرى مع بقاء معاملات الميل β_j ثابتة لكل مجموعة بيانات مقطعية (اي سوف نتعامل مع حالة عدم التجانس في التباين بين المجاميع)، وعليه فان نموذج التاثيرات الثابتة يكون بالصيغة الاتية:

$$y_{it} = \beta_{\circ(i)} + \sum_{j=1}^{k} \beta_j x_{j(it)} + \varepsilon_{it} \qquad , i = 1,2,....., N \quad t = 1,2,....., T \quad .....(3)$$

حيث ان $E(\varepsilon_{it}) = 0$ و $var(\varepsilon_{it}) = \sigma_{\varepsilon}^{2}$ (Greene, 2012). يقصد بمصطلح التاثيرات الثابتة بان المعلمة $\beta_{\circ}$ لكل مجموعة بيانات مقطعية لاتتغير خلال الزمن (time invariant) وانما يكون التغير فقط في مجاميع البيانات المقطعية (Gujarati, 2003). لغرض تقدير معلمات النموذج في المعادلة (3) والسماح لمعلمة القطع $\beta_{\circ}$ بالتغير بين المجاميع المقطعية عادة ما تستخدم متغيرات وهمية بقدر $(N-1)$ لكي نتجنب حالة التعددية الخطية التامة (Greene, 2012) ثم تستخدم طريقة المربعات الصغرى الاعتيادية. يطلق على نموذج التاثيرات الثابتة اسم نموذج المربعات الصغرى للمتغيرات الوهمية (Least Squares Dummy Variable Model). بعد اضافة المتغيرات الوهمية D في المعادلة (3) يصبح النموذج بالشكل الاتي:

$$y_{it} = \alpha_1 + \sum_{d=2}^{N} \alpha_d D_d + \sum_{j=1}^{k} \beta_j x_{j(it)} + \varepsilon_{it} \qquad ,i = 1,2,\ldots\ldots, N \quad t = 1,2,\ldots\ldots, T \quad \cdots(4)$$

حيث يمثل المقدار $\alpha_1 + \sum_{d=2}^{N} \alpha_d D_d$ التغير في المجاميع المقطعية لمعلمة القطع $\beta_\circ$.

ويمكن كتابة النموذج بالمعادلة (4) بعد حذف α_1 بالشكل الاتي (Gujarati, 2003) ، (Greene, 2012):

$$y_{it} = \sum_{d=1}^{N} \alpha_d D_d + \sum_{j=1}^{k} \beta_j x_{j(it)} + \varepsilon_{it} \qquad ,i = 1,2,\ldots\ldots, N \quad t = 1,2,\ldots\ldots, T \quad \cdots\cdots(5)$$

2-3 نموذج التاثيرات العشوائية

في نموذج التاثيرات الثابتة يكون حد الخطأ ε_{it} ذا توزيع طبيعي بوسط مقدراه صفر وتباين مساوي الى σ_ε^2، ولكي تكون معلمات نموذج التاثيرات الثابتة صحيحة وغير متحيزة عادة ما يفرض بان تباين الخطأ ثابت (متجانس) لجميع المشاهدات المقطعية وليس هناك اي ارتباط ذاتي خلال الزمن بين كل مجموعة من مجاميع المشاهدات المقطعية في فترة زمنية محددة. يعتبر نموذج التاثيرات العشوائية نموذجا ملائما في حالة وجود خلل في احد الفروض المذكورة اعلاه في نموذج التاثيرات الثابتة (Gujarati, 2003). في نموذج التاثيرات العشوائية سوف يعامل معامل القطع $\beta_{\circ(i)}$ كمتغير عشوائي له معدل مقداره μ اي:

$$\beta_{\circ(i)} = \mu + v_i \qquad ,i = 1,2,\ldots\ldots,N \qquad \ldots\ldots(6)$$

وبتعويض المعادلة (6) في المعادلة (3) نحصل على نموذج التاثيرات العشوائية وبالشكل الاتي:

$$y_{it} = \mu + \sum_{j=1}^{k} \beta_j x_{j(it)} + v_i + \varepsilon_{it} \qquad ,i = 1,2,.....,N \;\; t = 1,2,....,T \;\; .....(7)$$

حيث ان v_i يمثل حد الخطأ في مجموعة البيانات المقطعية **i**. يطلق على نموذج التاثيرات العشوائية احيانا نموذج مكونات الخطأ (Error Components Model) بسبب ان النموذج في المعادلة (7) يحوي مركبين للخطأ هما v_i و ε_{it}. يمتلك نموذج التاثيرات العشوائية خواص رياضية منها ان $E(\varepsilon_{it}) = 0$ ، $var(\varepsilon_{it}) = \sigma_\varepsilon^2$ ، $E(v_i) = 0$ ، $var(v_i) = \sigma_v^2$.

ليكن لدينا حد الخطأ المركب الأتي:

$$w_{it} = v_i + \varepsilon_{it} \qquad .....(8)$$

حيث ان:

$$E(w_{it}) = 0 \quad .....(9)$$

$$var(w_{it}) = \sigma_v^2 + \sigma_\varepsilon^2 \qquad .....(10)$$

تفشل طريقة المربعات الصغرى الاعتيادية في تقدير معلمات نموذج التاثيرات العشوائية كونها تعطي مقدرات غير كفوءة ولها اخطاء قياسية غير صحيحة مما يؤثر في اختبار المعلمات كون ان التباين المشترك بين w_{it} و w_{is} لايساوي الصفر اي:

$$cov(w_{it}, w_{is}) = \sigma^2_v \neq 0 \quad , t \neq s \qquad .....(11)$$

لغرض تقدير معلمات نموذج التاثيرات العشوائية بشكل صحيح عادة ما تستخدم طريقة المربعات الصغرى المعمة (Generalized Least Squares (GLS)) (Greene, 2012) .

3-1 أساليب اختيار النموذج الملائم للبيانات الطولية

كما ذكرنا سابقا بوجود ثلاثة نماذج رئيسية من نماذج البيانات الطولية وعلى هذا الاساس يطرح السؤال الأتي: ما هو النموذج الأكثر ملاءمة لبيانات دراسة ما؟ لغرض الاجابة عن مثل هكذا تساؤل سوف نقوم بعرض اسلوبين، الاول: اسلوب الاختيار بين نموذج الانحدار التجميعي ونموذج التاثيرات الثابتة والثاني: هو اسلوب الاختيار بين نموذج التاثيرات الثابتة ونموذج التاثيرات العشوائية.

لغرض الاختيار بين نموذج الانحدار التجميعي ونموذج التاثيرات الثابتة نستخدم اختبار **F** المقيد وبالصيغة الاتية:

$$\mathbf{F(N-1,NT-N-k)=\frac{(R_{FEM}{}^2-R_{PM}{}^2)/(N-1)}{(1-R_{FEM}{}^2)/(NT-N-k)}} \qquad .....(12)$$

حيث ان $\mathbf{k}$ هي عدد المعلمات المقدرة وان $\mathbf{R_{FEM}}$ يمثل معامل التحديد عند استخدام نموذج التاثيرات الثابتة و $\mathbf{R_{PM}}$ يمثل معامل التحديد عند استخدام نموذج الانحدار التجميعي. تقارن نتيجة المعادلة (12) مع $\mathbf{F(\alpha,N-1,Nt-N-k)}$ فإذا كانت قيمة المعادلة (12) اكبر او مساوية الى القيمة الجدولية (او اذا كانت قيمة $\mathbf{p-value}$ اقل من او تساوي **0.05**) عندئذ فان نموذج التاثيرات الثابتة هو النموذج الملائم لبيانات الدراسة (Greene, 2012). بعد اختيار نموذج

التاثيرات الثابتة بوصفه نموذجاً ملائماً نقوم بالاختيار بينـه وبـين نمـوذج التـاثيرات العشوائية لتحديد النموذج النهـائي الملائـم لبيانـات الدراسـة مـن خـلال اسـتخدام اختبار(Hausman(H حيث تكون فرضية العدم بالشكل الاتي:

$\mathbf{H_0}$: نموذج التاثيرات العشوائية هو النموذج الملائم

$\mathbf{H_1}$: نموذج التاثيرات الثابتة هو النموذج الملائم

وتكون صيغة الاختبار بالصيغة الاتية:

$$\mathbf{H} = (\hat{\beta}_{FEM} - \hat{\beta}_{REM})'[\mathbf{var}(\hat{\beta}_{FEM}) - \mathbf{var}(\hat{\beta}_{REM})]^{-1}(\hat{\beta}_{FEM} - \hat{\beta}_{REM}) \quad \ldots\ldots.(13)$$

حيث ان $\mathbf{var}(\hat{\beta}_{FEM})$ هو متجه التباين لمعلمات نموذج التاثيرات الثابتة و $\mathbf{var}(\hat{\beta}_{REM})$ هو متجه التباين لمعلمات نموذج التاثيرات العشوائية. حيث ان هذه الاحصائية لها توزيع مربع كاي وبدرجة حرية مقدارها $\mathbf{k}$. يكون نموذج التاثيرات الثابتة هو النموذج الملائـم اذا كانت قيمة الاحصائية اكبر من قيمة مربع كاي وعلى العكس سوف يكـون النمـوذج الملائم لبيانات الدراسة هو نموذج التاثيرات العشوائية (Hausman, 1978).

4-معايير اختيار افضل نموذج جزئي

في كثير من التطبيقات الاحصائية يرغب صانع القرار في الوصول إلـى أفـضل نموذج يحكم البيانات قيد الدراسة عن طريق ايجاد افضل المتغيرات التي تحـدد افـضل نموذج يمكن بناء القرار الامثل عليه. أن اسلوب اختيار افضل نموذج جزئي بعد توفيـق نموذج البيانات الطولية يحتاج الى انتباه شديد حيث سبق وان تطرقنا في المبحث الـسابق إلى كيفية اختيار النموذج الملائم للبيانات الطولية من بين نماذج البيانات الطولية الثلاثـة من حيث اجراء هذه الاختبارات للتاكد من ان اختيارنا للنموذج الملائم يبقى كما هو اثناء اختيار افضل نموذج جزئي. في بحثنا هذا سوف نتطرق الى اسـتخدام معيـارين همـا:

معامل التحديـد المـصحح (Adjusted $\mathbf{R^2}$ $\mathbf{(R^2_{adj})}$) ومعيـار أكـاكي للمعلومـات (Akaike Information Criteria (AIC)) حيث يعتمد المعيار الاول علـى اختيـار افضل نموذج جزئي ذي اعلى قيمة بينما يعتمد المعيار الثاني على اختيار افضل نمـوذج جزئي ذي اقل قيمة(Hughes & King, 2003)، (Lee & Russell, 2004)، (Shao et al,2009).

يعرف معامل التحديد المصحح بالصيغة الاتية:

$$\mathbf{R^2_{adj} = 1 - [\frac{NT-1}{NT-k-1}(1 - R^2_{panel\ data\ model})]} \quad \ldots\ldots(14)$$

أما معيار أكاكي للمعلومات فيعرف بالشكل الاتي:

$$\mathbf{AIC = -2\log(maximized\ likehood) + 2k} \quad \ldots\ldots(15)$$

5- الجانب التطبيقي

مما لاشك فيه ان نوعية البيانات التي تلائم الجانب التطبيقي من جانب ودقتهـا مـن جانب اخر يعدان ركيزة أساسية في الوصول الى نتائج يمكن الاعتماد عليهـا والركـون اليها في اي عمل بحثي. من هنا جاءت فكرة الخروج عن المؤلوف في اغلب الدراسـات والتطرق الى الجانب المالي من خلال دراسة الموازنة العامـة للدولـة مقاسـة بملايـين الدولارات كمتغير استجابة **(y)** وتأثيرها في عدد من المتغيرات التفسيرية وحسب وجهة نظر اهل الاختصاص وهي المعدل السنوي لسعر النفط لسلة خامات اوبك مقاسة بالدولار الامريكي**(x_1)**، والعجز في ميزان المدفوعات ويكون مقاساً بملايين الـدولارات**(x_2)**، حجم السكان مقاسا بالمليون نسمة**(x_3)**، واخيرا معدل التضخم**(x_4)**. شـملت الدراسـة الدول العربية البالغ عددها 18 دولة كبيانات مقطعية **($N = 18$)** باستثناء العراق الذي لـم تتوافر الإحصائيات كامل عنه والصومال وفلسطين لان ليس هنـاك دقـة فـي تـسجيل الاحـصائيات والـدول هـي (الاردن(a)، الامـارات(b)، البحـرين(c)، تـونس(d)،

الجزائر(e)، جيبوتي(f)، السعودية(g)، السودان(h)، سورية(i)، عمـان(j)، قطـر(k)، الكويت(l)، لبنان(m)، ليبيا(n)، مصر(o)، المغرب(p)، موريتانيا(q)، اليمن(r)) وللفترة الزمنية من 1995 الى 2009 **(T = 15)** وبحجم عينة 270 مشاهدة **(N * T = 270)** تم الحصول على بيانات الدراسة بالاعتماد على التقرير الاقتصادي العربي الموحـد لـسنة 2010 (صندوق النقد العربي واخرون، 2010). والنموذج الرياضي المقـدر للدراسـة حسب المعادلة (1) يكون بالصيغة الاتية:

$$\hat{y}_{it} = \hat{\beta}_{\circ(i)} + \sum_{j=1}^{4} \hat{\beta}_j \, x_{j(it)} \qquad ,i = 1,2,.....,18 \quad t = 1,2,....,15$$

5-1 تقدير معلمات النموذج

في هذا المبحث سوف نقدر معلمات النموذج المدروس باستخدام نمـاذج البيانـات الطولية الثلاثة وهي نموذج الانحدار التجميعي، ونمـوذج التـأثيرات الثابتـة، ونمـوذج التاثيرات العشوائية. الجدول(2) يوضح نتائج التقدير للمعلمات عند استخدام النماذج الثلاثة بالاعتماد على برنامج EViews 7. في ضوء نتائج تقدير المعلمات الموضحة بالجـدول (2) نلاحظ بان قيم معامل القطع (constant) تختلف من دولة الى اخرى حيـث يمثـل المقـدار (639.609-) قيمـة معامـل القطـع لدولـة الاردن فـي حـين تبلـغ قيمـة الاخـتلاف فـي دولـة الامـارات (10.341-) اي (639.609+629.268-) وهكذا لباقي الدول الاخرى ويعود سبب هذا الاختلاف في قـيم معامل القطع الى خصوصية كل دولة. من الجدول (2) يمكـن كتابـة معادلـة نمـوذج الانحدار التجميعي المقدرة، معادلة نموذج التاثيرات الثابتة المقدرة، ونمـوذج التـاثيرات العشوائية المقدرة على الترتيب بالشكل الاتي:

$$\hat{y}_{it} = -1246.636 + 48.863\,x_1 + 0.8011\,x_2 - 0.0857\,x_3 + 37.45\,x_4$$

$$\hat{y}_{it} = -639.609 + 87334\,x_1 + 0.747\,x_2 - 0.688\,x_3 + 7.759\,x_4 + 629.268D_b - 2269.77D_c + 2975.479D_d + 16760.03D_e - 2163.923D_f + 19536.45D_g + 19248.92D_h + 7564.28D_i - 1715.25D_j - 2031.49D_k + 5443.78D_l - 3679.22D_m + 1331.36D_n + 36817.78D_o + 15494.8D_p - 662.84D_q + 9883.53D_r$$

$$\hat{y}_{it} = -1272.65 + 52.066\,x_1 + 0.779\,x_2 - 0.088\,x_3 + 39.93\,x_4$$

الجدول(2): معلمات نموذج الدراسة المقدرة باستخدام النماذج الثلاثة

المتغيرات التوضيحية والوهمية	نموذج الانحدار التجميعي	نموذج التاثيرات الثابتة	نموذج التاثيرات العشوائية
constant	**-1246.636**	**-639.609**	**-1272.65**
x_1	**48.863 (*)**	**87.334 (*)**	**52.066 (*)**
x_2	**0.8011 (*)**	**0.747 (*)**	**0.779 (*)**
x_3	**-0.0857 (*)**	**-0.688 (*)**	**-0.0888 (*)**
x_4	**37.45**	**7.759**	**38.93**
D_b		629.268	
D_c		-2269.77	
D_d		2975.479	
D_e		16760.03 (*)	
D_f		-2163.923	
D_g		19536.45 (*)	

D_h		19248.92 (*)	
D_i		7564.28	
D_j		-1715.25	
D_k		-2031.49	
D_l		5443.78 (*)	
D_m		-3679.22	
D_n		1331.36	
D_o		36817.78 (*)	
D_p		15494.8 (*)	
D_q		-662.84	
D_r		9883.53 (*)	
R^2	**0.7062**	**0.688**	**0.683**

(*)معنوية المعلمة عند 0.05

بعد تقدير النماذج الثلاثة للنموذج المدروس سوف ننتقل الى استخدام اساليب الاختيار بين هذه النماذج الثلاثة من خلال اختبار **F** المقيد واختبار Hausman وكما هو موضح في الجدول (3).

الجدول(3): اختبار F المقيد واختبار Hausman

نوع الاختبار	قيمة الاختبار	p-value
اختبار **F** المقيد	2.779	0.0002 (*)
اختبار Hausman	29.63	0.000 (*)

(*)معنوية الاختبار عند 0.05

بالاعتماد على نتائج الجدول (3) نلاحظ بان النموذج الملائم لبياناتنا المدروسة هو نموذج التأثيرات الثابتة حيث لاحظنا من اختبار **F** المقيد بان نموذج التاثيرات الثابتة تغلب من ناحية الاهمية في الاستخدام مقارنة بنموذج الانحدار التجميعي، وعند اجراء اختبار Hausman ايضا تغلب نموذج التاثيرات الثابتة على نموذج التاثيرات العشوائية من ناحية اهمية هذا النموذج في ملاءمة للبيانات المدروسة. تجدر الاشارة هنا الى ان قيمة اختبار Hausman تحسب لمعلمات المتغيرات المدروسة وهي (x_1، x_2، x_3، و x_4) اي (β_1، β_2، β_3، وβ_4). في ضوء نتائج تقدير نموذج التاثيرات الثابتة نلاحظ بان سعر النفط وعجز ميزان المدفوعات وحجم السكان تمثل متغيرات رئيسية في الموازنة العامة للدولة في حين اظهر التضخم عدم معنويته في الموازنة العامة. لغرض الوصول الى افضل نموذج يحتوي على اهم المتغيرات المدروسة لنموذج التاثيرات الثابتة سوف نقوم بتوفيق عدد من النماذج مقدارها 15 نموذجا وفي كل مرة سوف نقوم بتوفيق نماذج البيانات الطولية الثلاثة للتاكد من ان نموذج التاثيرات الثابتة هو النموذج الملائم ثم بعد ذلك نقوم بحساب معيار معامل التحديد المصحح ومعيار اكاكي للمعلومات لكل نموذج.الجدول (4) يوضح قيم معياري R^2_{adj} و**AIC**.

الجدول (4): قيم معياري معامل التحديد المصحح واكاكي للمعلومات

النموذج	R^2_{adj}	AIC
x_1	0.244	21.52
x_2	0.717	20.538
x_3	0.1648	21.62
x_4	0.1562	21.631
x_1x_2	0.7256	20.511
x_1x_3	0.2472	21.52
x_1x_4	0.2447	21.523
x_2x_3	0.7164	20.544
x_2x_4	0.717	20.539
x_3x_4	0.1725	21.615
$x_1x_2x_3$	0.733	20.485
$x_1x_2x_4$	0.725	20.513
$x_1x_3x_4$	0.2452	21.526
$x_2x_3x_4$	0.716	20.546
$x_1x_2x_3x_4$	0.7323	20.786

في ضوء نتائج الجدول (4) يتضح بان معيار R^2_{adj} ومعيار **AIC** قد اختارا النمـوذج (x_1 ، x_2 ، و x_3) حيث اعطى اعلى قيمة لمعيار R^2_{adj} واقـل قيمـة لمعيـار **AIC** ويعبر عن افضل نموذج تاثيرات ثابتة جزئي بالشكل الاتي:

$$\hat{y}_{it} = -569.822 + 88.252x_1 + 0.747x_2 - 0.703x_3 + 615.06D_b - 2350.9D_c + 3043.6D_d + 17172D_e - 2232.6D_f + 19768.94D_g + 19826.85D_h + 7749.03D_i - 1769D_j - 2085.27D_k + 5400.6D_l - 3697.4D_m + 1328.55D_n + 37759.9D_o + 15845D_p - 678.5D_q + 10188.9D_r$$

6- الاستنتاجات

1- ان استخدام مفهوم نماذج البيانات الطولية أعطى دراسة شاملة عند استخدام البيانـات المقطعية والسلاسل الزمنية في ان واحد حيث اعطى نموذج رياضي واحد يحكم طبيعـة الدراسة مقارنة فيما لو تم استخدام 15 نموذج انحدار اعتيادي (نموذج واحد لكل سلـسلة زمنية) او مقارنة فيما لو تم استخدام 18 نموذج انحدار سلسلة زمنية (نمـوذج انحـدار سلسلة زمنية لكل دولة).

2- زيادة درجات الحرية عند استخدام نماذج البيانات الطولية حيث بلغت عدد المشاهدات 270 مشاهدة في حين بلغت درجة حرية الخطأ 265 عند اسـتخدام نمـوذج الانحـدار التجميعي ونموذج التاثيرات العشوائية في حين كانت درجـة حريـة الخطـا 249 عنـد استخدام نموذج التاثيرات الثابتة وهذا ما يعطي دقة في نتائج التقدير للمعلمات وان درجة الحرية هذه كانت اكبر من درجات الحرية البالغة 13 عند استخدام نمـوذج واحـد لكـل سلسلة زمنية و 10 عند استخدام نموذج انحدار سلسلة زمنية.

3- من نتائج الجدول (2) نلاحظ بان قيمة معامل التحديد تكون اكبر في حالـة اسـتخدام نموذج الانحدار التجميعي مقارنة باستخدام نموذج التاثيرات الثابتة والعشوائية وكذلك بان معامل التحديد الخاص بنموذج التاثيرات العشوائية يكون اقل مـن نظيـره فـي نمـوذج

التاثيرات الثابتة. لذا لا يمكن الاعتماد على معامل التحديد في اختيـار نمـوذج البيانـات الطولية الملائم لأية دراسة بسبب ان معامل التحديد يعتمد على مقاييس مختلفة في حسابه من نموذج لآخر.

4- من خلال اجراء اختبار **F** المقيد واختبار Hausman لاحظنا بان نموذج البيانـات الطولية الملائم لبياناتنا هو نموذج التاثيرات الثابتة ومن خلال اسـتخدام معيـار معامـل التحديد المصحح ومعيار اكاكي للمعلومات تبين بان افضل نموذج جزئي هـو النمـوذج الذي يحوي على المتغيرات التفسيرية وهي المعدل السنوي لسعر النفط، عجـز ميـزان المدفوعات، و حجم السكان.

7- المصادر

1. صندوق النقد العربي، واخرون، 2010، " التقرير العربي الاقتصادي الموحد"، شركة ابو ظبي للطباعة والنشر.

2. Baltagi, B., H., Jung, B., Ch. and Song, S., H., 2010, "Testing for heteroskedasticity and serial correlation in a random effects Panel Data Model", Journal of Econometrics, Vol.154, Iss.2, pp. 122-124.
3. Blatagi, B., H., 2005, "Econometric Analysis of Panel Data", 3rd ed., John Wiley & Sons, Ltd, West Sussex.

4. Bramati, M. and Croux, Ch., 2007, “Robust Estimators for the Fixed Effects Panel Data Model", Econometrics Journal, Vol.10, Iss.3, pp. 1-19.

5. Chuang, Ch. and Wang, Y., 2009, " Developed stock market reaction to political change: a panel data analysis", Quality and Quantity, Vol. 43, pp. 941-949.
6. Dustmann, Ch. And Engracia, M., 2007, " Selection correction in Panel Data Models: An Application to the Estimation of Female's Wage Equations", Econometrics Journal, Vol.10, Iss.2, pp. 263-293.

7. El-Gamal, M., A. and Inanoglu, H., 2005, " Inefficiency and Heterogeneity in Turkish Banking: 1990-2000", Journal of Applied Econometrics, Vol. 20, No. 5 , pp. 641-664.
8. Greene, W., H., 2012, “Econometrics Analysis", 7th ed., Pearson Education, Inc., NJ.
9. Hausman, J., 1978, "Specification Test in Econometrics", Econometrica, Vol.46, pp.1251-1271.
10. Hughes, A., W. and King, M., L., 2003, " Model Selecting using AIC in the Presence of One-Sided Information", Journal of Statistical Planning and Inference, Vol.115, pp. 397-411.
11. Kai,T. and Qin, L., 2011, " Analysis on consumption structure of Shanghai resident by using Panel data model", Proceeding of IEEE 2nd International Conference on Software Engineering and Service Science, pp. 819-822.
12. Lee, D., G. and Russell, J., S., 2004, " Panel Data Analysis of Factors Affecting As-Built Roughness of Asphaltic Concrete Pavements", Journal of Transportation Engineering, Vol.4, pp. 479-485.
13. Lee, L. and Yu, J., 2010, "Estimation of spatial autoregressive Panel Data Models with fixed effects", Journal of Econometrics, Vol.154, pp.165-185.
14. Lukas, M. and Jan, N., 2011, "Application of Econometric Panel Data Model for Regional Competitiveness Evaluation of Selected EU 15 Countries", Journal of Competitiveness, Issue 4, pp. 23-38.
15. Mikhed, V. and Zemcik, P., 2009, "Testing for Bubbles in Housing Markets: A Panel Data Approach", Journal of Real Estate Financial Economics, No.38, pp. 366-386.
16. Shao, X., Chen, Y., and Wang, H., 2009, " Research on Financial Warning for Chinese Listed Companies by Using Panel Data Model", Proceeding of IEEE International Conference on System, pp. 4663-4667.

17. Sun, Y., 2009, "Estimation of semiparametric regression model with longitudinal data ", Lifetime Data Analysis, Vol.16, pp. 271-298.