جمهورية مصر العربية

جامعة المنصورة
كلية التجارة
قسم الاحصاء التطبيقي و التأمين

تقديرات النماذج شبه المعلمية بالتطبيق على بيانات العشارى

تحت إشراف

أ.د/ فاطمة على محمد عبد العاطي	د/أشرف أحمد عبد العليم البدرى
أستاذ الاحصاء التطبيقي	مدرس بقسم الإحصاء التطبيقي
كلية التجارة–جامعة المنصورة	كلية التجارة–جامعة المنصورة

إعـداد

منى محمد أحمد بصل

٢٠٢٠

ملخص:-

تحتوي نماذج الانحدار على مكونين أحدهما مكون عشوائي يمثل حد الخطأ (Error Term) والأخر مكون غير عشوائي يمثل متغير الاستجابة (Response Variable) وهذا المكون هو محل الاهتمام فيتم تقديره بإحدى الطريقتين هما: الطريقة المعلمية أو الطريقة اللامعلمية أو كلاهما .

نموذج الانحدار المعلمي (Parametric Regression Model) هو النموذج الأكثر شيوعا حيث أنه يفترض أن متغير الاستجابة له صيغه دالية محدده مسبقا والتي تم الحصول عليها من معلومات سابقه حول هيكل دالة الانحدار من خلال ذلك يتم تقدير فئة من المعالم. ورغم أن النماذج المعلمية تعد أحد أهم أدوات تحليل البيانات؛ لأنها فعاله ويمكن تفسيرها بسهوله، إلا انها لم تستطع ان تكون كافية في بعض الحالات .

أما نموذج الانحدار اللامعلمي (Nonparametric Regression Model) فانه يعتمد على تقدير دالة الانحدار مباشرة من البيانات كما أنه لا يفترض فروضا معينه لتقدير معالم النموذج مثل التي يفترضها النموذج المعلمي وبالتالي فهو يتمتع بمرونة أكثر لاكتشاف البيانات التي قد تكون مفقودة او في الحالات التي لا تتوافر فيها معلومات سابقة.

يتم انشاء نموذج الانحدار شبه المعلمي (Semi parametric Regression Model)من خلال مزج الطريقتين السابقتين معا. فغالبًا ما يتم استخدام هذا النموذج في المواقف التي تكون فيها الافتراضات المعلمية غير محددة وغير متسقة، أو أن النموذج اللامعلمي لا يعمل بشكل كامل.

يشتمل هذا البحث على تطبيق نماذج الانحدار المعلمية وشبه المعلمية بالإضافة إلى النموذج اللامعلمى كحالة خاصة من نموذج الانحدار شبه المعلمى لتقدير الدالة

الافتراضية وتحديد شكل العلاقة بين المتغير التابع والمتغيرات المستقلة بالاضافة الى شكل انتشار البيانات مع خط (أو منحني) التقدير وذلك على بيانات واقعية تم الحصول عليها من الإتحاد المصري لألعاب القوى والمعروفة ببيانات لعبة العشاري .

الكلمات الافتتاحية:-

الانحدار المعلمى، الانحدار اللامعلمى، الانحدار شبه المعلمى.

Abstract:-

Regression models contain two components, one of which is a random component that represents the Error Term and the other is a non-random component that represents the response variable and this component is of interest and is estimated in one of two methods: the parametric method or the nonparametric method or both.

Parametric regression model is the most common model since it assumes that the response variable has a predetermined functional formula which was obtained from previous information about the structure of the regression function by which a class of parameters is estimated. Although parameter models are one of the most important tools for analyzing data; Because it is effective and can be explained easily, but it could not be sufficient in some cases.

As for the nonparametric regression model, it relies on estimating the regression function directly from the data, as it does not assume specific assumptions for estimating the model parameters, such as those assumed by the parametric model, and thus it has more flexibility to discover data that may be missing or in cases where previous information is not available.

The semi-parametric regression model is constructed by blending the two previous methods together. This model is often used in situations where the parametric assumptions are indeterminate and inconsistent, or the nonparametric model is not fully functional.

This research includes the application of parametric and semi-parametric regression models in addition to the nonparametric model as a special case of the semi-parametric regression model for estimating the hypothetical function and determining the form of the relationship between the dependent variable and the independent variables in addition to the data spread form with a line (or curve) estimation on realistic data On it from the Egyptian Athletics Federation, known as Decathlon data.

Key Words:
Parametric Regression, Nonparametric Regression, Semi-Parametric Regression.

مقدمة:

أثبت الانحدار شبه المعلمي أنه ذو قيمة كبيرة في العديد من التطبيقات في مجالات متنوعة مثل علم الفلك والبيولوجيا والطب والاقتصاد والتمويل ، باستخدام نماذج الانحدار شبه المعلمية ، يمكن استخراج معلومات مهمة من مجموعات البيانات التي غالباً ما تكون غير مستقرة.

يتضمن الانحدار شبه المعلمى نماذج انحدار تجمع بين :-

- النماذج المعلمية Parametric Model .
- النماذج اللامعلمية Nonparametric Model .

في الانحدار المعلمي تأثير كل متنبئ له شكل بسيط بحيث يكون الشكل العام للدالة من قبل النموذج وليس البيانات، كما يمكن لهذه الفئة من النماذج تلخيص العلاقات بين المتغيرات التابعة والمتغيرات المستقلة من خلال بعض المعلمات البسيطة .

للطريقة المعلميه عدة مزايا مثل أن تفسير المقدر يكون سهلاً ، وخوارزمية التقدير بسيطة ، ومعدل التقارب للمقدر يكون أسرع من معدل المقدر اللامعلمي. (Ravikumar et al., 2009)

ومع ذلك قد يكون أداء المقدر المعلمي غير مضمون حيث أنه يعاني من النقص المحتمل في تحديد النموذج الصحيح ، مما قد يؤدي إلى استنتاجات خاطئة. وبناءا على ذلك فغالبًا ما يكون من الضروري تخفيف بعض القيود المعلمية.

الانحدار اللامعلمي هو فئة من تحليل الانحدار لا يأخذ فيها المتنبئ شكلاً محددًا مسبقًا ولكن يتم إنشاؤه وفقًا للمعلومات المستمدة من البيانات. يتطلب الانحدار اللامعلمي أحجام عينات أكبر من النماذج المعلمية لأن البيانات يجب أن توفر بنية النموذج بالإضافة إلى تقديرات النموذج.

توفر نماذج الانحدار اللامعلمية هياكل قابلة للتفسير علميًا وذات مغزى ، فتكون الفكرة الرئيسية هي فرض بعض الهياكل القابلة للتتبع رياضيًا من أجل الحفاظ على مرونته و قابلية تفسيره ، وعليه فإن المعلمات الرئيسية ذات الأهمية هي الدوال أو المنحنيات ، والتي نشير إليها باسم "معلمات الدالة " أو "معلمات المنحنى" ، يكون النموذج مرنًا بدرجة كافية للسماح لأي اتجاه ممهد في البيانات .(Ravikumar et al., 2009)

وبناءا على ذلك يمكن القول بأن نموذج الانحدار شبه المعلمي نموذجا وسطا بين نموذجي الانحدار المعلمي والانحدار اللامعلمي. حيث أنه يجمع بتأثيرات النمذجة

المعلمية لبعض التنبؤات مع النمذجة اللامعلمية لتأثيرات المتغيرات الأخرى. كما أنه يسمح بإجراء تحويلات خطية للبيانات بسهولة نظرًا لمرونته. فغالبًا ما يتم استخدام هذا النموذج في المواقف التي تكون فيها الافتراضات المعلمية غير محددة وغير متسقة، أو أن النموذج اللامعلمي لا يعمل بشكل كامل. وعلى ذلك فانه يجمع بين مزايا كلا من نماذج الانحدار المعلمية من حيث الكفاءة وسهولة التفسير ونماذج الانحدار اللامعلمية من حيث المرونة. (Huang et al., 2012)

وحديثا تطورت تلك النماذج تطوراً كبيراً، وتُستخدم نماذج الانحدار شبه المعلمية لدراسة العلاقة بين المتغير التابع والمتغيرات المفسرة عندما تكون الصيغة الرياضية للعلاقة بين المتغير التابع وواحد أو أكثر من المتغيرات المفسرة معلومة مع وجود متغير مفسر على الأقل الشكل الرياضى للعلاقة بينه وبين المتغير التابع غير معلومة.(Yoshida, 2018)

الاستعراض المرجعى:

تناول(Yoshida, 2018) طريقة شبه معلمية باستخدام نماذج الانحدار المضافة وذلك من خلال إعداد نموذج معلمي والحصول على مقدرات لجميع المكونات المضافة. بعد ذلك ، بالنسبة للبيانات المتبقية المرتبطة بالمقدر المعلمية ، يتم تطبيق النموذج اللامعلمي. يتم إنشاء المقدر النهائي عن طريق جمع المقدر المعلمى والمقدر اللامعلمي للبيانات المتبقية. لكل مكون مضاف ، إذا أصبح المقدر اللامعلمي دالة صفرية ، يتم تقليل المقدر النهائي إلى مقدر معلمى. وبالتالى يتم اكتشاف بنية النموذج بعد ذلك وعرض الخصائص المقاربة للمقدر المقترح.

يقدم (Reda Abonazel & Gad, 2018) نسخة من تقنية البواقى الجزئية لتقدير المكونات المعلمية واللامعلمية في النموذج الخطي الجزئي شبه المعلمى. يتم إنشاء التقدير للمكون المعلمى بعد استبعاد تأثير المكون اللامعلمي على كل من الاستجابة

والمتغيرات المشتركة بناءً على البيانات الزائفة. أخيرًا ، يتم تقدير المكون اللامعلمي توضح دراسات المحاكاة وتحليل البيانات الحقيقي أن المقدر المقترح يؤدي بشكل أفضل من التقديرات الحالية عند القيم المتطرفة في مجموعة البيانات أو الأخطاء ذات الذيل الثقيل.

كما تناول(Huang et al., 2012) اقتراح طريقة متابعة الانحدار شبه المعلمى لتمييز المكونات الخطية والغير الخطية في النماذج الخطية الجزئية شبه المعلميه. ويحدد هذا المنهج المكونات المعلمية واللامعلمية فى النموذج شبه المعلمى على أساس البيانات، هذه الطريقة تختلف اختلافًا جوهريًا عن أسلوب الاستدلال شبه المعلمى القياسي حيث يتم تحديد المكونات المعلمية واللامعلمية في النموذج مسبقًا.هذه الطريقة مماثلة لمقياس الشبه المعلمى القياسي بافتراض أن بنية النموذج معروفة.كما أن النموذج الخطى الجزئى الشبه معلمى يتيح وضع نماذج مرنه للتأثيرات المشتركة على متغير الاستجابة فى الانحدار. فهو يجمع بين المرونة فى الانحدار اللامعلمى والثبات فى الانحدار الخطى.

استعرض (Li & Liang, 2008) كيفية اختيار المتغيرات المهمة في النمذجة شبه المعلمية. يتكون الاختيار المتغير لنماذج الانحدار شبه المعلمية من مكونين: اختيار النموذج للمكونات اللامعلمية واختيار المتغيرات المهمة للجزء المعلمى. وبالتالي ، يعد اختيار المتغير شبه المعلمى أكثر صعوبة من اختيار المتغير المعلمى وذلك لأن إجراءات الاختيار المتغيرة التقليدية بما في ذلك الانحدار التدريجي وأفضل اختيار مجموعة فرعية تتطلب الآن اختيار نموذج منفصل للمكونات غير المعلمية لكل نموذج فرعي. هذا يؤدي إلى عبء حسابي ثقيل للغاية. في هذا البحث تم اقتراح فئة من إجراءات الاختيار المتغيرة لنماذج الانحدار شبه المعلمية باستخدام الاحتمالية المعاقبة.

تناول (Naito, 2002) الانحدار شبه المعلمى باستخدام كل من النماذج المعلمية والتمهيد اللامعلمي. يُنظر إلى النموذج المعلمى على أنه تخمين خام لدالة

الانحدار الحقيقية ويتم تعديله بواسطة العامل اللامعلمي. يتم تحديد عامل الضبط اللامعلمي بمعيار يسمى تركيب L_2 المحلي. يتم تكوين فئة من مقدرات الانحدار مع تعديل مضاعف. المقدرون في الفصل متصلون بمعامل واحد ويظهر أن دالة التحيز للمقدر في الفصل تكون خطية في حين أن التباين مجاني. هذا يجعل من الممكن اكتشاف أفضل مقدر في الفصل.

الخصائص الرياضية لنماذج الانحدار:-

أولا: تقديرات النماذج المعلمية Estimation of Parametric Models :

في النماذج المعلمية يكون شكل العلاقة معرفة في إحدي صور الانحدار الخطي linear regressionأو غير الخطي nonlinear regression. سيتم فقط عرض نماذج الانحدار الخطي (المتعددة) كأحد أهم وأشهر صور نماذج الانحدار المعلمي، والصورة العامة لنماذج الانحدار الخطي باستخدام المصفوفات تأخذ الشكل التالي:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (1)$$

حيث أن Y تمثل متجه المتغير التابع بينما الـ X تمثل مصفوفة المتغيرات المستقلة مضافا اليها عمود الثوابت (العمود الأول) والـ β تمثل متجه المعاملات للمتغيرات المستقلة بما فيهم معامل ثابت الانحدار وأخيراً ε تمثل متجة الخطأ العشوائي والذي يفترض أنه يتبع التوزيع الطبيعي بمتوسط صفر وانحراف معياري قدره σ كما يفترض أن الأخطاء العشوائية مستقلة عن (X,Y). بينما الصورة العادية لمعادلة الانحدار الخطي المتعدد تأخذ الشكل التالي:

$$\mathbf{Y_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}\mathbf{X_{1i}} + \boldsymbol{\beta_2}\mathbf{X_{2i}} + \cdots + \boldsymbol{\beta_d}\mathbf{X_{di}} + \boldsymbol{\varepsilon_i} \qquad (2)$$

حيث أن d عدد المتغيرات المستقلة والـ i عدد المشاهدات (حجم العينة n).

ثانيا: تقديرات النماذج شبه المعلمية Estimation of Semi-Parametric Models :

في النماذج شبه المعلمية يكون شكل العلاقة معرفة في إحدي صور الانحدار الخطي linear Regression بالإضافة إلى إحدي صور الانحدار غير الخطي Nonlinear Regression أو معرفاُ بإحدي نماذج الانحدار الخطي بالإضافة إلى نماذج الانحدار غير المعلمي Non–Parametric Regression والتي فيها يكون شكل الدالة غير معروف وليس له دالة معلمية. سيتم فقط عرض نماذج الانحدار شبه المعلمي التي تحتوي على الانحدار الخطي ونماذج الانحدار غير المعملي. والصورة العامة لنماذج الانحدار شبه المعلمي باستخدام المصفوفات تأخذ الشكل التالي:

$$\mathbf{Y} = \mathbf{X\beta} + \mathbf{f(Z)} + \boldsymbol{\varepsilon} \qquad (3)$$

حيث أن Y تمثل متجه المتغير التابع بينما الـ X تمثل مصفوفة المتغيرات المستقلة مضافا اليها عمود الثوابت (العمود الأول) والـ β تمثل متجه المعاملات للمتغيرات المستقلة بما فيهم معامل ثابت الانحدار بينما الـ $f(Z)$ دالة ممهدة Smoothing Function للمتغيرات Z وأخيراً ε تمثل متجة الخطأ العشوائي والذي يفترض أنه يتبع التوزيع الطبيعي بمتوسط صفر وانحراف معياري قدره σ كما يفترض أن الأخطاء مستقلة عن أزواج المتغيرات المستقلة والتابعة. والصورة العامة لنماذج الانحدار شبه المعملي تأخذ الشكل التالي:

$$\mathbf{Y_i} = \mathbf{f(x_i)} + \boldsymbol{\varepsilon}_\mathbf{i} \qquad (4)$$

حيث أن الدالة $f(x_i)$ تمثل دالة الانحدار شبه المعلمي بشكل عام. إذا كانت الدالة $f(x_i)$ معرفة بصورة المتغيرات المستقلة والمعاملات لها فإنها تحول النموذج إلي نموذج انحدار معلمي، أما إذا كانت الدالة $f(x_i)$ تحتوي علي جزء معلمي وجزء لامعلمي

يتحول النموذج إلي نموذج انحدار شبه معلمي. ومن أبسط أشكال دالة الانحدار شبه المعلمي صورة الدالة الممهدة

$$\mathbf{f(x)} = \sum_{j=0}^{p} \boldsymbol{\beta}_j \mathbf{x}^j + \sum_{k=1}^{K} \boldsymbol{\mu}_k (\mathbf{x} - \boldsymbol{\kappa}_k)^{p}{}_{+} \qquad (5)$$

الدالة السابقة تحتوي علي جزئين هما:

الجزء المعلمي: $\sum_{j=0}^{p} \beta_j x^j$ ويمثل الجزء المعلمي وفيه الـ p هي درجة النموذج متعدد الحدود Polynomial بمعني أن p=2 تعني أن النموذج متعدد الحدود من الدرجة الثانية وأعلي أس للمتغير x هو 2.

الجزء اللامعلمي: $\sum_{k=1}^{K} \mu_k (x - \kappa_k)^{p}{}_{+}$ ويمثل دالة ممهدة أحادية Univariate Smoothing للمتغير X في ابسط صورة لها وهي صورة المقياس المطلق (ويرمز له بالرمز $(\ldots)_+$) للفرق بين المتغير X وقيمة تبعد عن هذا المتغير بمقدار κ_k حيث أن k تمثل عدد العقد knots والـ p تمثل درجة الانحدار المتعدد مع ملاحظة أن عدد الحدود الافتراضي p=1. في هذا الجزء ايضا نجد μ_k والتي تمثل معاملات الجزء اللاملعمي وعددها k وهي معرفة بالمعادلة التالية:

$$\mathbf{u} = [\mathbf{u_1}, \ldots \mathbf{u_k}]^{\mathbf{T}} \sim \mathbf{N}\left(\mathbf{0}, \boldsymbol{\sigma}_{\mathbf{u}}^{2} \boldsymbol{\Omega}^{-\frac{1}{2}} \left(\boldsymbol{\Omega}^{-\frac{1}{2}}\right)^{\mathbf{T}}\right) \qquad (6)$$

حيث أن Ω معرفة كالتالي:

$$\Omega = \underbrace{\left[(\kappa_k - \kappa_{k'})^{p}{}_{+}\right]}_{1 \le k,\ \kappa' \le K}$$

يتم تقدير النموذج السابق باستخدام طريقة ممهد الشريحة المعاقب Penalized Spline Smoothing وفي هذه الطريقة يتم تقدير معلمة التمهيد Smoothing Parameter باستخدام طريقة الإمكان الأعظم المقيدة Restricted Maximum

(REML) Likelihood ويتم تقدير الدالة f(x) باستخدام طريقة أفضل تنبؤ خطي غير متحيز (EBLUP) Estimated Best Linear Unbiased Prediction.

بالنسبة لتقدير الدالة f(x) يمكن الحصول عليه من المعادلة التالية باستخدام طريقة Penalized Spline Smoothing

$$\hat{\mathbf{f}} = \mathbf{C}(\mathbf{C}^{\mathbf{T}}\mathbf{C} + \boldsymbol{\lambda}^{\mathbf{p}}\mathbf{D})^{-\mathbf{1}}\mathbf{C}^{\mathbf{T}}\mathbf{Y} \qquad (7)$$

حيث أن

$$C = \begin{bmatrix} 1 & x_i & \dots & x_i^p & |x_i - \kappa_k|^p \end{bmatrix}$$

$$D = \begin{bmatrix} 0_{2\times 2} & 0_{2\times k} \\ 0_{k\times 2} & \left(\Omega^{\frac{1}{2}}\right)^{T} \Omega^{\frac{1}{2}} \end{bmatrix}$$

وايضا λ تسمي المقدر الممهد Smoothing Parameters ويتم حسابها من العلاقة التالية:

$$\lambda = \frac{\sigma_u^2}{\sigma_\varepsilon^2}$$

وبالتالي فإنه يلازم لعملية تقدير دالة الانحدار شبه المعلمي تحديد عدد الـ k (عدد العقد) المناسب وكذلك درجة المتغير المستقل p.

الجانب التطبيقى:-

يتناول الجانب التطبيقى بيانات تم تجميعها من الاتحاد المصري لألعاب القوى للعبة العشاري في بطوله الجمهورية للكبار والتي تمت اقامتها بالمركز الأولمبي بالمعادي خلال الأعوام (2017–2018–2019) منقسم إلى مسابقتين (الكأس والدرع) ، يحتوي إطار البيانات على ٣٨ سجل خلال١٠ رياضات علما بأن عدد المشاهدات الفعلية ٣٦

مشاهدة وذلك بعد حذف قيم السجلات والتي تشمل إما قيم مفقودة missing data أو قيم متطرفة outliers والبالغ عددها ٢ سجلات، ونقصد هنا بالسجل صف البيانات الممثل لجميع المتغيرات.

وصف المتغيرات المستخدمة في البحث:

متغير الاستجابة:

y : وهو متغير يصف أداء الرياضيين خلال لعبة رمى الرمحJavelin Throw .

المتغيرات المستقلة:

x_1 : وهو متغير يصف أداء الرياضيين خلال لعبة قذف الكرة Shout Put.

x_2 : وهو متغير يصف أداء الرياضيين خلال لعبة رمى القرص Discus Throw.

x_3 : وهو متغير يصف أداء الرياضيين خلال لعبة القفز بالزانة Pole Vaulting.

سوف نرمز للمتغيرات بعد إعادة القياس لها بالرموز التالية:

JT: لمتغير الـ Javelin Throw وذلك بعد إعادة القياس له.

SP: لمتغير الـ shot.put وذلك بعد إعادة القياس له.

DT: لمتغير الـ Discus Throw وذلك بعد إعادة القياس له.

PV: لمتغير الـ Pole.valut وذلك بعد إعادة القياس له.

أولا:تقديرات النماذج المعلمية Estimation of Parametric Models

بالتطبيق على المتغيرات المستقلة من واقع البيانات التي تم الحصول عليها باستخدام طريقة المربعات الصغرى OLS تم الحصول على النتائج التالية:

جدول (١): تقديرات نموذج الانحدار المعلمي

العلاقة	$JT_i = \beta_0 + \beta_1 * SP_i + \beta_2 * DT_i^2 + \beta_3 * exp(PV_i^2)$			
المعامل	**التقدير Estimate**	**الخطأ المعياري Standard error**	**ت المحسوبة t value**	**المعنوية المحسوبة p-value**
الثابت: $\widehat{\beta_0}$	−0.0981	0.0856	−1.15	0.2603
معامل SP: $\widehat{\beta_1}$	0.9671	0.2385	4.05	0.0003***
معامل DT: $\widehat{\beta_2}$	0.2241	0.2105	1.06	0.2950
معامل PV: $\widehat{\beta_3}$	−0.0555	0.0706	−0.79	0.4374
<u>معنوية النموذج ككل:</u>				
الخطأ المعياري للبواقي = 0.102 ودرجات الحرية 32				
معامل التحديد R^2 = 0.762				
معامل التحديد المعدل = 0.739				
قيمة ف المحسوبة = 34.1 والمعنوية المحسوبة لها = 0.000***				

*** معنوي عند مستوي الدلالة الإحصائية 1%

التعليق علي الجدول السابق:

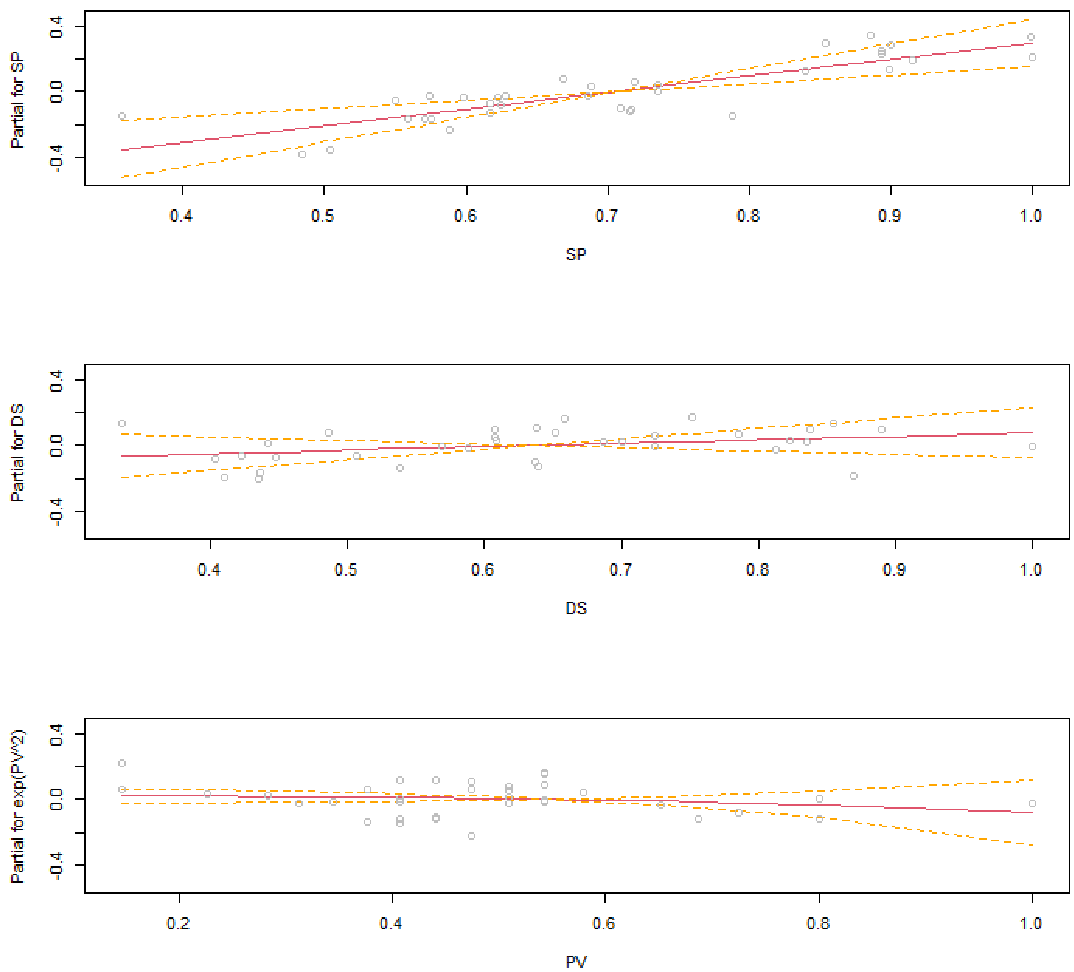
١. تقدير ثابت الانحدار $\widehat{\beta_0}$ كان 0.0981− بخطأ معياري قدره 0.0856 والمعنوية المحسوبة له 0.2603 مما يعني عدم معنوية ثابت الانحدار في معادلة التقدير. بينما كان تقدير $\widehat{\beta_1}$، معامل المتغير SP، 0.9671 بخطأ معياري قدره 0.2385 والمعنوية المحسوبة له 0.0003 مما يعني معنوية هذا المتغير في معادلة الانحدار. في حين كان تقدير $\widehat{\beta_2}$، معامل المتغير DT، 0.2241 بخطأ معياري قدره 0.2105 والمعنوية المحسوبة له 0.2950 مما يعني عدم معنوية هذا المتغير في معادلة الانحدار. أخيراً نجد أن تقدير $\widehat{\beta_3}$، معامل المتغير PV،

كان 0.0555- بخطأ معياري قدره 0.0706 والمعنوية المحسوبة له 0.4374 مما يعني عدم دلالة هذا المتغير في معادلة الانحدار.

٢. معامل التحديد R^2 للنموذج تساوي 0.762 مما يعني أن القوة التفسيرية للمتغيرات المستقلة للمتغير التابع قد بلغت 76.2% مما يعني أن مجمل التغيرات في المتغير التابع بسبب المتغيرات المستقلة هي 76.2% والباقي يرجع لعوامل أخري. يسري هذا التفسير ايضاً علي معامل التحديد المعدل والذي بلغت قيمته 73.9%.

٣. بالنسبة لمعنوية النموذج ككل نجد أن قيمة ف المحسوبة 34.1 وذلك بدرجات حرية الخطأ 32 ودرجات حرية الانحدار 3 (عدد المقدرات -1) كما أن المعنوية المحسوبة للإحصائية ف المحسوبة كانت أقل من 0.001 وهي أقل من مستوي الدلالة الإحصائية 5% (أو أي مستوي أخر يُرغب في استخدامه) مما يعني أن علاقة الانحدار معنوية وذلك عند مستوي دلالة إحصائية 1% .

الرسم التالي يوضح شكل الانتشار وخط الانحدار العام والبواقي لكل متغير تبعاً مع المتغير التابع .



شكل (١): علاقة الانتشار بين المتغيرات المستقلة وخط التقدير العام والبواقي

ثانيا: تقديرات النماذج شبه المعلمية Estimation of Semi-Parametric Models :

فيما يلي مجموعة من نماذج الانحدار شبه المعلمي المفترضه والتي تشمل نوعي الانحدار المعلمي واللامعلمى بالاضافة الى التقديرات والاشكال التوضيحية للتقديرات.

الحالة الأولي: بافترض أن متغير PV غير معلمي والمتغيرين الأخرين: SP, DT معلمين كالتالى:

جدول (٢-أ): تقديرات نموذج الانحدار شبه المعلمي الحالة الأولي

العلاقة	$JT \sim SP + DT^2 + f(\exp(PV^2))$			
تقديرات الجزء المعلمي	**التقدير Estimate**	**الخطأ المعياري Standard error**	**ت المحسوبة t value**	**المعنوية المحسوبة p-value**
الثابت: $\widehat{\beta_0}$	−0.04114	0.1144	−0.3596	0.7202
معاملSP: $\widehat{\beta_1}$	1.0370	0.2349	4.4130	0.000***
معامل DT: $\widehat{\beta_2}$	0.1254	0.1605	0.7812	0.4373
تقديرات الجزء اللامعلمي	**درجات الحرية Df**	**تقدير معلمة التمهيد** $\hat{\lambda}$	**تقدير عدد العقد K**	
دالة $\exp(PV^2)$	1.0295	19.63	3	

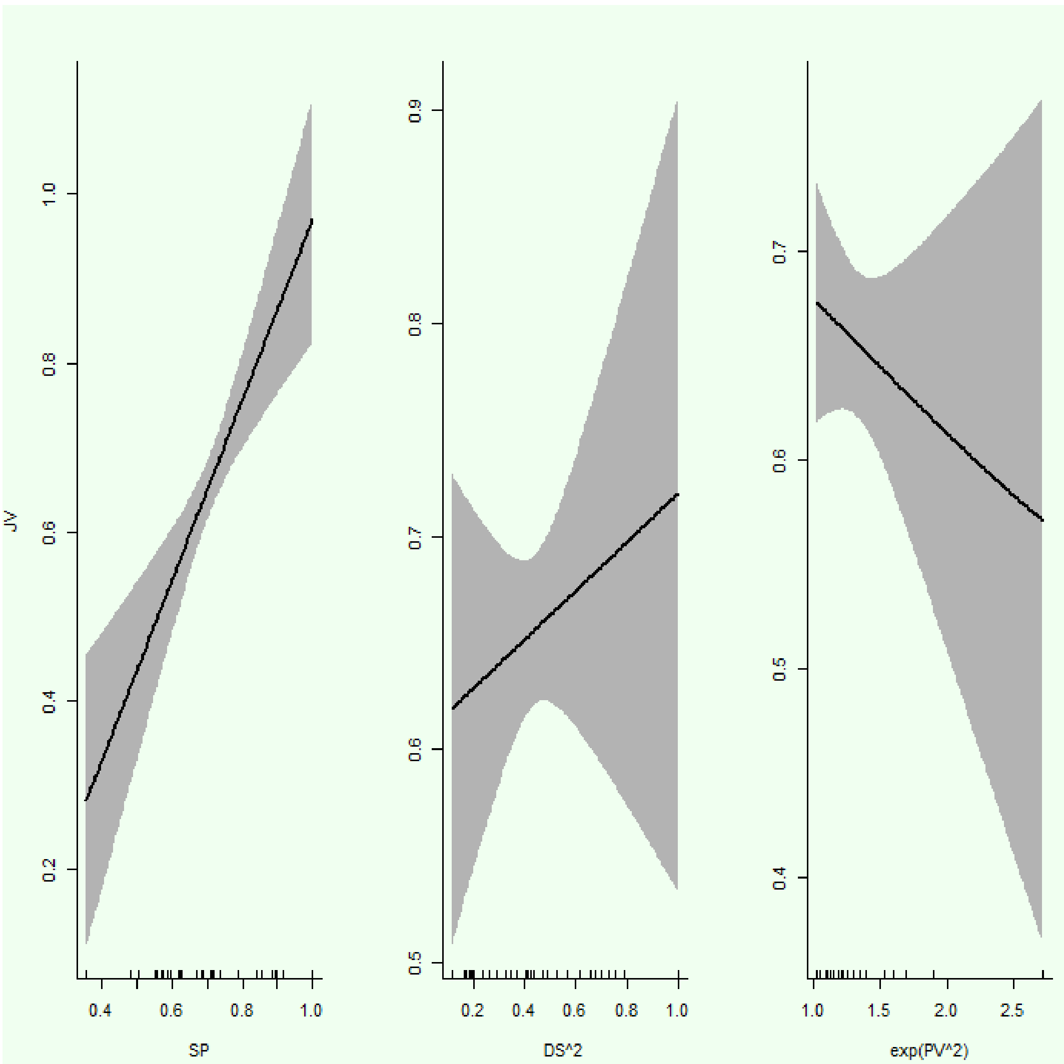
*** معنوي عند مستوي الدلالة الإحصائية ١%

التعليق علي الجدول السابق:

١. بالنسبة للجزء المعلمي: تقدير ثابت الانحدار $\widehat{\beta_0}$ كان −0.04114 بخطأ معياري قدره 0.1144 والمعنوية المحسوبة له 0.7202 مما يعني عدم معنوية ثابت الانحدار في معادلة التقدير شبه المعلمي. بينما كان تقدير $\widehat{\beta_1}$، معامل المتغير SP، 1.0370 بخطأ معياري قدره 0.2349 والمعنوية المحسوبة له 0.0000 مما يعني معنوية هذا المتغير في معادلة الانحدار شبه المعلمي. في حين كان تقدير $\widehat{\beta_2}$، معامل المتغير DT، 0.1254 بخطأ معياري قدره 0.1605 والمعنوية المحسوبة له 0.4373 مما يعني عدم معنوية هذا المتغير في معادلة الانحدار شبه المعلمي. يلاحظ أن التقديرات للجزء المعلمي في

نموذج الانحدار شبه المعلمي كانت قريبة من التقديرات في نموذج الانحدار المعلمي.

٢. بالنسبة للجزء اللامعلمي: نجد أن تقدير معلمة التمهيد كان 19.63 وعدد العقد المناسب للانحدار شبه المعلمي هو 3 .



شكل (٢): علاقة الانتشار شبه المعلمي بين المتغيرات المستقلة والمتغير التابع في الحالة الأولي

الحالة الثانية: بافتراض أن كلا من: متغير PV وDT غير معلمي بينما متغير SP معلمي و يكون شكل العلاقة والنتائج كما فى الجدول التالى :

جدول (٢-ب): تقديرات نموذج الانحدار شبه المعلمي الحالة الثانية

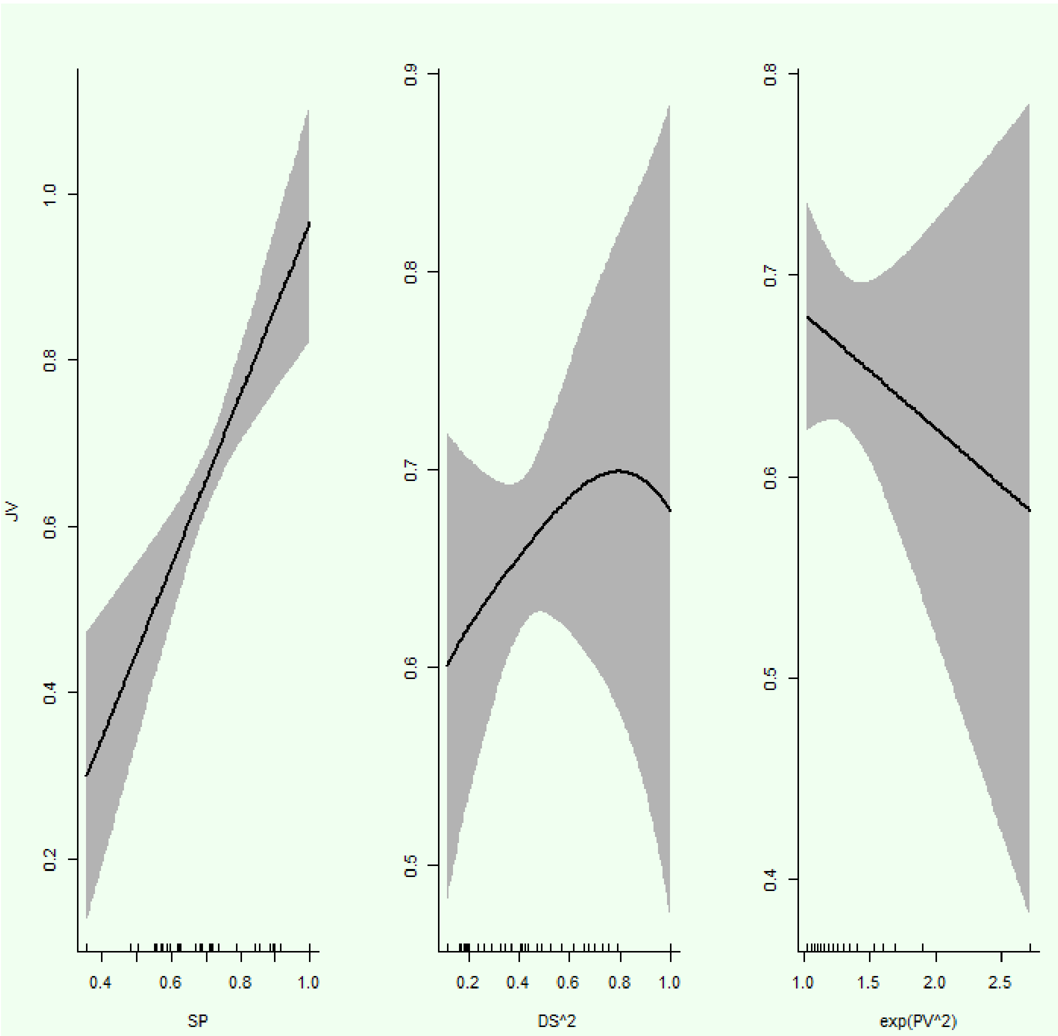
العلاقة	$JT \sim SP + f(DT^2) + f(\exp(PV^2))$			
تقديرات الجزء المعلمي	التقدير Estimate	الخطأ المعياري Standard error	ت المحسوبة t value	المعنوية المحسوبة p-value
الثابت: $\widehat{\beta_0}$	−0.0486	0.1121	−0.4334	0.6673
معاملSP: $\widehat{\beta_1}$	1.0020	0.2341	4.2800	0.001***
تقديرات الجزء اللامعلمي	درجات الحرية Df	تقدير معلمة التمهيد $\hat{\lambda}$	تقدير عدد العقد K	
دالة DT^2	1.392	1.88	7	
دالة $\exp(PV^2)$	1	3031	3	

*** معنوي عند مستوي الدلالة الإحصائية ١%

التعليق علي الجدول السابق:

١. بالنسبة للجزء المعلمي: تقدير ثابت الانحدار $\widehat{\beta_0}$ كان −0.0486 بخطأ معياري قدره 0.1121 والمعنوية المحسوبة له 0.6673 مما يعني عدم معنوية ثابت الانحدار في معادلة التقدير شبه المعلمي. بينما كان تقدير $\widehat{\beta_1}$، معامل المتغير SP، 1.0020 بخطأ معياري قدره 0.2341 والمعنوية المحسوبة له 0.001 مما يعني معنوية هذا المتغير في معادلة الانحدار شبه المعلمي. يلاحظ أيضا أن التقديرات للجزء المعلمي في نموذج الانحدار شبه المعلمي كانت قريبة من التقديرات في نموذج الانحدار المعلمي.

٢. بالنسبة للجزء اللامعلمي: نجد أن تقدير معلمة التمهيد لمتغير DT كان 1.88 وعدد العقد المناسب للانحدار شبه المعلمي له 7 ، في حين نجد أن تقدير معلمة التمهيد لمتغير PV كان 3031 وعدد العقد المناسب للانحدار شبه المعلمي له 3.



شكل (٣): علاقة الانتشار شبه المعلمي بين المتغيرات المستقلة والمتغير التابع في الحالة الثانية

الحالة الثالثة: بافترض أن كلا من متغير SP و DT غير معلمي بينما متغير PV معلمي كالتالي:

جدول (٢-ج): تقديرات نموذج الانحدار شبه المعلمي الحالة الثالثة

العلاقة	$JT \sim f(SP) + f(DT^2) + \exp(PV^2)$			
تقديرات الجزء المعلمي	التقدير Estimate	الخطأ المعياري Standard error	ت المحسوبة t value	المعنوية المحسوبة p-value
الثابت: $\widehat{\beta_0}$	−0.0852	0.1429	−0.5964	0.5547
معاملPV: $\widehat{\beta_3}$	−0.0657	0.0697	−0.9432	0.3519
تقديرات الجزء اللامعلمي	درجات الحرية Df	تقدير معلمة التمهيد $\hat{\lambda}$	تقدير عدد العقد K	
دالة SP	1.773	0.8515	7	
دالة DT^2	1.822	1.1050	7	

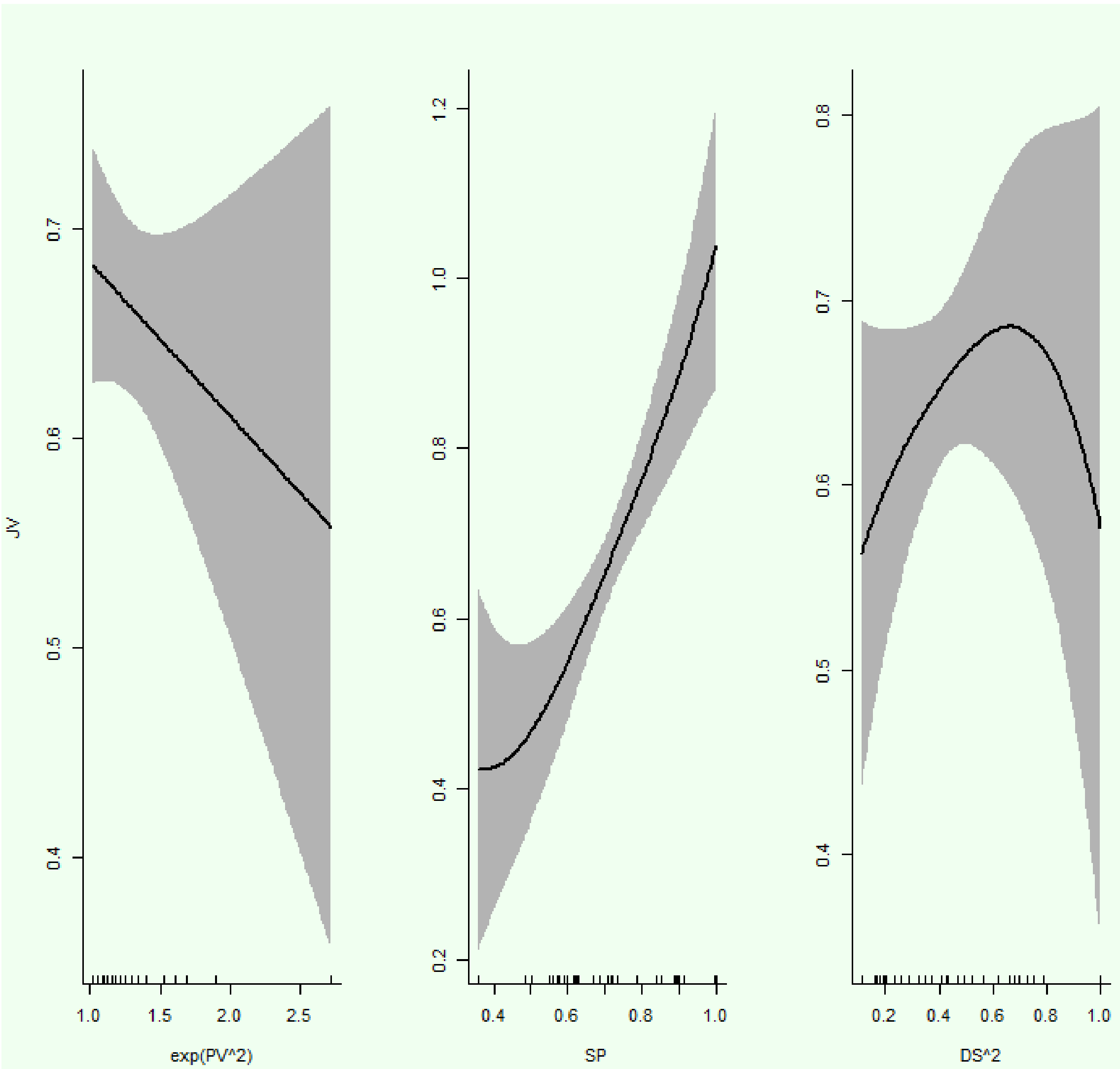
*** معنوي عند مستوي الدلالة الإحصائية ١%

التعليق علي الجدول السابق :-

١. بالنسبة للجزء المعلمي: تقدير ثابت الانحدار $\widehat{\beta_0}$ كان 0.0852− بخطأ معياري قدره 0.1429 والمعنوية المحسوبة له 0.5547 مما يعني عدم معنوية ثابت الانحدار في معادلة التقدير شبه المعلمي. بينما كان تقدير $\widehat{\beta_3}$، معامل المتغير PV، 0.0657− بخطأ معياري قدره 0.0697 والمعنوية المحسوبة له 0.3519 مما يعني عدم معنوية هذا المتغير في معادلة الانحدار شبه المعلمي.

٢. بالنسبة للجزء اللامعلمي: نجد أن تقدير معلمة التمهيد لمتغير DT كان 1.1050 وعدد العقد المناسب للانحدار شبه المعلمي له 7، في حين نجد أن

تقدير معلمة التمهيد لمتغير SP كان 0.8515 وعدد العقد المناسب للانحدار شبه المعلمي له 7.



شكل (٤): علاقة الانتشار شبه المعلمي بين المتغيرات المستقلة والمتغير التابع في الحالة الثالثة

الحالة الرابعة: حالة خاصة من الانحدار شبه المعلمي بافترض ان جميع المتغيرات تأخذ الشكل اللامعلمي وبالتالي يصبح النموذج في الحالة اللامعلمية كالتالى:

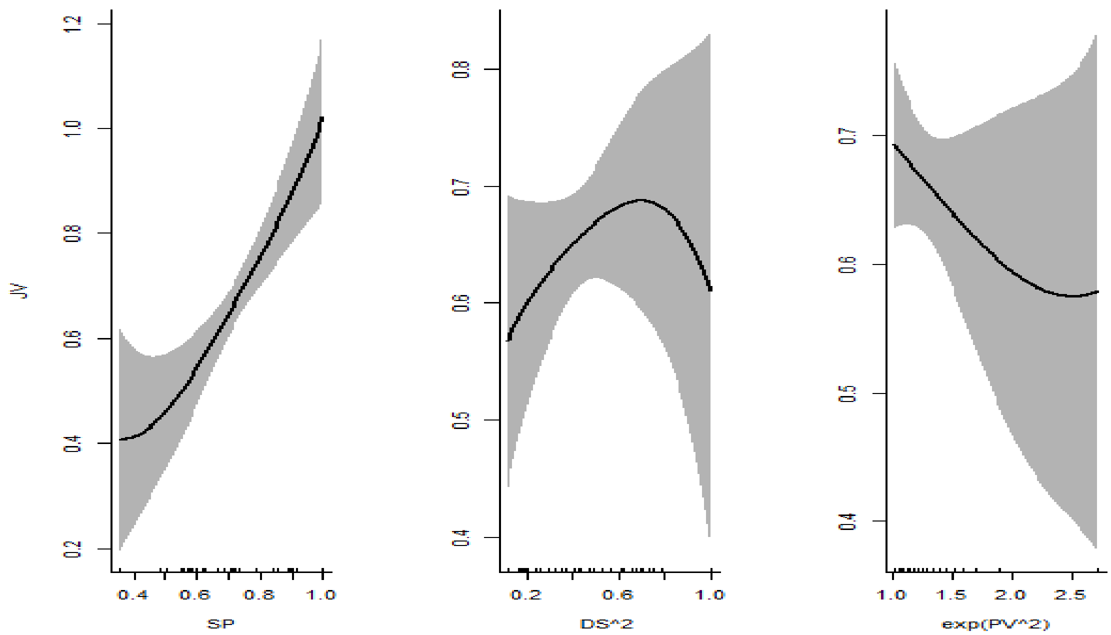
جدول (٢-د): تقديرات نموذج الانحدار اللامعلمي (الحالة الرابعة)

العلاقة	$JT \sim f(SP) + f(DT^2) + f(\exp(PV^2))$		
تقديرات الجزء اللامعلمي	درجات الحرية Df	تقدير معلمة التمهيد $\hat{\lambda}$	تقدير عدد العقد K
دالة SP	1.882	0.8077	7
دالة DT^2	1.825	1.0940	7
دالة $\exp(PV^2)$	1.288	12.0700	3

*** معنوي عند مستوي الدلالة الإحصائية ١%

التعليق علي الجدول السابق :-

١. بالنسبة للجزء المعلمي: لا يوجد تقديرات حيث أن النموذج كله اصبح لا معلمي.
٢. بالنسبة للجزء اللامعلمي: نجد أن تقدير معلمة التمهيد لمتغير SP كان 0.8077 وعدد العقد المناسب للانحدار شبه المعلمي له 7، في حين نجد أن تقدير معلمة التمهيد لمتغير DT كان 1.0940 وعدد العقد المناسب للانحدار شبه المعلمي له 7، بينما تقدير معلمة التمهيد لمتغير PV كان 12.0700 وعدد العقد المناسب للانحدار شبه المعلمي له 3.



شكل (٥): علاقة الانتشار اللامعلمي بين المتغيرات المستقلة والمتغير التابع في الحالة الرابعة

النتائـج:-

يمكن إيجاز النتائج التي تم التوصل إليها في هذا البحث في النقاط التالية:

١) تقديرات الجزء المعلمى فى جميع حالات نموذج الانحدار شبه المعلمى ماعدا الحالة الرابعه كانت قريبه من التقديرات فى نموذج الانحدار المعلمى.

٢) فى الحالة الثانية من نموذج الانحدار شبه المعلمى يتضح من رسم شكل انتشار المتغيرات اللامعلمية أن متغير DT يؤول إلى الشكل التربيعي المتزايد ثم المتناقص مما يدل على أنه الانحدار شبه المعملي ملائم لهذا المتغير .كما يلاحظ ان المتغيرين SP و PV غالباً ما يأخذا الشكل الخطي.

٣) شكل انتشار المتغيرات اللامعلمية فى الحالة الثالثة من نموذج الانحدار شبه المعلمى قريب من الشكل فى الحالة الثانية ولكن متغيرSP يأخذ شكل الانحدار غير الخطي ويظل متغير DT غير خطي ملائم أكثر من الشكل الخطي، في حين نجد أن متغير PV لم يتغير كثيراً.

٤) فى الحالة الخاصة من النموذج شبه المعلمى (النموذج اللامعلمى) يتضح أن كل المتغيرات تأخذ شكل الانحدار غير الخطى.

التوصيات:-

بناءَ على ما تم التوصل إليه البحث من نتائج يوصى بالآتي:

١. استخدام نماذج الانحدار شبه المعلمية للبيانات التي غالباً ما تكون غير مستقرة حيث أن هذه النماذج أعم وأشمل من النماذج المعلمية.

٢. استعمال النموذج الخطى الجزئى فى النماذج شبه المعلمية ووضع نماذج مرنة للتأثيرات المشتركة على متغير الاستجابة.

٣. يوصى البحث بالتطرق الى تقديرات اخرى للنموذج اللامعلمى ومنها تقدير معلمة التمهيد Smoother Parameter.

٤. يمكن تعميم النتائج وذلك بزياده حجم العينة المستخدمه فى الجانب التطبيقى.

المراجع:

(1) Greenberg, E., & Greenberg, E. (2012). Semiparametric Regression. In Introduction to Bayesian Econometrics (pp. 148–168). https://doi.org/10.1017/cbo9781139058414.011

(2) Harezlak, J., Ruppert, D., & Wand, M. P. (2018). Semiparametric regression with R. Springer New York.

(3) Huang, J., Wei, F., & Ma, S. (2012). Semiparametric regression pursuit. Statistica Sinica, 22(4), 1403–1426. https://doi.org/10.5705/ss.2010.298

(4) Li, R., & Liang, H. (2008). Variable selection in semiparametric regression modeling. Annals of Statistics, 36(1), 261–286. https://doi.org/10.1214/009053607000000604

(5) Martins-Filho, C., Mishra, S., & Ullah, A. (2008). A class of improved parametrically guided nonparametric regression estimators. Econometric Reviews, 27(4–6), 542–573. https://doi.org/10.1080/07474930801960444

(6) Naito, K. (2002). Semiparametric regression with multiplicative adjustment. Communications in Statistics - Theory and Methods, 31(12), 2289–2309. https://doi.org/10.1081/STA-120017226

(7) Robinson, P. M. (1988). Root-N-consistent semiparametric regression. Econometrica: Journal of the Econometric Society, 931-954.

(8) Ruppert, D., Wand, M. P., & Carroll, R. J. (2009). Semiparametric regression during 2003–2007. Electronic journal of statistics, 3, 1193.

(9) Ravikumar, P., Lafferty, J., Liu, H., & Wasserman, L. (2009). Sparse additive models. Journal of the Royal Statistical Society. Series B: Statistical Methodology, 71(5), 1009–1030. https://doi.org/10.1111/j.1467-9868.2009.00718.x

(10) Reda Abonazel, M., & Gad, A. A. E. (2018). Robust partial residuals estimation in semiparametric partially linear model. Communications in Statistics: Simulation and Computation, 49(5), 1223-1236. https://doi.org/10.1080/03610918.2018.1494279

(11) Yoshida, T. (2018). Semiparametric method for model structure discovery in additive regression models. Econometrics and Statistics, 5, 124-136. https://doi.org/10.1016/j.ecosta.2017.02.005