

الاندوجينية (Endogeneity)

Foundations • The Valuation Engineer

William Bert Craytor • October 1, 2026

Version 1.2.0

DOI: 10.67330/z9feps66

تعريف صوري. الداخلية هي الحالة التي يتحرك فيها المتغير المستقل --- أي السمة التي يستخدمها الانحدار لتفسير السعر --- بالتوازي مع الخطأ الهيكلي --- أي السعر مطروحاً منه ما تنص عليه العلاقة السببية بأن يكون عليه السعر، وليس بقايا النموذج في عملية الملازمة --- بحيث لا يعكس معامل تقديره التأثير السببي محل الاهتمام. ومن الأسباب الشائعة لذلك: استبعاد إحدى السمات، أو تأثير السعر العكسي على السمة، أو كون المبيعات مجموعة مختارة، أو قياس السمة مع وجود خطأ.

تقريب حدسي. قد يكون المعدل خاطئاً لأن الميزات والسعر يؤثران على بعضهما البعض. تعتمد المدة التي يظل فيها المنزل معروضاً في السوق على السعر المطلوب: فالميزات هي جزئياً نتيجة للسعر، أو شيء آخر يؤدي أيضاً إلى ارتفاع السعر، مثل ارتفاع السوق، والمعدل الذي يُقاس بناءً على تلك المبيعات يخلط بين هذين الاتجاهين معاً. لا يمكن لأي تحليل انحداري أن يفصل بينهما استناداً إلى المبيعات وحدها، سواء كانت الملازمة معدلاً واحداً على النطاق الكامل لميزة ما، أو كما يفعل نموذج MARS (الانحدار التكيفي متعدد المتغيرات، فريدمان 1991)، حيث تُطبق معدلات مختلفة على نطاقات مختلفة من السمة --- فهذا يتطلب مصدرًا خارجيًا للتباين، أو افتراضاً لا يمكن للبيانات المبيعاتية التحقق منه --- لأن التشابك يكمن في السوق، وليس في الحسابات. ما يمكن للمقيمين فعله هو معرفة الطرق الأربع التي يمكن أن تتحرك بها السمة مع ما يستبعده النموذج (الموضحة أدناه)، والتعرف عليها في مجموعة البيانات، وتحديد أي منها ساري المفعول قبل استخدام معدل من تلك المجموعة كما لو كان يقيس ما يدفعه المشتري مقابل تلك السمة.

للقارئ الذي يرغب في معرفة الآلية. يُطلق الإحصائيون على السمة مصطلح «الداخلية» عندما تتغير بالتزامن مع عامل يستبعده النموذج، ويقدم هذا المقال الصيغة الدقيقة لهذا الشرط ويفصل بين مصادره الأربعة. أما الحلول التي تقدمها الأدبيات الإحصائية، ولا سيما المتغيرات الآلية، فهي مخصصة للانحدار العادي ونادراً ما تتوفر في بيانات التقييم؛ وهذا ما يوضحه هذا المقال. ويختلف التشخيص عن الحل في أنه ينطبق على أي طريقة للملازمة.

يحتوي هذا البند على سياق إحصائي ويتضمن أيضاً استدلالاً سببياً، حيث يتناول الحالات التي لا يعكس فيها معامل الانحدار التأثير السببي الذي يريده المقيمين. وهو موجه للمقيمين الذين يفسرون المعاملات أو تعديلات المبيعات المقترنة على أنها دلالات سببية. وعلى غرار البند 008 (انحياز المتغير المحذوف)، يستخدم هذا البند سوق «باسيفيكا» المصنّع، الذي يُدرج الشرط في الانحدار كما يفعل التقييم التقليدي؛ أما الطريقة الخاصة بالمجلة، وهي التقييم القائم على تحليل RCA، فلا تفعل ذلك (البند 007، التوقع الشرطي)، والمثال هنا هو مثال توضيحي، وليس نموذجاً يجب اتباعه.

لا ينبغي النظر إلى «تحيز المتغير المهم» --- وهو موضوع المقالة رقم 008 (تحيز المتغير المهم) --- باعتباره نوعاً مستقلاً من أنواع الأخطاء له معادلة خاصة به، بل باعتباره مثالاً على حالة أكثر عمومية، وهي: الداخلية.

لنتذكر تحليل المدخل رقم 007 (التوقع الشرطي): $P = E[P | Z] + \varepsilon$ ، حيث $E[P | Z]$ هو متوسط سعر المنازل ذات الخصائص Z و ε هو مدى ابتعاد سعر المنزل، سواء للأعلى أو للأسفل، عن هذا المتوسط، بحيث ينطبق بحكم التعريف أن $E[\varepsilon | Z] = 0$. وينطبق هذا على دالة التوقع الشرطي (CEF)، كما ورد في ذلك المدخل) مهما كانت الخصائص التي تحتوي عليها Z . إذا أراد المقيمين بدلاً من ذلك معرفة التأثير السببي --- أي مقدار الزيادة في السعر في حالة وجود إطلالة، مع بقاء جميع العوامل الأخرى المتعلقة بالمنزل ثابتة فعلياً --- فيجب إضافة شرط آخر: يجب ألا تتغير المتغيرات المستقلة مع أي عامل آخر يؤثر على السعر. وعندما يتغير أحدها، فإن هذا المتغير المستقل يكون داخلياً. اكتب العلاقة السببية على النحو التالي: $P = \beta_0 + \beta_1 z_1 + \dots + \beta_k z_k + u$ ، حيث يمثل كل β_j التأثير الذي سيجده تغير بمقدار وحدة واحدة في السمة z_j على السعر. ويُعرف مصطلح الخطأ الخاص بها بالخطأ الهيكلي، ويُكتب على النحو التالي: u ، أي السعر مطروحاً منه ما تنص عليه العلاقة السببية أنه ينبغي أن يكون. وهذا يختلف عن ε ، وهو البعد عن المتوسط الشرطي، الذي يكون متوسطه صفرًا نظرًا لـ Z بحكم البناء؛ وقد يتغير u مع المتغيرات المستقلة، ويشير شرط الداخلية إليه.

يعود أصل هذا المصطلح إلى نمذجة المعادلات المتزامنة، حيث يُطلق على المتغير وصف «الداخلي» (endogenous) إذا كان يتم تحديده داخل النظام بدلاً من أن يتم توفيره من الخارج. أما في الاستخدام الحالي، فيحمل المصطلح معنىً أكثر ميكانيكية: $Cov(z_j, u) \neq 0$ (تقيس التباين المشترك (Cov) كيفية تحرك كميتين معاً) --- يتحرك المتغير المستقل والخطأ الهيكلي معاً، لذا فإن نموذج الملازمة (OLS) --- انحدار المربعات الصغرى العادية --- أو (MARS) لا يستطيع الفصل بين مساهماتهما، ويُعزو جزءاً من حركة الخطأ إلى المتغير المستقل. هناك أربع آليات تسبب ذلك، ومن المهم التمييز بينها لأنها تتطلب استجابات مختلفة.

المتغيرات المهمة. يتم استبعاد متغير يؤثر على السعر ويتحرك بالتوازي مع متغير مُدرج؛ فيصبح جزءاً من u ، ويترتب على ذلك $Cov(z_j, u) \neq 0$. وهذا هو الحال في الإدخال رقم 008 (تحيز المتغير المهم)، حيث تم استبعاد مؤشر «remodel» ويتمثل الحل في

قياس هذه السمة وإدراجها، أو تحديد أكبر وأصغر القيم التي يمكن أن يتخذها هذا التحيز بشكل معقول. **التزامن.** المتغير المستقل لا يرتبط بالسعر فحسب، بل إنه ناتج جزئياً عنه، أو يتم تحديده بالتزامن معه؛ فالعلاقة السببية تسير في كلا الاتجاهين، وبما أن الملاءمة (OLS) أو (MARS) لا تشير إلى أي اتجاه، فإنها تُظهر مزيجاً من الاثنين. والحل هو استخدام متغير وسيط --- وهو متغير يؤثر على المتغير المستقل المعني دون أن يكون له أي ارتباط آخر بالسعر (بالمعنى الاقتصادي القياسي للكلمة، وليس بالمعنى القانوني لوثيقة رسمية) --- أو استخدام تصميم مختلف.

الاختيار. العينة ليست عينة عشوائية مأخوذة من المنازل التي تتناولها النتيجة، وما يحدد إدراج منزل ما في العينة مرتبط بسعره. والحل هو وضع نموذج يوضح كيفية اختيار العينة، أو إعادة تعريف المجموعة التي تتناولها النتيجة.

خطأ القياس. يتم تسجيل السمة بشكل غير دقيق --- فقد يمنحها مفتشان تقيمين مختلفين، أو يتم أخذ مساحة المعيشة من سجل بدلاً من قياسها. يصبح خطأ التسجيل جزءاً من u ويتغير مع القيمة المسجلة، وبالتالي يكون المعامل متحيزاً. في الحالة النموذجية --- خطأ القياس الكلاسيكي، حيث لا علاقة للخطأ بالقيمة الحقيقية أو بأي شيء آخر في النموذج --- ينجرّف المعاملات الخاصة بالسمة التي تم قياسها بشكل خاطئ نحو الصفر. قد لا تكون CQA من هذا النوع: فقد يعتمد خطأها على مدى جودة المنزل فعلياً، وعلى عوامل أخرى يراها المُقيّم. وبالنسبة لمثل هذا الخطأ، لا يكون للانحياز اتجاه عام. والحل هو إجراء قياس أفضل، أو الحصول على قياس ثانٍ مستقل لنفس السمة في نفس المنزل، والذي يمكن بعد ذلك لأي طريقة مصممة لمعالجة خطأ القياس أن تستخدمه لتصحيح التقدير.

تتشترك الآليات الأربع في أمر واحد: إن إضافة المزيد من المبيعات من النوع نفسه لا يزيل الخطأ الذي تسببه. وهذا هو الفرق العملي الأوضح بين الداخلية وعدم الدقة (تشتت التقدير الناتج عن وجود عدد محدود فقط من المبيعات): فعدم الدقة يقلص مع تزايد حجم العينة، أما الداخلية فلا.

قبل اختيار المتغيرات، قم بتسمية التأثير. يعتمد ما إذا كانت السمة تنتمي إلى النموذج أم لا على ما يتم تقديره، وهناك نوعان من السمات يسهل الخلط بينهما. النوع الأول هو سبب لكل من المتغير المستقل والسعر: فالمواقع المرغوبة داخل الحي (مثل شارع مسدود هادئ، أو زاوية ذات إطلالة) تجذب أعمال تجديد باهظة التكلفة وتحقق أسعاراً مرتفعة بعد ذاتها. يُطلق الإحصائيون على هذه السمة اسم العامل المشوش. ويجب إبقاؤها ثابتة، وإلا فإن المعاملات التي يحددها المتغير المستقل ستحمل تأثيرها. والنوع الثاني هو القناة التي يعمل من خلالها المتغير المستقل على السعر: حيث تؤدي عملية إعادة التصميم إلى رفع السعر من خلال تحسين الحالة والجودة. هذه الميزة تُعرف بـ الوسيط، وإبقاؤها ثابتة يزيل من المعاملات المستقلة بالضبط الجزء من تأثيره الذي يمر عبر القناة. لذا فإن إعادة التصميم في البند 008 (تحيز المتغير المحذوف) يطرح سؤالين مختلفين. في التقييم التقليدي، إذا كان السؤال يتعلق بسعر السوق للحالة، فإن الحالة هي المتغير المستقل (التقييم القائم على تحليل التباين المركب (RCA) لا يطرح هذا السؤال أبداً؛ البند 007، التوقع الشرطي، يوضح السبب)، وتجديد المنزل هو مجرد الطريقة التي وصلت بها بعض المنازل إلى حالتها الحالية؛ فهي لا تصنف أي قيمة تتجاوز الحالة والجودة التي أنتجتها، بمجرد تسجيلها بالكامل، ويكون المنزل الذي خضع لإعادة التصميم والمنزل الذي تم الحفاظ عليه جيداً متساويين في الحالة والجودة، وبالتالي لهما نفس القيمة. إذا كان السؤال هو: ما تأثير عملية التجديد على سعر منزل معين؟، فإن الحالة والجودة تكونان متغيرين وسيطين، والنموذج الذي يقيّمها ثابتين يجب على سؤال مختلف --- ما هي قيمة عملية التجديد بصرف النظر عن الحالة والجودة التي أنتجتها، والتي قد تكون ضئيلة، وهي في أي حال كمية مختلفة عن التأثير الإجمالي لعملية التجديد. «تضمن المتغير» هي القاعدة الصحيحة بالنسبة لعامل التشوش، وهي القاعدة الخاطئة بالنسبة للوسيط؛ ولا يمكن اختيار أي منهما حتى يتم تحديد التأثير المطلوب.

أين يصادفه المقيّم. السببية العكسية هي مسألة تتعلق بالموقع. وتُعدّ الداخلية الناتجة عن التزامن مثلاً واضحاً على جودة المدارس، وهي مشكلة حقيقية وليست افتراضية. إذا أجريت تحليلاً انحدارياً للسعر بناءً على تصنيف المدرسة، فسيبدو أن المعاملات توضح المبلغ الذي يدفعه المشترون مقابل المدرسة. لكن العلاقة السببية تسير في الاتجاه المعاكس أيضاً: فالأحياء الأكثر تكلفة تتمتع بقاعدة ضريبية أكبر، وأولياء أمور أكثر انخراطاً وأفضل موارد، ويمكنها جذب المعلمين والاحتفاظ بهم بشكل أفضل. يمكن لجودة المدارس أن ترفع الأسعار، ويمكن للأسعار المرتفعة أن تحسن جودة المدارس. ولا يلتقط المعاملات أيّاً من هذين التأثيرين؛ بل يعكس مزيجاً يعتمد على القوة النسبية للثلاثين في ذلك السوق. ويمكن أن تظهر نفس العلاقة المتبادلة مع معدلات الجريمة، وكثافة المرافق التجارية، وتغطية الأشجار. وفي كل حالة، قد يكون العامل سبباً ونتيجة في آن واحد لمستويات أسعار الأحياء.

تعد عملية إعادة تصميم المطبخ «معالجة»، لكنها لا تُطبق بشكل عشوائي. وفي مصطلحات الاستدلال السببي، تُعرف المعالجة بأنها أي تدخل يُستفسر عن تأثيره، وإعادة تصميم المطبخ تُعدّ أحد هذه التدخلات. وتكمن الصعوبة في تحديد من يتلقى هذه «المعالجة». يقوم العديد من مالكي المنازل بالتجديد عندما يتوقعون أن يتجاوز العائد التكلفة، ويستند هذا التوقع إلى ما يعرفونه عن منزلهم ولا تسجله أي مجموعة --- مثل مخطط طوابق جيد، أو موقع متميز داخل الحي، أو قطعة أرض تدعم الاستثمار. تؤدي تلك الصفات نفسها إلى ارتفاع السعر بشكل مباشر: فهي عوامل تشوش على عملية التجديد، ولا تقيسها مجموعة البيانات. لذا فإن النموذج الذي يتضمن مؤشر التجديد قد عالج إغفال المؤشر وليس العوامل المشوشة. يجمع المعاملات بين تأثير عملية التجديد على السعر والسعر الإضافي --- الذي يتجاوز ما تفسره السمات المسجلة --- للمنازل التي اختار أصحابها تجديدها. وقد أدى إدراج المتغير إلى إزالة الإغفال؛ لكنه لم يجعل المتغير المستقل خارجياً --- أي غير مرتبط بالخطأ الهيكلي.

عينة العقارات المباعة ليست عشوائية. يبيع الناس عندما تكون الظروف مواتية أو الأسعار جذابة، بينما يبيع آخرون لأنهم مضطرون لذلك، لذا قد تختلف العقارات التي تُباع عن تلك التي لا تُباع بطرق لا يراها النموذج؛ وعندما يحدث ذلك، فإن العنصر الخاص بسعر العقار

المباع ليس عشوائياً: فالمجموعة الأولى تدفعه في اتجاه، بينما تدفعه حالات الضائقة المالية والمبيعات القسرية في الاتجاه الآخر. ضمن المبيعات، تشكل المعاملات التي تتم وفقاً لمبدأ الاستقلالية والمبيعات المتعثرة مجموعات فرعية أخرى غير عشوائية. يتم استخلاص كل عقار مقارن في كل شبكة من هذه المبيعات المختارة. ولا يمكن توسيع نطاق أي استنتاج تقييمي ليشمل العقارات غير المباعة إلا بناءً على افتراضات حول كيفية اختيار العقارات المباعة، وهو ما يجب أن يذكره التقرير.

قد يتأثر تقييم حالة العقار بالسعر. فالهدف من التقييم هو قياس حالة المنزل دون أي شيء آخر، ويقوم المُقيّمين بتقييم العقار المعني بناءً على الفحص الميداني، قبل التوصل إلى أي قيمة. أما العقارات المقارنة فهي مسألة أخرى: فقد يتم تقييمها استناداً إلى الصور والملاحظات الواردة في الإعلانات، بعد معرفة أسعار بيعها. وإذا كان السعر المعروف يؤثر على CQA --- حيث يُنظر إلى المنزل الذي يبيع بسعر مرتفع على أنه محفوظ بشكل أفضل قليلاً --- فإن متغير الانحدار يصبح جزئياً دالة للنتيجة. هذه ليست سمة من سمات الحالة نفسها، بل عيب في عملية القياس: خطأ قياس يعتمد على النتيجة، حيث يؤثر السعر على CQA. وهو من النوع غير الكلاسيكي الذي تم وصفه للتو، مع اختلاف واحد. بالنسبة لتأثير من هذا النوع، الذي ينقل جزءاً من انحراف السعر إلى CQA، يكون الاتجاه معروفاً: تتحرك CQA نحو CQA التي تشير إليها سعر البيع، ويتم دفع المعاملات إلى الأعلى (قد تستمر الأخطاء الأخرى في CQA في سحبها في الاتجاه المعاكس). وتتمثل الحماية في الإجراءات المتبعة: تقييم العقارات المقارنة بناءً على الأدلة المتعلقة بحالتها قبل النظر إلى السعر الذي بيعت به، وتسجيل أساس كل CQA في ملف العمل.

لماذا يهم في تماسك الرأي. يتم تحديد ما إذا كان المتغير المستقل متأثراً داخلياً أم لا من خلال الحجة المنطقية، وليس من خلال أي شيء يمكن أن يُظهره الانحدار العادي. ولا يمكن لأي من الفحوصات المعتادة لمدى ملائمة النموذج --- مثل رسم بياني للفروق، أو نسبة التباين السعري المفسر (R^2)، أو اختبار النموذج على المبيعات المستبعدة من عملية الملاءمة (التحقق المتقاطع)، أو اختبار ما إذا كان المعامل يمكن أن يساوي صفراً (اختبار الدلالة الإحصائية) --- لا يمكن لأي منها إثبات أن متغير الترجيع خارجي، لأن الخطأ الهيكلي غير قابل للملاحظة، وكذلك ارتباطه بمتغيرات الترجيع. توجد اختبارات رسمية --- حيث يقارن اختبار هاوسمان التقدير العادي بتقدير تم الحصول عليه بطريقة أخرى --- لكنها تحتاج إلى أداة أو مُقدّر ثانٍ، أي افتراضات إضافية لا يمكن للبيانات نفسها التحقق منها. ما يمكن تقديره دائماً هو تفسير منطقي للآلية: لماذا قد يكون هذا المتغير المستقل مرتبطاً بالخطأ، وفي أي اتجاه، وإلى أي مدى. التقرير المعقول يقدم هذا التفسير بدلاً من الإيحاء بأن المشكلة غير موجودة من خلال عدم ذكرها.

اذكر القيمة الفعلية للمعامل، لا القيمة التي يتمنى المرء أن تكون عليها. فالمعامل الخاص بمتغير تابع داخلي هو رقم حقيقي ومحدد بدقة: وهو الفرق في السعر المُقدّر لكل وحدة من السمة، مع إبقاء السمات المقاسة ثابتة. وهو ليس التأثير السببي. في عمليات البيع الست عشرة لمنازل «باسيفيكا» التي تم استبعاد مؤشر إعادة التصميم منها (المدخل 008، انحياز المتغير المُهمَل)، من الصحيح القول: «باعت هذه المنازل ذات الإطلاقات بمبلغ \$79,178 أكثر، مع إبقاء الخصائص المقاسة ثابتة». أما القول بأن «الإطالة تساهم بمبلغ \$79,178 في القيمة» فهو تأكيد لسبب لا يدعمه التقدير. العبارة الأولى ستصمد أمام الاستجواب؛ أما الثانية فلن تصمد.

التصميم يتفوق على التقدير. حيث يمكن كسر التأثير المتبادل بين إحدى الميزات والسعر من خلال طريقة تنظيم البيانات ---- من خلال مقارنة العقارات قبل وبعد حجب الرؤية، أو باستخدام حدود تتغير فيها التوزيعات المدرسية دون أن يتغير فيها مخزون المساكن، أو من خلال اعتبار عملية تجديد فرضتها أحداث خارجية مثل الحريق أو الفيضان بدلاً من عملية اختارها المالك بنفسه --- لا يزال كل تصميم يعتمد على افتراضات يجب الدفاع عنها (أنه لم يتغير أي شيء آخر في الوقت نفسه، وأن الحدود تفصل فقط بين التوزيع المدرسي)، لكن التصميم الذي يتم الدفاع عن افتراضاته يحقق قابلية للدفاع أكبر من أي مُقدّر يعمل بالبيانات القائمة على الملاحظة. هذا ما تقوله الأدبيات المتعلقة بالمتغيرات الآلية من الناحية العملية، وتكمن أهميتها بالنسبة للتقييم في أن مجموعة العقارات المقارنة المختارة بعناية تشكل في حد ذاتها تصميمًا.

مثال تقييمي محلول. يمكن لمجموعة بيانات «باسيفيكا» أن توضح مفهوم الاندوجينية، لكنها لا تستطيع معالجتها، والسبب وراء عدم قدرة هذه المجموعة على ذلك هو الدرس الحقيقي الذي يقدمه هذا المقال.

أولاً، السؤال. لنفترض أن أحد المُقيّمين سأل عما إذا كان من الممكن استرداد تكلفة مشروع تجديد المطبخ المقترح بقيمة \$45,000 لمنزل معين عند بيعه. هذا سؤال يتعلق بمنزل واحد وسعرين لا يمكن أن يتواجدوا في آن واحد: سعره بعد التجديد وسعره في حالته الحالية، $P_i(1) - P_i(0)$ ، حيث $P_i(1)$ هو سعر المنزل i في حالة التجديد و $P_i(0)$ هو سعره في حالة عدم التجديد. وهذا السؤال لا يتعلق بسعر السوق للحالة، وإبقاء الحالة ثابتة أثناء الإجابة عليه سيكون الخطأ الوسيط الموصوف أعلاه.

تحتوي العقارات الست عشرة للمدخل رقم 008 (تحيز المتغير المحذوف) على مؤشر «إعادة التصميم» الذي يبلغ معامل \$60,000: وهو الفرق في السعر الذي تربطه البيانات المركبة بعملية إعادة التصميم --- \$60,000 أكثر من منزل يتمتع بنفس المساحة الإجمالية القابلة للتأجير (GLA) والمساحة الأرضية والإطالة ودرجة CQA. هذا المبلغ \$60,000 ليس سعرًا لعملية إعادة التصميم نفسها؛ بل يمثل كل ما أضافته عملية إعادة التصميم ولم تسجله الخصائص الأخرى. في مجموعة البيانات المركبة تلك، تم تحديد المنازل التي تم إعادة تصميمها من قبل المؤلف. مؤشر إعادة التصميم خارجي بطبيعته، ولا يرتبط بأي شيء متبقي في الخطأ لأنه لم يتبق شيء في الخطأ، ولذا فإن \$60,000 يمثل، في تلك المجموعة من البيانات، ما تضيفه إعادة التصميم إلى السعر مع إبقاء CQA ثابتة؛ أما إعادة التصميم التي رفعت CQA أيضاً فستضيف أكثر من ذلك.

لنفترض الآن أن نفس العقارات الستة عشر جاءت من سوق حقيقي. يختار الملاك إجراء أعمال تجديد، ولنفترض --- كما هو معقول --- أن مالكي المنازل التي تتمتع بصفات قوية غير مرئية (لنسمي هذا المكون من السعر u) هم أكثر ميلاً إلى إجراء أعمال تجديد مقارنة بمالكي

المنازل المماثلة من نواحٍ أخرى. الانحدار الذي يتضمن مؤشر إعادة التصميم الآن يتقارب، مع تراكم المبيعات، ليس إلى \$60,000 بل إلى

$$\hat{\beta}_{\text{remodel}} \xrightarrow{p} \underbrace{60,000}_{\text{that brought remodel the what record not do features other the}} + \underbrace{\frac{\text{Cov}(\tilde{r}, u)}{\text{Var}(\tilde{r})}}_{\text{remodeling into selected houses the of price extra}},$$

حيث $\tilde{r}$ هو الجزء من مؤشر إعادة التصميم الذي لا تتنبأ به الخصائص المضمنة، و Var (التباين) يقيس مدى تباينه، والسهم المسمى p يعني «يقترّب أكثر فأكثر، مع تراكم المبيعات». إذا كانت المنازل التي خضعت للتجديد تحمل، في المتوسط، \$25,000 نسبة أكبر من u مقارنة بالمنازل غير المُجددة المماثلة لها بشكل ملحوظ، فإن المعاملات تتقارب إلى ما يقارب \$85,000، ولا توجد أي كمية من البيانات الإضافية من نفس السوق تقلل من هذا الفائض. مؤشر التجديد، الذي درس إدخاله عند الحذف 008 (تحيز المتغير المحذوف)، يتم قياسه الآن وإدراجه، ومع ذلك لا يزال المعامل لا يمثل التأثير السببي: فما زال هناك ما تم استبعاده هو u ، وهو ما لا تسجله أي مجموعة بيانات.

هذا التمييز له تأثير ملموس في الواقع العملي. فالمعاملات تصف، في المتوسط، المنازل التي خضعت لأعمال تجديد، وهي المنازل التي رأى أصحابها أن العائد يستحق العناء. أما الموضوع فهو منزل معين لم يتخذ مالكه قراراً بعد، ولا تنطبق عليه المعاملات إلا إذا كان يشبه تلك المنازل من حيث ما كان يعرفه مالكها --- وهو بالضبط ما لا يستطيع النموذج رؤيته. أما المُقيّمين الذين يجيبون على سؤال العائد من النموذج \$85,000 فيبالغون في تقدير العائد؛ أما المُقيّم الذي يجيب عليه استناداً إلى نموذج قام أيضاً بقياس جميع الشروط والجودة التي نتجت عن عملية التجديد، واعتبرها ثابتة، فإنه يقلل من العائد، ليقصر على ما تضيفه عملية التجديد بخلاف تلك الشروط والجودة، وهو ما يُعد قليلاً أو لا شيء وفقاً لمبدأ البند 008 (تحيز المتغير المحذوف). كلاهما فسّر الفرق الشرطي في السعر المرتبط بالتجديد على أنه تأثير للتجديد نفسه. وبالتالي، فإن السؤال التشخيصي الذي يجب طرحه بشأن أي معاملات يُراد تفسيرها من منظور سببي ليس «هل النموذج محدد بشكل جيد؟» بل «ما هو التأثير المطلوب، وكيف وصل هذا المتغير إلى القيم التي اتخذها؟» وإذا كانت الإجابة تنطوي على اختيار اتخذها شخص يمكنه رؤية أمور لا يستطيع النموذج رؤيتها، فإن المعاملات تصف هذا الاختيار بقدر ما تصف السمة نفسها.

References.

- Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, 19(1), 1–67.
<https://doi.org/10.1214/aos/1176347963>
- Haavelmo, T. (1943). The Statistical Implications of a System of Simultaneous Equations. *Econometrica*, 11(1), 1–12.
- Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6), 1251–1271.
- Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge University Press.
- Wright, P. G. (1928). *The Tariff on Animal and Vegetable Oils*. Macmillan.

identification. parameter ;expectation conditional ;bias variable omitted Cross-references:

Formalization Prolog Addendum:

knowledge Prolog journal's the in entry this formalizes (009-endogeneity.pl) companion machine-readable following The article. this alongside galley browsable a as published also is It base.

later-or-0.3-AGPL :Identifier-License-SPDX %
 Engineer Valuation The by Published .Craytor Bert William 2026 (C) Copyright %
 can you :software free is program This .(Net Vista Pacific) Journal %
 General Affero GNU the of terms the under it modify or/and it redistribute %
 3 version either ,Foundation Software Free the by published as License Public %
 distributed is It .version later any (option your at) or ,License the of %
 or MERCHANTABILITY of warranty implied the even without ,WARRANTY ANY WITHOUT %
 or it beside file COPYING the see :PURPOSE PARTICULAR A FOR FITNESS %
 .<html.0.3-agpl/licenses/org.gnu.www//:https> %

pl.endogeneity-009 %%
 %%

009 Entry ,2 Issue ,1 .Vol Foundations to companion Prolog %%

```

Endogeneity :Title %%
Craytor Bert :Author %%
endogeneity :Concept %%
0.2.1 :Version %%
%%

chosen are variables the before named is estimand the :0.1.v1 %%
randomly not is that treatment a is remodel a ;(mediator .vs confounder) %%
data constructed the difference price the is 000,60 its ;assigned %%
rating a ;case classical the in only attenuates error measurement ;remodel a to attach %%
.error measurement dependent–outcome is price known the by contaminated %%

so ,time run at loaded not is (dictionary/tools/Issues) dictionary The %
The .Scrier and Prolog–SWI ,Trealla in unchanged runs file the that %
.linter dictionary the by read ,facts plain are below declarations %
.('1.2026–dictionary')dictionary_release_pinned

,omitted_variable_bias ,endogeneity ])uses_terms
,hedonic_regression ,conditional_expectation
,comp_selection ,residual ,latent_variable
,adjustment ,paired_sales_analysis
.([ sales_comparison_approach
.([2/formalized_by ,2/introduced_by ,2/related ,2/depends_on])uses_relations

.(endogeneity)article_concept
.((2026)year ,(2)number ,(1)volume)article_issue
.('009.2026.foundations')article_id

.(instrument)local_term
.(simultaneity)local_term
.(selection)local_term
.(causal_effect)local_term
.(conditional_association)local_term
.(treatment)local_term
.(confounder)local_term
.(mediator)local_term
.(estimand)local_term

,endogeneity)defining_condition
.("term error structural s'model the with correlated is regressor The"

,omitted_variables)mechanism
.("error the inside sits now it ;regressor included an with correlates and excluded is variable affecting–price A"
,simultaneity)mechanism
.("directions both of blend a returns squares Least .outcome the ,with determined jointly or ,by caused partly is regressor The"
,selection)mechanism
.("sample the in being on conditional is relationship estimated the so ,outcome the to related is membership Sample"
,measurement_error)mechanism
classical Under .coefficient the biases and error structural the enters error recording the ;error with recorded is regressor The"
classical–non under ;zero toward attenuated is coefficient the (else everything of and value true the of independent) error ⇔
.("follows direction general no error ⇔

.(attenuates_toward_zero ,classical)measurement_error_kind
.(no_general_direction ,non_classical)measurement_error_kind
.(toward_the_outcome ,outcome_dependent)measurement_error_kind
.(non_classical ,condition_rating)measurement_error_of
.(outcome_dependent ,condition_rating_after_price)measurement_error_of

.("bias the bound or ,variable the include and Measure" ,omitted_variables)remedy
.("loop the breaks that design a or ,instrument an Find" ,simultaneity)remedy

```

.("about is conclusion the population the redefine or ,selection the Model" ,selection)remedy
 .("characteristic same the of measurement independent second a obtain or ,better Measure" ,measurement_error)remedy

 .(omitted_variable_bias ,endogeneity)generalizes
 .(endogeneity ,omitted_variable_bias)special_case_of

 .(imprecision)shrinks_with_sample_size

 -: (endogeneity)distinguishable_from_imprecision
 .(endogeneity)shrinks_with_sample_size +\

 .("price the and regressor the both of cause A" ,confounder)role
 .("price the on acts regressor the which through channel The" ,mediator)role
 .("coefficient s'regressor the on rides otherwise effect Its" ,include ,confounder)include_rule
 .("effect total the from quantity different a ,effect direct the only leaves fixed it Holding" ,exclude ,mediator)include_rule

 .(remodel)treatment
 suitability ,location-micro ,plan floor % .(price ,remodel ,unobserved_quality_u)confounder_of
 .(price ,remodel ,condition)mediator_of
 .(price ,remodel ,quality)mediator_of

 ,condition_rate)estimand
 to came houses some how is remodel the ;regressor the as condition with ,condition for rate s'market the :appraisal Traditional"
 .("this asks never appraisal based-RCA own its of value no earns and condition their ↔
 ,remodel_total_effect)estimand
 .("fixed held be not must and mediators are quality and condition :price s'house particular a to does remodel a What"
 ,remodel_direct_effect)estimand
 be may ;estimates fixed those holding model a what :produced it quality and condition the from apart effect s'remodel The"
 .("effect total the not ;small ↔

 .(price ,remodel ,V)mediator_of -: (exclude ,V ,remodel_total_effect)variable_choice
 .(price ,remodel ,V)confounder_of -: (include ,V ,remodel_total_effect)variable_choice
 .(regressor ,condition ,condition_rate)variable_choice

 ,simultaneity ,school_quality)encounter
 then coefficient the ;resources parental and base tax via quality school raise can prices and prices raise can quality School"
 .("effect neither recovers ↔
 ,omitted_variables ,renovation_choice)encounter
 u qualities on ,it justify to return the expect they when remodel owners :assigned randomly not is that treatment a is remodel A"
 omission the cures indicator remodel the including so ,measure not does dataset the confounder a is u .price raise also that ↔
 .("confounding the not ,indicator the of ↔
 ,selection ,comparable_set)encounter
 ;other the sellers forced and distressed ,way one price the push moment their choose who sellers :selected are properties Sold"
 .("see not does model the what to related is selection the when matters this ↔
 ,measurement_error ,condition_rating)encounter
 the ,rating the nudges price known the if ;known are prices sale their after photographs from rated be may Comparables"
 the ,score the into deviation price the of part carries that nudge a For .outcome the of function a partly is regressor ↔
 .("upward pushed is coefficient condition ↔
 ,condition_rating)protection
 .("rating each of basis the record ;price the at looking before condition their of evidence the from comparables the Rate"

 .(school_quality)bidirectional_with_price
 .(crime_rate)bidirectional_with_price
 .(retail_amenity_density)bidirectional_with_price
 .(tree_canopy)bidirectional_with_price

 .(omitted_variables ,measuring_the_variable)cures

 .(unobserved_quality_u ,renovation_choice)omitted_component

```

.(remodel_numeric)measurable

-: (remodel_numeric)remains_endogenous_after_inclusion
,(_ ,omitted_variables ,renovation_choice)encounter
,(U ,renovation_choice)omitted_component
.(U)measurable +\

,what_the_work_produced ,remodel_numeric)coefficient_component
it ;fixed score condition the and view ,lot ,GLA holding ,remodel a to attach data constructed the difference price the :000,60"
.("record not do features other the that brought remodel the whatever for stands ⇔
,value_of_selected_properties ,remodel_numeric)coefficient_component
under negative ,selection positive under positive ;remodel to chose owners whose properties of quality unobserved The"
.("selection negative ⇔

-: (Limit ,SelectionDifference)coefficient_limit
.SelectionDifference + 60000 is Limit

,renovation_payback_question)consequence
not need it ;remodeled were that properties the among association conditional the is data observational from coefficient A"
.(".(0)P_i - (1)P_i ,subject particular the remodeling of effect causal the equal ⇔
,two_ways_wrong)consequence
quality and condition holds that model a reading ;selection positive under return the overstates coefficient selected the Reading"
.("alone effect direct the to ,it understates fixed ⇔

,not_from_the_regression_itself)defensibility_consequence
need (Hausman) tests formal ;exogeneity establish cannot tests significance and validation-cross ,squared-R ,plots Residual"
.("check cannot data the assumptions .e.i ,estimator second a or instrument an ⇔
,state_what_it_is)defensibility_consequence
asserts contribution causal a as it reporting ,association conditional defined-well a is regressor endogenous an on coefficient A"
.("supports estimate the than more ⇔
,design_beats_estimation)defensibility_consequence
is set comparable chosen-well a ;afterward applied estimator any than more does assembled are data how in loop the Breaking"
.("design a itself ⇔

.(before_after_view_obstruction)design
.(boundary_discontinuity_school_assignment)design
.(externally_forced_renovation)design

)diagnostic_question
model the what observe could chooser a If ?took it values the take to come variable this did how and ,wanted is effect What"
.("characteristic the as much as choice the describes coefficient the ,cannot ⇔

;request on predicates list the and 2/forall loads Prolog Scryer %
below loads guarded The .in built them have Trealla and Prolog-SWI %
.three all in unchanged run file the let %
.((true ,_ ,((lists)library)use_module)catch)initialization -:
.((true ,_ ,((iso_ext)library)use_module)catch)initialization -:

```