

تحديد المعالم (Parameter Identification)

Foundations • The Valuation Engineer

William Bert Craytor • October 1, 2026

Version 1.2.0

DOI: 10.67330/xjszh949

تعريف صوري. يتم تحديد المعلمة عندما يتم تحديدها بشكل فريد من خلال التوزيع المشترك للبيانات القابلة للرصد (عدد المرات التي تحدث فيها كل تركيبة من السمات والسعر ضمن جميع عمليات البيع هذه)، في ضوء افتراضات النموذج ومواصفاته. ويأتي التحديد منطقياً قبل التقدير، وهو أمر متميز عنه.

تقريب حدسي. هناك بعض المعدلات التي لا يمكن أن توفرها لك المبيعات على الإطلاق، ولن تغير أي طريقة من ذلك. إذا كانت كل المنازل التي خضعت للتجديد في السوق في حالة ممتازة، وكانت كل المنازل التي في حالة ممتازة قد خضعت للتجديد، فإن المبيعات تُظهر القيمة الإجمالية لهاتين الفئتين معاً، ولا تُظهر أي شيء عن كل منهما على حدة؛ والانحدار الذي يُبلغ عن معدل لكل منهما --- كما هو الحال عندما تتوافق هاتان الفئتان معاً في الغالب --- يُبلغ عن انقسام لم يحدث في السوق إلا نادراً. إذا كان لكل عقار مقارن إطلالة أفضل من موضوع الدراسة، فإن معدل الإطلالة يُقرأ بما يتجاوز نهاية البيانات. وإذا كانت المبيعات الوحيدة التي تُظهر تغيراً في ظروف السوق بمرور الوقت هي أيضاً المبيعات الوحيدة التي تُظهر تغيراً في معدلات الفائدة على الرهن العقاري، فلا يمكن التمييز بين الاثنين. هذه أسئلة حول محتوى المبيعات، تُطرح قبل أي عملية ملاءمة، وتكون الإجابة هي نفسها سواء كانت الطريقة ثلاثية معدلاً واحداً عبر نطاق السمة أو معدلات مختلفة عبر فترات مختلفة منها، كما يفعل نظام MARS (الخطوط المنحنية للانحدار التكيفي متعدد المتغيرات، فريدمان 1991، طريقة الانحدار الرئيسية التي تستخدمها المجلة): لا يمكن لطريقة MARS تحديد مكان تغير السعر إلا عندما تحتوي المبيعات على التباين الذي يوضح ذلك. ولذلك، فإن السؤال الأول الذي يطرحه المُقيمين ليس «ما هو السعر» بل «هل يمكن لهذه المبيعات أن تخبرني».

للقارئ الذي يرغب في فهم الآلية. ويُعرف هذا إحصائياً باسم «التحديد»: أي ما إذا كان إجمالي المبيعات، عند قياسه دون أخطاء، سيحدد المعلمة من الأساس. تُظهر طريقة المربعات الصغرى العادية الحالة المتطرفة من خلال استبعاد إحدى السمتين اللتين تتحركان معاً تماماً، وتخفي الحالة الخطيرة، حيث تتحركان معاً تقريباً ويستند الرقمان إلى نفس المبيعات القليلة. يتناول هذا المقال كلا الحالتين، بالإضافة إلى حالات الاستقرار وتأثيرات الوقت المربكة التي تواجهها الشبكة يومياً.

يتناول هذا المقال موضوع الإحصاء: فهو يبحث في ما قد تكشفه مجموعة من البيانات من حيث المبدأ قبل استخدام أي مُقدّر (طريقة لحساب المعدلات استناداً إلى المبيعات، مثل OLS أو MARS) وهو موجه إلى أولئك الذين يعملون على تجميع مجموعات بيانات قابلة للمقارنة، وكذلك إلى أولئك الذين يقومون بتطبيق نماذج الانحدار.

المعلمة هنا هي رقم يتعين على النموذج توفيره --- معدل أو المعاملات --- يُحسب على أساسه التعديل. والسؤال الذي يطرحه التقدير هو: بناءً على العينة، ما الذي يجب أن اعتبره أفضل تقدير للمعلمة؟ أما تحديد المعلمة فيطرح سؤالاً مسبقاً: لو كان لدي المجموعة الكاملة --- أي كل معاملة ستحدث في هذا السوق، مقياساً بدقة --- فهل سيتم تحديد المعلمة أصلاً؟

إذا كانت الإجابة «لا»، فإن المعلمة غير محددة ولا يمكن تصحيح ذلك بأي إجراء لاحق. ولن يُحدث المُقدّر المُحسن أي فرق، ولن تفيد عينة أكبر من نفس نوع المبيعات، ولن يساعد النموذج الأكثر تعقيداً أيضاً، ما لم يصف افتراضاً لا يمكن للبيانات اختباره (الحالة السببية أدناه)، بل وقد يخفي المشكلة. إن الفشل في تحديد المعلمة يرجع إلى السؤال والبيانات المعتمدة عليه --- ما أنتجه السوق وما جمعه المُقيمين --- وليس إلى طريقة المحلل.

يكون فشل التمييز أسهل في ملاحظته عندما تظهر خاصيتان معاً دائماً. تخيل أن كل عقار تم تجديده حصل أيضاً على تقييم «ممتاز» (00.9)، وأن كل عقار حصل على تقييم «ممتاز» قد تم تجديده أيضاً؛ في هذه الحالة، تكون مجموعتي العقارات التي تم تجديدها وتلك التي حصلت على تقييم «ممتاز» متطابقتين. يدفع السوق سعراً أعلى لتلك المجموعة --- أي لمنزل ممتاز تم تجديده مقارنةً بمنزل جيد لم يتم تجديده. وبغض النظر عن عدد المبيعات التي نلاحظها، لا يمكننا تقسيم فرق السعر هذا بين الحالة المسجلة وما أضافه التجديد من ميزات لا تسجلها الخصائص الأخرى، لأن في هذا السوق لا يتغير أحدهما أبداً دون الآخر. نحن نعرف المجموع؛ لكننا لا نعرف الأجزاء.

وعندما توضح طريقة المربعات الصغرى ذلك في حالة التوازي التام (حيث تكون إحدى السمات دالة خطية تامة للسمات الأخرى)، فإن السبب في ذلك هو أن مصفوفة التصميم --- أي الجدول المكون من أعمدة المتغيرات المفسرة (المتغيرات التفسيرية) التي يستخدمها المُقدّر --- تعاني من نقص في الرتبة: أحد أعمدها هو تركيبة دقيقة من الأعمدة الأخرى. ولذلك، يتجاهل البرنامج أحد المعاملات أو يعرض قيمة مفقودة. وهذا مجرد إجراء مجاملة. أما الحالات الخطيرة فهي الحالات القريبة --- تلك التي يتغير فيها المتغيران معاً في معظم الأحيان --- نظراً لأن المعلومات في مثل هذه الحالات قابلة للتحديد من الناحية الفنية، فيقدم البرنامج رقمين، وكلاهما يستند إلى العدد القليل من الملاحظات التي تخرج عن النمط. يتم تحديدها؛ ولكن مع وجود عدد قليل فقط من المبيعات التي تخرج عن النمط، فإنها تكون غير دقيقة لدرجة أنه من

الناحية العملية لا يمكن للبيانات التمييز بينها. وهنا يتشابه مشكلتي التحديد والدقة. وهذا هو ما تعنيه التعددية الخطية (السمات التي تتحرك معًا بشكل متطابق تقريبًا) من منظور التحديد.

قد نقشل عملية تحديد العلاقة حتى في حالة عدم وجود أي ترابط خطي على الإطلاق، عندما تكون المتغيرات المعنية ذات علاقة سببية وتكون البيانات قائمة على الملاحظة، ما لم يتم إضافة افتراض لا يمكن للبيانات اختياره. والمثال الوارد في البند 009 (الداخلية)، حيث جاءت نفس المبيعات الستة عشر من سوق فعلي، هو من هذا النوع: يمكن تحديد العلاقة بين إعادة البناء والسعر --- نظرًا لأنها سمة من سمات التوزيع المشترك، وستقوم عينة كبيرة بتقديرها بدقة --- في حين أن التأثير السببي لعملية التجديد لا يمكن تحديده، لأن التوزيع القابل للملاحظة لا يحتوي على أي معلومات تسمح بالتمييز بين القيمة التي أضافتها عملية التجديد والقيمة الإضافية التي كانت تتمتع بها المنازل أصلاً، والتي دفعت مالكيها إلى تنفيذ العمل. لتحديد المعلمة السببية، من الضروري وضع افتراض لا يمكن اختياره في حد ذاته استنادًا إلى البيانات، وبالتالي فإن موثوقية الاستنتاج تعتمد على مصداقية ذلك الافتراض.

أين يصادف المقيم. شرط التدرج. القاعدة التي تنص على ضرورة تدرج الموضوع --- أي تضمين العقارات المقارنة أعلى وأدنى منه بالنسبة لكل سمة مهمة --- تُعرض باعتبارها قاعدة عامة، وهي تجيب على سؤالين يجدر الفصل بينهما. السؤال الأول يتعلق بالتحديد: هل تحتوي هذه المبيعات على تباين في السمة، بحيث تختلف عن غيرها؟ إذا كانت كل عقار مقارن يتمتع بنفس الإطلالة، فلا يمكن الحصول على سعر خاص بالإطلالة منها. والثاني هو الدعم: هل يقع الموضوع قيد الدراسة ضمن النطاق الذي تم تحديد المعدل بناءً عليه؟ إذا كان لكل عقار مقارن إطلالة أفضل من الموضوع، فقد يكون هناك معدل بالفعل --- مستمد من هذه المبيعات أو من مجموعة بيانات أكبر --- لكن تطبيقه على الموضوع يعني الاستقراء خارج نطاق البيانات. يحمي تحديد النطاق العامل الثاني؛ ولا يحمي العامل الأول سوى هذا التباين. يمكن أن يظل الموضوع المحدد النطاق ضمن منطقة لا تختلف فيها خاصيتان أبدًا، ويمكن أن يحصل الموضوع غير المحدد النطاق على سعر تم تحديده في مكان آخر، وهو الحل المتمثل في «مجموعة البيانات الأكبر» الذي يسمح به قسم للدفاع.

ومن الأمثلة على الخصائص التي لا تتباين بشكل مستقل داخل سوق فرعي: منطقة سكنية تكون فيها كل منزل على الجانب المطل على الشارع ذو المساحة الأكبر؛ فلا يمكن التعامل مع الإطلالة والحجم كعاملين منفصلين في تلك المنطقة. والمقيمون الذين يجري تعديلات منفصلة لكل منهما يقدم فعليًا رقمين لا يمكن لبيانات المنطقة التمييز بينهما، وتوزيع المبلغ الإجمالي بينهما يعتمد على تقدير المقيم أو على سوق آخر. وقد يكون ذلك مناسبًا، لكنه يختلف عن مصطلح «مستمد من السوق».

العمر، والحالة، وفترة البناء. غالبًا ما ترتبط هذه العوامل ارتباطًا وثيقًا. تميل المنازل الأقدم إلى أن تكون في حالة أسوأ، وقد شُيدت وفقًا لمعايير مختلفة. ويشير استخدام شبكة تعديل تحتوي على خطوط منفصلة للعمر والحالة إلى أن السوق يقيم هذين العاملين بشكل منفصل، وأن البيانات قادرة على التمييز بينهما. لكن مسألة ما إذا كانت البيانات المتاحة قادرة على التمييز بينهما هي أمر يجب التحقق منه، لا الافتراض به، وغالبًا ما نقشل البيانات في اجتياز هذا الاختبار.

الزمن وظروف السوق. يختبئ هنا نوعان مختلفان من الأخطاء. إذا كانت الشبكة تحتوي على خط لظروف السوق ولم تحتوي على خط للأمر الأخرى التي تغيرت خلال تلك الفترة --- مثل أسعار الفائدة، والمخزون، وصاحب العمل الجديد --- فإن معامل الزمن يتم تحديده، لكنه معامل مركب ولا يمثل تأثير ظروف السوق وحدها: يُحمل التغيير بأكمله على التسمية الوحيدة المتاحة، وهي حالة المتغير المُهمَل في الإدخال 008 (تحيز المتغير المُهمَل). إذا كان كلا الخططين موجودين، لكن المبيعات الوحيدة التي تُظهر تغييرًا في أسعار الفائدة هي المبيعات التي تُظهر تغييرًا في ظروف السوق، فإن الفصل بينهما لا يتم تحديده، وهذه هي حالة هذا البند.

لماذا يهم في تماسك الرأي. عند طرح سؤال التحديد، يجب طرحه قبل سؤال التقدير لأن الترتيب مهم. فإذا نظر المقيمون أولاً في السؤال «ما الذي تسمح لي بياناتي بتمييزه؟» ثم بعد ذلك فقط في السؤال «ما هي أفضل طريقة لتقديره؟»، فإن التعديلات التي يتم التوصل إليها ستكون تلك التي تدعمها البيانات. وعلى العكس من ذلك، إذا طرح سؤال التحديد في النهاية، فستكون النتيجة أرقامًا موثوقة لكميات لم يكن بوسع أحد قياسها.

يجب الإبلاغ عن أي متغير لم يتم تحديده على هذا النحو. وعندما تحدد البيانات تأثيرًا مركبًا دون تحديد كيفية تقسيمه، فإن النهج المعقول لا يتمثل في حذف التعديل، بل في تقديم وصف صادق لكيفية تقسيمه: أي أن العقارات المقارنة لا تستطيع الفصل بين هذين التأثيرين، وأن التقسيم المستخدم يستند إلى مجموعة بيانات أكبر أو إلى تقدير، وما هو التأثير المركب، لأن التأثير المركب هو ما تحددته البيانات فعليًا. وبذلك يمكن للقارئ أن يحكم على هذا التقسيم. في المقابل، لا يمكن للقارئ أن يحكم على رقمين يُقدَّمان على أنهما مستمدان من السوق، في حين أن مجموعهما فقط هو المستمد من السوق في الواقع.

يُعد التحديد نقطة التقاء المنهجية ومعايير الممارسة المهنية الموحدة للتقييم (USPAP). وتُعرف معايير الممارسة المهنية الموحدة للتقييم (USPAP) بأنها المعايير الموحدة لممارسة التقييم المهني. وهي تُلزم المقيمين بجمع جميع المعلومات اللازمة للحصول على نتائج موثوقة للمهمة والتحقق منها وتحليلها (القاعدة رقم 1-4 من المعايير). تُركز إرشادات المقرضين بشكل مباشر على التعديلات: حيث يتطلب دليل البيع الصادر عن فاني ماي (Fannie Mae) إجراء تعديلات تعكس رد فعل السوق تجاه الاختلافات بين العقارات. ولا تنص أي من القاعدتين على ضرورة تحديد كل خط في الشبكة بشكل منفصل عن العقارات المقارنة المعروضة أمام القارئ؛ فقد تنبئ ردود فعل السوق من عمليات بيع مقترنة خارج الشبكة، أو من استخلاص أوسع نطاقًا، أو من حكم مُعلن، وهو الحل المذكور أعلاه. وفقًا لتفسير المجلة، فإن ما تفعله القواعد هو جعل مسألة التحديد إلزامية: يجب أن يكون المقيمون الذين يدعون وجود دعم من السوق لتعديل ما قادرين على تحديد البيانات التي تدل عليه. وبصياغة السؤال بهذه الطريقة، يصبح قابلاً للتحقق --- حيث يمكن فحص البيانات التي تم الاعتماد عليها ومعرفة ما إذا كانت

تحتوي على التباين --- دون تحويل الاختبار الرياضي للتحديد إلى اختبار للامتثال. يكمن الحل في التصميم، وليس في الحساب. فإذا لم يتم تحديد أحد المعلمات، فإن الحل يكمن في تغيير البيانات المستخدمة --- أي توسيع النطاق الجغرافي، أو إطالة الفترة الزمنية، أو البحث عن العقارات التي تخرج عن النمط --- أو تعديل الفرضية. فلا يمكن لأي مُقدّر استخلاص معلومات لا تحتوي عليها البيانات.

مثال تقني محلول. لنأخذ مجموعة بيانات «باسيفيكا» (Pacifica) المكونة من ستة عشر عقارًا، وبالتحديد السجل رقم 008 (تحيز المتغير المحذوف). ويُقصد بـ«الحالة» هنا درجة الحالة التي تحددها المجلة: وهي تقدير لحالة المنزل مقارنةً بحالة المنازل المحيطة به في منطقة السوق المحددة. وتستند هذه النتيجة إلى فهم المُقيمين لتلك المنطقة، ولا ينبغي أن تتعارض بشكل جوهري مع فهم المُقيمين الأقران العاملين في نفس المنطقة. وتتراوح النتيجة بين 00.0 و 00.10: وهي تمثل النسبة المئوية للمنازل في المنطقة التي تكون في حالة أسوأ، مقسومة على عشرة. تسجل مجموعة البيانات ثلاثة مستويات: متوسط (00.5)، جيد (00.7)، وممتاز (00.9)، وتدفع البيانات المُعدة \$20,000 لكل نقطة. وكما في المدخلتين 008 (تحيز المتغير المحذوف) و 009 (الداخلية)، فإن إدراج «الحالة» في معادلة الانحدار هو ما تفعله طرق التقييم التقليدية، وليس طريقة المجلة، أي التقييم القائم على تحليل RCA (المُدخلة 007، التوقع الشرطي). قم بتقييم مجموعة البيانات لتشمل العقارات العشرة التي تنتمي إلى نمط واحد --- حالة «جيدة» ولم يتم تجديدها، أو حالة «ممتازة» وتم تجديدها: العقارات المقارنة A، B، C، D، E، G، I، K، M، N، O. يتم تكوين هذه العينة الفرعية من خلال عمليتي تصفية، وتؤدي كل منهما وظيفة مختلفة. وإستبعاد العقارات E و H و J (العقار الممتاز الذي لم يتم تجديده أبدًا والعقارين الجيدين اللذين تم تجديدهما) يترك نمطًا واقعيًا: كيف يبدو السوق عندما يكون التجديد هو السبب الرئيسي وراء حصول العقار على أعلى درجة CQA في البيانات (هنا 00.9)، كما قد يحدث. إن استبعاد المنازل الثلاثة المتوسطة، F و L و P، هو أمر مصطنع، ويتم القيام به لجعل الفشل دقيقًا وليس شبه دقيق؛ أما الحالة شبه الدقيقة، وهي الحالة الخطيرة، فسيتم تناولها بعد الجدول. تعامل مع العقارات العشرة على أنها عالم البيانات الكامل للمُقيمين --- سوق فرعي لا يحدث فيه التجديد والحالة المسجلة الأعلى أبدًا بشكل منفصل. بالنسبة لتلك المجموعة، فإن الانقسام غير محدد؛ وتُظهر العقارات الستة عشر جميعها أن الفشل جاء من بيانات السوق الفرعي، وليس من السوق.

ضمن هذه العينة الفرعية، ترتبط الحالة المسجلة و عملية إعادة التصميم بمعادلة تطابقية دقيقة: $cond = 7 + 2 \cdot remodel$ في كل واحدة من العقارات العشرة. لتلبية الملاءمة الكاملة --- المساحة الإجمالية للمعيشة (GLA) قطعة الأرض، الإطلالة، الحالة، إعادة التصميم --- يجب على البرنامج حذف أحد العمودين المتوازيين، ويحدد العمود الذي يتم حذفه ما يتم طباعته. يتم عرض كلا الحذفين:

العينة الفرعية (10)		العينة الكاملة (16)	
تم إلغاء عملية إعادة التصميم		تم حذف الشرط	
اعتراض	-240,000	110,000	-30,000
GLA	600	600	600
لوت	25	25	25
عرض	100,000	100,000	100,000
الشرط (لكل نقطة)	50,000	سقط	20,000
إعادة التصميم	سقط	100,000	60,000

هناك ثلاثة أمور في هذا الجدول تستحق الانتباه.

أولاً، لا تُعتبر معاملات الحالة ولا معاملات إعادة التصميم مجرد معاملات غير دقيقة؛ فإذا تم استبعاد واحدة منها، فإنها تغيب تمامًا، بينما يأخذ الآخر قيمة تعتمد فقط على عملية الاستبعاد. تحتوي مصفوفة التصميم على ستة أعمدة (عمود من الأرقام «1» لكل من القاطع، و GLA، والقطعة، والإطلالة، والحالة، وإعادة التصميم) ولكن خمسة منها فقط تحمل معلومات منفصلة (المرتبة 5)، ولا تقدم طريقة المربعات الصغرى إجابة واحدة: فكل زوج $(\beta_{cond}, \beta_{remodel})$ على الخط $\beta_{cond} + \frac{1}{2} \beta_{remodel} = 50,000$ ، لكل منهما قاطعه الخاص، تتم الملاءمة فيها للعشرة مبيعات بشكل متساوٍ ودقيق تمامًا. عمودا العينة الفرعية في الجدول هما نقطتان على ذلك الخط. تختلف نقاط التقاطع لنفس السبب: فكل منها يمثل السعر المطابق لمنزل بدون مساحة معيشة، وبدون قطعة أرض، وبدون إطلالة --- بدرجة CQA تساوي صفرًا في أحد المطابقات، وغير مُجدد في الآخر --- وهو ما لا ينطبق على أي منزل، ولا يمثل حقيقة سوقية.

ثانيًا، ما لا يعتمد على الانخفاض هو الفرق في السعر بين منزل ممتاز تم تجديده ومنزل جيد لم يتم تجديده: \$100,000، ويُقرأ على أنه \$50,000 نقطة عبر نقطتين في عمود واحد، وعلى أنه \$100,000 على إعادة التصميم في العمود الآخر، وفي كلتا الحالتين يكون الرقم بالضبط $20,000 + 60,000$ --- النقطتان الخاصتان بالحالة المسجلة بالإضافة إلى كل ما جلبته إعادة التصميم معها. القارئ الذي يضع \$50,000 بجانب \$20,000 في السوق ويصفه بأنه مبالغ، قد قرأ طريقة تعسفية واحدة لطباعة النتيجة (تلك التي تم حذف «إعادة التصميم» منها) كما لو كانت معدلًا؛ بينما يوضح عنوان العمود ماهيتها. تحدد المبيعات العشر، بالدولار، الفرق في السعر بين منزل ممتاز تم تجديده ومنزل جيد لم يتم تجديده، ولا تقول شيئًا عن كيفية توزيع هذا الفرق. لاحظ أن هذا ليس تحيزًا بالمعنى الوارد في البند 008 (تحيز المتغير المحذوف). الرقم \$100,000 صحيح تمامًا فيما يتعلق بالسؤال الذي يمكن للعينة الفرعية الإجابة عليه؛ إنه ببساطة سؤال مختلف.

ثالثاً، يتم استرداد قيم GLA و«lot» و«view» بقيمتها الأصلية --- 600 و25 و100,000 --- دون تغيير عن العينة الكاملة. ويتم التحديد معلماً تلو الأخرى. يمكن للنموذج تحديد معظم المعاملات ويفشل في ثلاثة منها --- وهنا الحالة، وإعادة التصميم، ومعهما نقطة التقاطع، التي تعتمد قيمتها المطبوعة فقط على أي من العمودين المتوازنين يستبعده البرنامج --- ويكون هذا الفشل غير مرئي في المعاملات التي نجحت. ما تحدده العقارات العشر هو الفرق السعري المجمع لنقطتين من «الحالة» و«إعادة التصميم» المسجلتين: \$100,000. يكمن جوهر هذا البند في التباين مع البند رقم 008 (تحيز المتغير المحذوف). ففي تلك الحالة، كانت المعلومات المطلوبة موجودة في البيانات، لكن النموذج فشل في استخدامها؛ وقد أدى إدراج مؤشر إعادة النمذجة إلى حل المشكلة بالكامل. أما هنا، فإن إدراج مؤشر إعادة البناء لا يغير شيئاً، لأن المعلومات غير موجودة في العينة الفرعية على الإطلاق. نفس المواصفات، ونفس المُقدَّر، ونفس البرنامج --- وفشل من نوع مختلف تماماً، لا يمكن علاجه إلا بخصائص خارج الخط $cond = 7 + 2 \cdot remodel$. تقوم العقارات الستة المخصصة بذلك بطريقتين مختلفتين. E، وهو منزل ممتاز لم يتم تجديده أبداً، و H و J، وهما منزلان تم تجديدهما وحالتهما جيدة، يفصل كل منهما بين العمودين مباشرةً: فكل تباين يحدد الانقسام من المبيعات نفسها، مع الأخذ في الاعتبار شكل النموذج فقط، دون أي افتراض بشأن المعدلات خارج نطاق البيانات. أما F و L و P، وهي المنازل المتوسطة، فتكسر هذه الهوية أيضاً، لكنها تفصل بين العمودين فقط من خلال افتراض: أن معدل الحالة بين 00.5 و 00.7، الذي توفره هذه المنازل، ينطبق أيضاً بين 00.7 و 00.9، حيث تقع جميع المنازل التي خضعت للتجديد. أعد إدخال F و L و P فقط، وسيرجع البرنامج رقمين مرة أخرى --- وهما هنا القيم المُؤدَّة بالضبط، لأن البيانات لا تحتوي على أي تشويش --- لكن كلاهما يستند إلى ثلاث مبيعات وإلى ذلك الافتراض. وهذا قريب جداً من الحالة المشابهة التي وصفها المدخل بأنها «خطيرة»: تم تحديدها شكلياً، لكنها تستند إلى المبيعات الثلاث التي تخرق النمط وإلى الافتراض بأن المعدل بين 00.5 و 00.7 ينطبق أيضاً بين 00.7 و 00.9. وفي حالة وجود تشويش حقيقي، فإن الرقمين سيتحركان معاً على طول الخط $\beta_{cond} + \frac{1}{2}\beta_{remodel} = 50,000$. إعادة إضافة E و H و J هي ما يميز الانقسام عن المبيعات وحدها. لم يكن التصحيح أبداً إحصائياً.

References.

- Fisher, F. M. (1966). *The Identification Problem in Econometrics*. McGraw-Hill.
- Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, 19(1), 1–67.
<https://doi.org/10.1214/aos/1176347963>
- Koopmans, T. C. (1949). Identification Problems in Economic Model Construction. *Econometrica*, 17(2), 125–144.
- Manski, C. F. (1995). *Identification Problems in the Social Sciences*. Harvard University Press.

expectation. conditional ;bias variable omitted ;endogeneity Cross-references:

Formalization Prolog Addendum:

knowledge Prolog journal's the in entry this formalizes (010-identification.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.identification-010 %%

%%

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

Identification Parameter :Title %%

Craytor Bert :Author %%

identification :Concept %%

0.2.1 :Version %%

%%

```

good ,00.5 average) score 00.10–00.0 s'Journal the on condition :0.1.v1 %%
the carries base knowledge The .(point a 000,20 ;00.9 excellent ,00.7 %%
    .units those in function generating the and score %%
        %%
    the from failure identification s'entry the reconstructs galley The %%
collinearity the proves 0/not_identified_in_subsample :itself dataset %%
    .it asserting than rather %%

so ,time run at loaded not is (dictionary/tools/Issues) dictionary The %
The .Sryer 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

----- dictionary the From ---- %
    ,econometric_method , "Identification" ,identification)term %
the by determined uniquely is it when identified is parameter A" %
    assumptions the given data observable the of distribution joint %
    reveal could data the what concerns identification ;model the of %
        ,from separate and ,to prior logically is and principle in %
        .('1.2026–dictionary' ,published , "estimation %
            %
        .(conditional_expectation ,identification)depends_on %
        .(endogeneity ,identification)depends_on %
----- %

    ,omitted_variable_bias ,endogeneity ,identification ])uses_terms
    ,multicollinearity ,conditional_expectation
    ,comp_selection ,hedonic_regression
    ,uspap ,adjustment ,paired_sales_analysis
    .([ standards_rule
    ,2/introduced_by ,2/related ,2/depends_on])uses_relations
    .([2/generalized_by ,2/formalized_by

        .(identification)article_concept
        .((2026)year ,(2)number ,(1)volume)article_issue
        .('010.2026.foundations')article_id

    .('pacificas_comps_extended/kb/...')consult)initialization –:
;request on predicates list the and 2/forall loads Prolog Sryer %
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 –:

        .(rank_deficiency)local_term
        .(design_matrix)local_term
        .(bracketing)local_term
        .(estimand)local_term

===== %
DISTINCTION THE %
===== %

.("?parameter the at guess best the is what ,sample this Given" ,estimation)asks
.("?all at down parameter the pin population entire the Would" ,identification)asks

    .(estimation ,identification)prior_to

    .failure identification an repair NOT does What %

```

```

.(better_estimator)does_not_repair
.(larger_sample)does_not_repair
.(more_sophisticated_model)does_not_repair

===== %
IDENTIFIED NOT IS REMODEL WHICH ON SUBSAMPLE THE %
===== %

%
,00.7 good ,00.5 average :levels three s'dataset the from score Condition %
.00.9 excellent %
.(S ,cond_score ,P)attribute_value -: (S ,P)cond_score

.base knowledge the from ,function generating The %
.(V ,K>true_coefficient -: (V ,K>true_coefficient_score

submarket a in ,properties (00.9) excellent or (00.7) good to Restrict %
three the) was one good no and remodeled was house excellent every where %
the Within .(exact illustration the keep to aside set are homes average %
are columns two the so ,identically remodel*2 + 7 = cond ,restriction %
.intercept an given collinear perfectly %

-: (P)in_subsample
,(P)property
,(C ,P)cond_score
,(R ,remodel_numeric ,P)attribute_value
.( 0 =:= R ,0.7 =:= C ; 1 =:= R <- 0.9 =:= C )

.(Ps ,(P)in_subsample ,P)findall -: (Ps)subsample

.semidet is collinear_in_subsample !%
%
is This .subsample the of property every at remodel*2 + 7 = cond Proves %
.asserted than rather data the from derived ,failure identification the %

-: collinear_in_subsample
,(P)in_subsample)forall
,(C ,P)cond_score )
,(R ,remodel_numeric ,P)attribute_value
.(( R*2 + 7 =:= C

.semidet is (ParamB? ,ParamA?)not_identified_in_subsample !%

-: (remodel_numeric ,cond_score)not_identified_in_subsample
.collinear_in_subsample

separate J and H ,E :kinds Two .pattern the BREAK that properties six The %
houses remodeled ;remodeled never house excellent an) directly columns the %
the that assumption the under only them separate P and L ,F ;(good at %
.00.9 to 00.7 from holds 00.7 to 00.5 from rate condition %
.(P)in_subsample +\ , (P)property -: (P)breaks_pattern
.0.7 =< C ,(C ,P)cond_score ,(P)breaks_pattern -: (P)direct_contrast
.0.7 > C ,(C ,P)cond_score ,(P)breaks_pattern -: (P)linearity_contrast

===== %
DETERMINE NOT DOES AND DOES SUBSAMPLE THE WHAT %
===== %

see ;dropped column remodel the with (10 = n) subsample the on Fitted %
,000,110 intercept :instead dropped column condition the With .tex.entry %

```

```

on lie Both .rates three same the ,000,100 remodel %
.000,50 = 2/beta_remodel + beta_cond %
.(240000- ,intercept)subsample_coefficient
.(600 ,gla_sqft)subsample_coefficient
.(25 ,lot_sqft)subsample_coefficient
.(100000 ,view_numeric)subsample_coefficient
point per % .(50000 ,cond_score)subsample_coefficient
.(deficiency rank) estimable not :remodel_numeric %

.(subsample ,remodel_numeric)not_estimable
.(110000 ,intercept ,condition_dropped)alternative_normalization
.(100000 ,remodel_numeric ,condition_dropped)alternative_normalization

.semidet is normalizations_agree_on_the_difference !%
.000,60 + 000,20*2 = remodel the on 000,100 = 000,50 * points 2 %
-: normalizations_agree_on_the_difference
,(S ,cond_score)subsample_coefficient
,(M ,remodel_numeric ,condition_dropped)alternative_normalization
,M := S * 2
,(B ,remodel_numeric)true_coefficient_score ,(A ,cond_score)true_coefficient_score
.B + A * 2 := M

.semidet is identified_sum_is_correct !%
%
recorded the equals ,point per ,coefficient condition s'subsample The %
points two the over spread 000,60 s'remodel the plus rate s'condition %
Across .000,50 = 000/2,60 + 000,20 :excellent from good separate that %
the difference price the ,000,60 + 000,20*2 = 000,100 ,points two those %
.divide cannot and determine data %

-: identified_sum_is_correct
,(S ,cond_score)subsample_coefficient
,(A ,cond_score)true_coefficient_score
,(B ,remodel_numeric)true_coefficient_score
,2 / B + A := S
.B + A * 2 := S * 2

.semidet is identification_is_parameter_by_parameter !%
%
the in values generating their at recovered are view and lot ,GLA %
.identified not are parameters other two though even subsample %

-: identification_is_parameter_by_parameter
,([view_numeric ,lot_sqft ,gla_sqft] ,R)member)forall
.(( (X ,R)true_coefficient_score ,(X ,R)subsample_coefficient )

===== %
008 ENTRY WITH CONTRAST %
===== %

,omitted_variable_bias)failure_kind
.(" .it fixes variable the Including .it use to fails model the and data the in present is information The"
,identification)failure_kind
.(" .nothing fixes variable the Including .upon relied data the from absent is information The"

.(include_the_variable ,omitted_variable_bias)remedy
.(change_the_data_or_the_claim ,identification)remedy

===== %

```

ENCOUNTERS APPRAISAL %

===== %

,bracketing)encounter

and (?all at identified rate the is) characteristics other the from apart varying contrast :questions two answers Bracketing"
guards contrast such only ;support guards Bracketing .(?found was it where range the inside lie subject the does) support ⇔
.("identification ⇔
,tract_subdivision)encounter
.("tract that within separated be cannot size and view ,plan larger the also is home side-view every Where"
,age_and_condition)encounter
.("support not may data the separation a assert lines grid separate ;together move often construction of era and condition ,Age"
,time_adjustment)encounter
,period the over changed that else everything time to attributes it :composite but identified is coefficient conditions-market A"
.("alone conditions market of effect the not is and ⇔

.not practically ,identified formally :case dangerous the are misses-Near %
,near_collinearity)dangerous_case

is this ;pattern the break that observations of handful the by determined but identified technically coefficients Two"
.("side identification the from seen multicollinearity ⇔

===== %

DEFENSIBILITY %

===== %

,order_of_questions)defensibility_consequence

for numbers confident produces order reverse the ;it estimate to best how asking before distinguish can data the what Ask"
.("quantities unmeasurable ⇔
,report_the_sum)defensibility_consequence
any of basis the disclose and determine do data the effect combined the report ,identified separately not are effects two Where"
.("applied split ⇔
,uspap_link)defensibility_consequence
(adjustments based-market) Guide Selling Mae Fannie the and (results credible) 4-1 SR USPAP ,reading s'Journal the In"
to line grid each requiring without — ?adjustment this identify data which — obligatory question identification the make ⇔
.("reader the of front in comparables the from identified be ⇔
,remedy_is_design)defensibility_consequence
absent recovers estimator no ;correlation the break that properties the find or ,period the extend ,geography the Widen"
.("information ⇔