

PATRICK KRIEGER
CARSTEN LAUSBERG

Immobilien-Risiko-Scorings: Neue Anforderungen an eine bekannte, aber unterschätzte Entscheidungsunterstützung

Patrick Krieger¹, Carsten Lausberg²

**IMMOBILIEN-RISIKO-SCORINGS:
NEUE ANFORDERUNGEN AN EINE BEKANNTE, ABER UNTER-
SCHÄTZTE ENTSCHEIDUNGSUNTERSTÜTZUNG**

Arbeitspapier Nr. 4/2014

Ursprüngliche Fassung: Oktober 2014. Diese Fassung: April 2020.

Dieses Paper wurde bei einer Zeitschrift zur Veröffentlichung eingereicht.
Bitte kontaktierten Sie die Autoren bevor Sie es zitieren.

Zusammenfassung: Scoring ist eine weit verbreitete, lang bewährte, universell anwendbare und leicht verständliche Methode zur Messung von Risiken. Leider wird sie in der immobilienwirtschaftlichen Praxis oft fehlerhaft eingesetzt und in der Wissenschaft unterschätzt. Eine „Risikoscoring-Theorie“ könnte dem abhelfen, doch die gibt es noch nicht. Es ist auch nicht unsere Absicht, sie zu schreiben. Wir möchten allerdings die Literatur um einen Vorschlag bereichern, nach welchen Regeln Scorings konstruiert und validiert werden sollten, damit sie als Risikoinstrumente taugen. Hierzu übertragen wir zunächst die Axiome für kohärente Risikomaße von Pedersen/Satchel (1998) und anderen auf Scores. Anschließend schlagen wir Verfahren zur Erstellung von Risiko-Scorings vor.

Keywords: Immobilien; Risiko; Scoring; Entscheidungsunterstützung

Herausgeber:

Campus of Real Estate an der Hochschule für Wirtschaft und Umwelt Nürtingen-Geislingen e.V.
Parkstr. 4, 73312 Geislingen
Homepage: www.campus-of-real-estate.de
E-Mail: immobilienwirtschaft@hfwu.de
Telefon: 07331 / 22 -555

Autoren:

1) Patrick Krieger, M.Sc., B.Sc.
2) Prof. Dr. Carsten Lausberg, Hochschule für Wirtschaft und Umwelt Nürtingen-Geislingen, Kontakt: carsten.lausberg@hfwu.de, Tel. 0 17 17 18 17 66

Immobilien-Risiko-Scorings: Neue Anforderungen an eine bekannte, aber unterschätzte Entscheidungsunterstützung

Zusammenfassung

Zweck

Das Scoring ist eine weit verbreitete, seit langem etablierte und universell anwendbare Methode zur Messung von Risiken, insbesondere von solchen, die sich nur schwer quantifizieren lassen. Leider wird die Scoringmethode in der Immobilienpraxis oft fehlerhaft eingesetzt und in der Wissenschaft unterschätzt. Der Zweck dieses Artikels besteht darin, die Literatur um einen Vorschlag bereichern, nach welchen Regeln Scorings konstruiert und validiert werden sollten, damit sie zu zuverlässigen Risikoinstrumenten werden können.

Gestaltung/Methodik/Ansatz

Das Paper kombiniert die aus der Literatur bekannten Regeln bzw. Axiome für kohärente Risikomaße mit denen für Scorings. Das Ergebnis ist ein Regelwerk, das ein Risikoscoring erfüllen sollte. Der Ansatz ist theoretisch und basiert auf einer Literaturübersicht und Argumentation.

Ergebnisse

Zunächst stellt das Paper klar, dass eine Risikoeinstufung für Immobilien die Änderung ihrer Rendite und nicht ihrer Qualität ausdrücken sollte, wie es in der Praxis oft geschieht. Dann werden die Axiome für ein kohärentes Risikoscoring abgeleitet, z.B. die Unabhängigkeit der Risikofaktoren. Am Schluss schlägt das Paper Verfahren für valide und verlässliche Risikoscorings vor, z.B. die Out-of-Time-Validierung.

Praktische Auswirkungen

Obwohl es sich um eine theoretische Arbeit handelt, liegt der Schwerpunkt der Arbeit auf der praktischen Anwendbarkeit. Die Ergebnisse werden anhand eines fiktiven Beispiels eines Bewertungssystems veranschaulicht.

Originalität/Wert

Regeln für Risikomaße und für Bewertungssysteme wurden schon vor langer Zeit aufgestellt, aber die Kombination ist eine Premiere. Auf diese Weise leistet das Paper einen Beitrag zur Immobilienrisikoforschung und zur Praxis des Risikomanagements.

Schlüsselwörter

Immobilien; Risiko; Punktbewertung; Entscheidungsunterstützung

1 Einleitung

Die betriebswirtschaftliche Entscheidungstheorie befasst sich mit den Entscheidungen, die ein Manager treffen sollte (normativer Ansatz) oder trifft (deskriptiver Ansatz). Sie umfasst die Beschreibung und Erstellung von Methoden, Instrumenten, Regeln, Prozessen und anderen Hilfsmitteln zur Verbesserung der Entscheidungsfindung. In der Geschichte der Entscheidungstheorie sind viele solcher Hilfsmittel entwickelt worden – grob gesagt in drei Generationen. Zunächst wurden Entscheidungen im Hinblick auf eine einzige monetäre Zielgröße wie Wert, Gewinn oder Rendite analysiert. Dies erwies sich für die meisten praktischen Anwendungen als zu einfach. Später wurden Zielsysteme geschaffen, die auch nicht-monetäre, qualitative Ziele wie die Marktposition eines Unternehmens einschlossen. Dabei mussten die Präferenzen der Entscheidungsträger einbezogen werden, da die beste Alternative nicht nur auf der Basis quantitativer Daten berechnet werden konnte. Noch später wurden Erkenntnisse aus der Psychologie und anderen Disziplinen genutzt, um die Präferenzen der Entscheider zu verstehen und irrationales Verhalten in die Modelle zu integrieren. Das Scoring ist ein typischer Vertreter der zweiten Generation. Es ermöglicht dem Entscheidungsträger, eine subjektive Wahl zu treffen – z.B. ein Haus als schön, weniger schön oder hässlich zu klassifizieren –, aber auf transparente Weise und auf der Grundlage eines Zielmodells. Dadurch wird eine intersubjektive Überprüfbarkeit geschaffen, die als Vorteil dieser Methode angesehen wird.

Eine Bewertung besteht normalerweise aus einer endlichen Anzahl von Kriterien, die auf einer Ordinalskala, einem Gewichtungssystem und einer Aggregationsregel bewertet werden. Alle Elemente können aus Expertenwissen ("Judgmental scorecard"), per Datenanalyse ("Statistical scorecard") oder durch einer Kombination aus beiden ("Hybrid scorecard") abgeleitet werden (Lux, 2012, 74f.). Dies macht das Scoring zu einem äußerst vielseitigen Instrument, das für zahlreiche Zwecke im Management, aber auch in Psychologie, Medizin, Technik und anderen Bereichen eingesetzt wird. Die Scoringmethode kann auch zur Messung von Risiken verwendet werden. Sie wird typischerweise zur Bewertung des Kreditrisikos eines Kreditnehmers eingesetzt. Zu den Anwendungen in der Immobilienbranche gehört die Messung des Investitionsrisikos, wobei Teilscores für Gebäudemerkmale, Marktbedingungen und andere Faktoren, die einen Einfluss auf den zukünftigen Wert einer Immobilie haben, vergeben werden. Diese Methode ist insofern attraktiv, als das Risiko in operationalisierbare Faktoren zerlegt wird. Im Vergleich zu anderen Risikomessungen bietet sie den Vorteil, dass Probleme schnell aufgedeckt werden können und dass ein Wert für Risikofaktoren ermittelt werden kann, die sonst – z.B. aufgrund fehlender Daten – kaum in Zahlen ausgedrückt werden können. Andererseits erschwert dieser Datenmangel die Validierung von Scorings. Es besteht auch die Möglichkeit, dass ein Risikoscore ein anderes Ergebnis liefert als ein auf Wahrscheinlichkeiten basierendes Risikomaß. Es überrascht nicht, dass Kröll und Lausberg (2012) in einer Studie über die Validität von

Risikoscorings, die in deutschen Immobilienunternehmen verwendet wurden, bemerkenswerte Mängel entdeckten. Keine Methode in ihrer Stichprobe hat das Immobilienrisiko in Übereinstimmung mit den gängigen Risikodefinitionen gemessen. Die Gründe dafür könnten darin liegen, dass die Scorings zu stark auf Expertenwissen beruhten, ursprünglich nicht für die Risikomessung, sondern für andere Zwecke geschaffen und weder validiert noch kalibriert worden waren.

Obwohl die Scoringmethode aus englischsprachigen Ländern stammt, sind dort sowohl praktische Anwendungen als auch akademische Forschung zum Zweck der Immobilienrisikomessung selten zu finden, mit Ausnahme von Blundell et al. (2005) und Adair und Hutchison (2005). In Deutschland scheinen die praktischen Richtlinien von TEGoVA (2003) und VÖB (2006) besonders weit verbreitet zu sein. Auch in der deutschen wissenschaftlichen Literatur wird das Risikoscoring behandelt, z.B. von Lorenz et al. (2006) und Lausberg und Wiegner (2009). Die Forschung zum Risikoscoring scheint jedoch etwas losgelöst vom Mainstream der Risikoforschung zu sein, in der in den letzten zwei Jahrzehnten große Fortschritte erzielt wurden. Unter anderem wurden Regeln aufgestellt, an die sich alle Risikomessungen halten müssen. Das Ziel dieses Papers ist es, geeignete Regeln für die Schaffung von Scorings zu definieren, damit sie als gültige und zuverlässige Risikoinstrumente eingesetzt werden können.

2 Eine kurze Geschichte der Scoringmethode

Wann eine Kriterienliste zum ersten Mal als Scoring bezeichnet wurde, ist nicht dokumentiert. Es ist jedoch bekannt, dass bereits zu Beginn des 20. Jahrhunderts Scorings in Bereichen wie Bildung, Medizin und Psychologie eingesetzt wurden. Mit dem Fortschritt der Wissenschaft entwickelten sich auch diese Methoden (und umgekehrt), so dass Instrumente entstanden, die bis heute verwendet werden, beispielsweise zur Beurteilung der Fähigkeiten von Studienbewerbern ("Scholastic Aptitude Test" (siehe Schonemann, 2005, S. 196)), zur Messung der Schwere eines Schädel-Hirn-Traumas ("Glasgow Coma Scale" (Teasdale und Jennett, 1974)) oder als Persönlichkeitstests ("House-Tree-Person-Test" (Buck, 1948)).

Die ersten in der Geschäftswelt angewandten Bewertungsmodelle erschienen wahrscheinlich Mitte des 20. Jahrhunderts, beispielsweise zur Auswahl zwischen Mitarbeitern oder verschiedenen Projekten. Banken verwendeten Bewertungsmodelle auch für Kreditanträge ihrer Kunden (McGrath, 1960). All diese Instrumente waren Ausfluss der Entscheidungstheorie, die zu dieser Zeit noch in den Kinderschuhen steckte. Wie zu Beginn dieses Papers angedeutet, bestand ihr Ziel darin, Entscheidungen zu verbessern und Entscheidungshilfen zu entwickeln.

Schon früh wurde dabei das Scoring als ein geeignetes Instrument für die Entscheidungsfindung identifiziert, unter anderem als indirekte Methode zur Bewertung von Wahrscheinlichkeiten, wenn eine direkte Messung nicht möglich war. Seitdem gibt es in einer Literaturströmung die Tendenz, sich auf Bewertungsregeln zu konzentrieren, vermutlich zurückgehend auf Good (1952). Daraus haben sich axiomatische Systeme entwickelt, die in Abschnitt 4 diskutiert werden. Eine andere, spätere Strömung sieht das Scoring als ein Mittel zur Aufteilung einer komplexen Entscheidungssituation in überschaubare Abschnitte an (Slovic et al., 1977). Diese Trennung in normative/ökonomische Entscheidungstheorie und deskriptive/psychologische Entscheidungstheorie besteht im Prinzip immer noch, und Scoring ist in beiden Welten zu Hause.

Die Scoringmethoden florierten in den 1970er Jahren, als sie systematisch erforscht und weiterentwickelt wurden. Im deutschsprachigen Raum sind die Arbeiten von Zangemeister (1970), Bechmann (1978) und später Lillich (1992) erwähnenswert. International bekannt wurden Saaty's Analytic Hierarchy Process (Saaty, 1980), eine Methode zur Lösung von multikriteriellen Entscheidungsproblemen, und Kaplan und Nortons Balanced Scorecard (Kaplan und Norton, 1992), ein Konzept zur Integration von finanziellen und anderen Leistungsparametern. Darüber hinaus perfektionierten die Banken die Kreditwürdigkeitsprüfung. Ein entscheidender Faktor war dabei die Arbeit am sogenannten "Z-Score" von Altman (1968). Für eine lange Zeit danach galten die Insolvenzprognose und die Bonitätsbeurteilung als ideale Anwendungen der Scoringmethodik, da ein klar definiertes Entscheidungsproblem ("Kreditvergabe: ja oder nein?") gelöst werden musste – und zwar auf der Grundlage einer Vielzahl quantitativer und qualitativer Faktoren (z.B. Eigenkapitalquote und persönlicher Eindruck) mit klarem Bezug zur Zielgröße (die (Un-)Fähigkeit des Kreditnehmers, zu zahlen). Als jedoch viele Unternehmen Anfang der 2000er Jahre infolge einer globalen Rezession in Konkurs gingen, wurde ein Problem vieler Scorings deutlich: Sie hatten die Ausfälle nicht mit ausreichender Genauigkeit vorhergesagt (Altman, 2002). Folglich ersetzten die Banken weltweit im Zuge der Einführung von Basel II das Scoring durch das Rating, eine ähnliche, aber geeignetere Methode zur Messung der Ausfallwahrscheinlichkeit. Heute ist das Scoring die bevorzugte Methode zur Ableitung eines Kreditratings in Fällen, in denen die Datenlage die empirisch-statistische Erstellung einer Scorecard erlaubt. Dazu muss aber die Anzahl der homogenen Fälle groß und die Anzahl der Vorhersagefaktoren klein sein (Lausberg und Wiegner, 2009).

In den frühen 2000er Jahren wurden Scoring und Rating – die Begriffe werden oft synonym verwendet – auch in der Immobilienbranche populär, zum Beispiel bei Investoren, die sie zur Bewertung von Investitionen und Märkten verwendeten. Traditionelle Rating-Agenturen wie Moody's Investors Service und neu gegründete Unternehmen wie Scope Ratings wagten sich an die Bewertung von hypothekarisch gesicherten Wertpapieren, Immobilienportfolios und Immobilienfonds (Schulte et al.,

2005) – übrigens nicht immer erfolgreich (Dörwald, 2004). Aber das Scoring hat eine viel längere Tradition. Frühe Beispiele sind Fines (1968) für die Landschaftsqualität, Zerbst (1974) für die Wohnqualität und Marans und Spreckelmeyer (1982) für die architektonische Qualität. 1992 schlug Diekstra (1992) ein Instrument zur Bewertung der Qualität von Bürogebäuden und Standorten vor, das später zu einer nationalen Norm in den Niederlanden wurde. Diese Beispiele verwenden die Qualität als Zielvariable, was sie von den am Ende von Kapitel 1 genannten Beispielen unterscheidet. Das nächste Kapitel befasst sich mit der Frage, ob Qualität ein geeigneter Risikoindikator ist.

3 Indikatoren für das Risiko

3.1 Qualität

Trotz ihrer zum Teil sehr individuellen Ausgestaltung sind die verwendeten Risikobewertungsmethoden in der Praxis ähnlich. Dies liegt daran, dass zur Risikomessung de factor nur endliche viele Attribute vorhanden sind, die Informationen über die vermuteten Einflussfaktoren enthalten. Dazu gehören in der Regel Objekt- und Lagequalität. Aus einer Reihe von Gründen ist die Qualität jedoch kein geeigneter Indikator für das Risiko einer Immobilie. Zwei davon erscheinen uns besonders wichtig:

- Weder im allgemeinen Sprachgebrauch noch in Wissenschaft oder Praxis werden die Begriffe Risiko und Qualität synonym, antonym oder in einer sonstigen eindeutigen Beziehung verwendet. Im Gegenteil, es gibt komplexe und variable Beziehungen zwischen den beiden Begriffen. Beispiel 1: Umgangssprachlich würde man sagen, dass es ein Zeichen schlechter Qualität ist, wenn ein Bauteil häufig kaputt geht. Übersetzt in die Fachsprache bedeutet dies, dass die Wahrscheinlichkeit eines Ausfalls ein Indikator für die Qualität ist. (Andere Indikatoren können z.B. das Herstellungsland, die Marke oder das verwendete Material sein). Andererseits kann die Qualität als Indikator für das Reparaturkostenrisiko der Komponente dienen. (Als weitere Indikatoren bieten sich Garantiebedingungen, Reparaturdauer oder Preis des Bauteils an). Daher können Qualität und Risiko jeweils als ein Teil des anderen betrachtet werden. Beispiel 2: Das Bankenviertel in Frankfurt ist einer der Top-Bürostandorte Europas. Es hat unbestreitbar eine sehr hohe Lagequalität, was sich in einem relativ geringen Leerstandsrisiko widerspiegelt (Clamor et al., 2011). Gleichzeitig besteht ein sehr hohes Risiko schwankender Marktmieten im Frankfurter Bankenviertel (IPD, 2012). Die Qualität kann sich also auf unterschiedliche Risiken unterschiedlich auswirken. Sie kann sich auch auf unterschiedliche Assets unterschiedlich auswirken während die eine Immobilie von der guten

Lage profitiert, weil ein Leerstand schnell beseitigt wird, leidet die andere darunter, weil ihre Mieter sinkende Marktmieten als Druckmittel zur Anpassung der Vertragsmieten verwenden.

- Die unklare Beziehung zwischen Risiko und Qualität ist auch auf die uneinheitliche Verwendung und Bewertung der beiden Begriffe zurückzuführen. In den letzten Jahren wurden von der Royal Institution of Chartered Surveyors (RICS, 2017) und anderen eine Reihe von Vorschlägen zur Standardisierung von Immobilienrisiken gemacht, die jedoch noch nicht den Status eines branchenweiten Standards erreicht haben. Dasselbe gilt für den Begriff der Qualität im Zusammenhang mit Immobilien. Nach Wissen der Autoren gibt es keinen ganzheitlichen Ansatz zur Definition und Messung von Immobilienqualität, obwohl es in der Architektur und bei Immobilien eine lange Tradition gibt, Unterthemen wie Landschaftsqualität und Standortqualität zu definieren (siehe Kapitel 2).

Eine etwas technischere Begründung ist bei hedonischen Regressionansätzen zu finden. Für gewerbliche Zwecke wird normalerweise der Quadratmetermietpreis als Zielvariable herangezogen. Ein Marktteilnehmer bewertet das Risiko einer Immobilie anhand der erwarteten Einnahmen $E[I]$ und des Kaufpreises P . Diese Einschätzung muss dem inhärenten Risiko der Immobilie entsprechen, welches vereinfachend mittels eines Single-Index-Modells ausgedrückt werden kann. Als Gleichung kann diese Aussage wie folgt formuliert werden

$$\frac{E[I]}{P} = r_f + \beta_r \quad (1),$$

wobei r_f den risikofreien Zins und β_r das "intrinsische" Risiko der Immobilie darstellt. Wenn man davon ausgeht, dass die Qualitätsmerkmale einer Immobilie die erklärenden Faktoren für das Risiko sind, kann Gleichung (1) nun erweitert werden zu

$$\frac{E[I(Q)]}{P(Q)} = r_f + \beta_r \quad (2),$$

da nach dem hedonischen Ansatz Merkmale sowohl Einfluss auf die Miete als auch auf den Kaufpreis haben. Der Preis einer Immobilie kann auch durch die Einnahmen multipliziert mit der ewigen Rente dargestellt werden, wodurch sich

$$P(Q) = E[I(Q)] \cdot \frac{1}{i} \quad (3)$$

ergibt. Wird weiterhin davon ausgegangen, dass i der Zins auf dem Investmentmarkt ist, auf dem sich der Preis aus Angebot S und Nachfrage D ergibt, folgt nach dem Einsetzen und Kürzen in Gleichung (2)

$$i(S, D) = r_f + \beta_r \quad (4).$$

Die Qualität wird aus der Gleichung gestrichen, nur Angebot und Nachfrage bleiben als Eingangsvariablen übrig. Aus diesem Grund wird die Cap-Rate in hedonischen Bewertungsmodellen immer separat modelliert.

Die Tatsache, dass Qualität für ein Risikoscoring ungeeignet ist, wird durch das folgende Gedankenexperiment veranschaulicht: Eine hochpreisige Immobilie kann in Stadt A oder in Stadt B errichtet werden. Die jeweiligen Grundstücke sind von der Lage her qualitativ gleichwertig. Mehrere Dienstleistungsunternehmen, die sehr gut bezahlte Fachkräfte beschäftigt hatten, sind aus Stadt A abgewandert. Die Abwanderung dieser Spezialisten wird in kurzer Zeit zu einem Überangebot an hochpreisigem Wohnraum führen. In Stadt B findet eine genau umgekehrte Entwicklung statt, da ein starker Zustrom von Experten zu erwarten ist.

Auch ohne genauere Informationen zu den beiden Märkten bietet Stadt B wesentlich bessere Aussichten für den Bau einer Immobilie im Hochpreissegment. Dies kann zur Konstruktion eines qualitativen Indikators genutzt werden. Sowohl die Immobilien als auch die Grundstücke sind von gleicher Qualität, so dass beide Optionen einen identischen Score erhalten und ein Entscheidungsträger zwischen beiden neutral sein muss. Ein weiterer zu berücksichtigender Faktor ist natürlich, dass Qualität normalerweise auf der aktuellen Situation und nicht auf zukünftigen Entwicklungen basiert.

Nun wandeln wir das Gedankenexperiment etwas ab. In der Stadt A kann entweder eine hochpreisige Immobilie in sehr guter Lage oder eine einfache Immobilie in mittelmäßiger Lage gebaut werden. Es ist bekannt, dass die ansässigen Spezialisten aufgrund der abwandernden Dienstleister die Stadt verlassen werden, und dadurch ein Überangebot an hochpreisigem Wohnraum entsteht. Gleichzeitig expandieren jedoch mehrere Industriebetriebe in der Stadt A und ziehen in steigendem Maße Hilfs- und Facharbeiter an, so dass mit einer Unterversorgung an bezahlbarem Wohnraum zu rechnen ist.

In diesem Beispiel würde eine subjektive Einschätzung die Basisimmobilie begünstigen. Die Gesamtaussichten für diese Immobilie sind günstiger, und sie wäre mit einem geringeren Risiko verbunden. Die Berücksichtigung eines auf Qualitätsfaktoren konstruierten Scorings ergibt jedoch ein paradoxes Ergebnis. Da sowohl die hochpreisige Immobilie als auch das sehr gut gelegene Grundstück eine höhere Qualität aufweisen, müsste ein rationaler Entscheider die hochpreisige Immobilie wählen. Insgesamt würde diese Auswahlheuristik zu einer ausschließlichen Allokation in hochpreisige Immobilien führen. Dies entspricht natürlich nicht der Realität.

In VÖB (2006) wird zum Beispiel kein sachlicher Unterschied zwischen Risiko und Qualität gemacht. Dies bedeutet nicht, dass die Qualität der Bauteile keinerlei Bedeutung für die Immobilie hat, sondern dass Ausreißer u.a. aus unerwarteten Schäden und Instandsetzungskosten resultieren. Außerdem handelt es sich um eine relative Variable, wie aus Gleichung (4) hervorgeht. Ein einfacher

Punktwert für die Qualität von Ausstattungsmerkmalen gibt jedoch nur Auskunft über einen nominalen Parameter, so dass die Korrelation (bestenfalls) gering ist.

3.2 Renditen als Risikoindikatoren

Ein Scoring ist ein allgemeines Instrument, das an das jeweilige Messproblem angepasst werden muss. Für die Risikomessung sollte das Scoring das allgemeine Verständnis von Risiko widerspiegeln bzw. aufgrund der subjektiven Natur des Risikos das, was der Scoringersteller als Risiko wahrnimmt. In der Finanztheorie besteht ein gewisser Konsens darüber, dass das Risiko im Allgemeinen die Veränderung der Rendite des investierten Kapitals Δi_C , ausdrückt, die auf verschiedene Weise gemessen werden kann, z.B. mittels Volatilität. Dies soll hier nicht weiter untersucht werden. Stattdessen soll gezeigt werden, dass die Verwendung der Rendite ebenfalls problematisch ist. Häufig wird die Rendite als log-Rendite ausgedrückt:

$$i_C = \ln \left(\frac{V_t}{V_{t-1}} \right) \quad (5).$$

V_t ist der Wert oder der Schlusskurs einer Aktie zum Zeitpunkt t und V_{t-1} der Vorperiode $t-1$ (der Einfachheit halber wird hier angenommen, dass keine Einnahmen erzielt werden). Dies ist für Immobilien impraktikabel, da sie nicht an einer Börse gehandelt werden. Das Risiko ist daher während der Haltezeit nicht erkennbar. Eine Möglichkeit, ein tieferes Verständnis der Risikosituation einer Immobilie zu gewinnen, ist die Verwendung eines Bewertungsparameters V^* . Gleichung (1) kann dann neu formuliert werden zu

$$i_C = \ln \left(\frac{V_t^*}{V_{t-1}^*} \right) \quad (6).$$

In diesem Fall muss die geometrische Rendite zur Risikomessung verwendet werden (Geltner et al., 2003; Giliberto, 1988). Clayton et al. (2001) zeigen jedoch, dass ein systematischer Fehler auf dem Level der Einzelimmobilie besteht und nicht in der Berechnungsmethodik selbst. Geltner et al. (2003) liefern eine detaillierte Darstellung über die Problematik der Bewertungsglättung. "Unsmoothing"-Methoden scheinen jedoch auf aggregierter Ebene ungeeignet zu sein, da der Fehler der Bewertungsglättung durch die Aggregation teilweise kompensiert wird (Bond et al., 2012; Cheng et al., 2011; Lai und Wang, 1998). Aus diesen technischen Gründen bergen Renditemaße auf Basis von Bewertungsparametern ein entsprechendes Fehlerpotenzial. Dies ist zwar kein ausschließliches Problem des Risikoscorings, jedoch für die Risikomessung von essentieller Bedeutung.

Dass Renditemaße auf Basis von Bewertungsparametern eher problematisch sind, lässt sich auch aus nutzwerttheoretischen Gründen argumentieren. Unter der Annahme "flacher" Erwartungen während der Haltedauer wäre Gleichung (2) die Nettoanfangsrendite. In der Praxis sind "flache" Erwartungen jedoch eine Ausnahme, so dass die berechnete Rendite aus einem Investmentmodell angemessener

erscheint. Eine Dekomposition scheint jedoch insofern nicht plausibel, als "eine erste grobe, aber entscheidende Messung des Risikos einer Position die Frage gilt, ob ihr Zukunftswert zur Teilmenge der akzeptierten Risiken gehört oder nicht" Artzner et al. (1999, S. 205). Dies entspricht eher

$$\frac{E[I]}{P} < r_f + \beta_r^* \quad (7),$$

so dass das Risiko nur bei einer Abweichung von den Erwartungen entsteht. r_f ist wieder der risikofreie Zinssatz und β_r^* ist nach Wheaton et al. (2001) das intrinsische Risiko einer Investition. Würde man jedoch Gleichung (4) zur Messung verwenden, würde nur die a priori Erwartung des Investors zerlegt, nicht aber das intrinsische Risiko. Da der Bewertungsparameter eine unvoreingenommene Bewertung des Marktes sein sollte, gilt

$$\frac{E[I]}{P} \sim \frac{E[I]}{V_t} \quad (8).$$

Dies entspricht der Übereinstimmung des Marktwertes mit dem Preis, den ein Investor zu zahlen bereit ist. Die Zerlegung einer Renditemessung auf der Grundlage von Bewertungsparametern würde sich daher nur auf die Erwartungshaltung des Wertermittlers beziehen. Diese Maße sind nicht falsch – im Gegenteil, diese Renditemessungen spiegeln das Marktgleichgewicht mit seinen theoretischen und statistischen Merkmalen wider. Es erscheint jedoch akzeptabel, ein unvollkommenes Renditemaß in einem unvollkommenen Markt anzuwenden (Lusht, 1988). Bei einem Risikoscoring steht das die Messung des Immobilien(objekt)risikos im Vordergrund und nicht die Messung des Marktrisikos. Aus diesem Grund müssen die Begriffe "Risiko" und "historische Volatilität" getrennt betrachtet werden. Risiko ist eine subjektive Wahrnehmung unsicherer Ereignisse, während historische Volatilität die Realisation von Ereignissen ist. Diese Ereignisse sind durch die Entscheidungen des Investors und Dritter charakterisiert, die nicht mit den Erwartungen übereinstimmen müssen. Dieser Ansatz richtet sich wesentlich stärker an Brunswicks (1952) Linsenmodell aus.

Die historische Volatilität lässt sich mit dem Renditemaß

$$i_t = \ln \left(\frac{I_t + I_G}{P + E_t} + 1 \right) \quad (9)$$

messen, wobei I_t die Mieteinnahmen im laufenden Jahr, I_G der Verkaufsgewinn, P der Kaufpreis einschließlich Investitionen der Vorperioden und E_t Kapitalkaufwendungen zum Zeitpunkt t sind. Dies entspricht einem adjustierten Return on Investment, wobei die Transaktionskosten (Sullivan et al., 1991) und das Liquiditätsrisiko bei der Veräußerung besonders berücksichtigt werden müssen (Cheng et al., 2013), um den von Wheaton et al. (2001) kritisierten Schwächen bei der Kalibrierung des Scorings entgegenzuwirken.

Um dieses Kapitel zusammenzufassen: Trotz seiner Probleme ist die Rendite besser geeignet als die Qualität und sollte als Zielvariable für das Scoring des Risikos einer Immobilie verwendet werden.

4 Anforderungen an ein Risikoscoring

Bei der Konzeption von Scorings werden die technischen Grundlagen des Scoringverfahrens in der Praxis oft nicht ausreichend berücksichtigt. Es wurden jedoch Axiome für das Scoring eingeführt, um sicherzustellen, dass die Verwendung einer additiven Präferenzfunktion erlaubt ist. Nach Lillich ist dies nur praktikabel, "wenn die folgenden fünf Voraussetzungen erfüllt sind:

- (1) $\succsim$ ist eine schwache Ordnung auf X .
- (2) Die Attribute sind gegenseitig präferenzunabhängig.
- (3) Für jedes Attribut gilt unbegrenzte Substituierbarkeit
- (4) Jede streng begrenzte Standardsequenz ist endlich.
- (5) Jedes Attribut ist wesentlich." (Lillich, 1992, 40f.)

Die schwache Ordnung auf X bedeutet, dass es keine Inkonsistenzen in der Rangfolge der Alternativen gibt. Die Präferenzunabhängigkeit gewährleistet, dass bei der Entscheidung ein "Framing" der Alternativen sowie die Schiefe der Wahrscheinlichkeitsverteilung der Alternativen keine Auswirkungen auf die Entscheidung haben dürfen (Hastie und Dawes 2010). Unbegrenzte Substituierbarkeit bedeutet, dass eine Reduktion eines Faktors durch Erhöhung eines anderen Faktors um den gleichen Betrag zu einer indifferenten Präferenz der beiden Alternativen führen muss. Das vierte Axiom – auch archimedisches Axiom genannt – führt eine messtheoretische Betrachtung der Skala ein, die in n gleich große Abschnitte unterteilt wird. Das Axiom der Wesentlichkeit postuliert, dass eine Entscheidung nicht auf unwesentlichen Kriterien beruhen darf. (Lillich, 1992)

Da die ursprüngliche Methodik nicht auf die Risikoscorings übertragbar ist, müssen die Axiome modifiziert werden. Je nachdem, wie das Risiko betrachtet wird, kann ein Maß als Abweichungsgröße oder Verlustgröße betrachtet werden (Albrecht, 2003). Das eingeführte System für Scorings bezieht sich zum einen auf ein Präferenz-Nutzen-Verhältnis, und zum anderen werden keine Axiome für Risikomaße einbezogen. Im einfachsten Fall kann das axiomatische System von Pedersen und Satchell (1998) mit der Erweiterung von Rockafellar et al. (2002) verwendet werden, so dass eine Überführung in das kohärente System von Artzner et al. (1999) gewährleistet ist. In den letzten Jahren wurden zusätzliche Anforderungen an Risikomaße – insbesondere an komonotonische Additivität, Robustheit und Erhebbarkeit – beispielsweise von Emmer et al. (2015) aufgestellt. Diese werden hier nicht weiter diskutiert, weil sie entweder alternative Definitionen der oben genannten Axiome liefern oder sich auf verteilungsbasierte Risikomessungen beziehen.

Das System von Artzner et al. (1999) besteht aus den folgenden Axiomen, die wünschenswerte Eigenschaften für alle Risikomaße sind:

- a) $P(A) > 0 \quad \forall A \in \Omega$

- b) $P(A) > 0 \quad \forall A \in \Omega$
- c) $P(cA) = cP(A)$
- d) $P(X_1) + P(X_2) \geq P(X_1 + X_2) \quad P(X + c) = P(X)$

Das erste Axiom (a) impliziert, dass Risiko nicht negativ werden darf. Die positive Homogenität (b) legt fest, dass eine Erhöhung einer Risikoposition zu einer Erhöhung des Risikos um den gleichen Faktor dieser Position führen muss. Die Subadditivität (c) besagt, dass die Summe der Einzelrisiken größer oder gleich dem Risiko einer Verbundmessung ist. Dies beschreibt die Korrelation zwischen Risiken. Die Shift-Invarianz (d) besagt, dass das Hinzufügen einer risikofreien Position zu einer risikotragenden Funktion nicht zu einer Erhöhung des Gesamtrisikos führen darf.

Axiome für die Risikoeinstufung können durch die Kombination der beiden Systeme miteinander abgeleitet werden. Zusammenfassend ergibt sich das folgende axiomatische System:

- I. $, \geq '$ ist eine schwache Ordnung auf X .
- II. Die Risikofaktoren sind voneinander unabhängig.
- III. Unbegrenzte Substitution von Risikofaktoren.
- IV. In jedem Risikofaktor gilt Subadditivität.
- V. Jede streng begrenzte Standardreihenfolge ist endlich.
- VI. Jeder Risikofaktor ist positiv.
- VII. Jeder Risikofaktor ist wesentlich.

Als erstes muss die Präferenzunabhängigkeit in eine allgemeine Unabhängigkeit (Bedingung II.) der Risikofaktoren umgewandelt werden, da jetzt die Umweltbedingungen (und nicht mehr die Präferenzen) im Vordergrund stehen. Darüber hinaus ist eine positive Homogenität für eine unbegrenzte Substituierbarkeit unerlässlich (III.), da eine unbegrenzte Substituierbarkeit einen linearen Zusammenhang zwischen zwei Risikofaktoren erfordert, der durch eine positive Homogenität gewährleistet wird. Da unbegrenzte Substituierbarkeit positive Homogenität impliziert, kann letztere auch weggelassen werden. Die Forderung positiver Risikofaktoren (VI.) setzt zwar voraus, dass keine unwesentlichen Faktoren einbezogen werden dürfen, ändert aber nichts daran, dass wesentliche Risikofaktoren (VII.) nicht weggelassen werden dürfen.

Das folgende hypothetische Beispiel zeigt ein Risikobewertungsmodell für einzelne Anlageobjekte, das die festgelegten Bedingungen erfüllt (siehe Tab. 1).

Attribute (a)	Scala	Werte (v)	Gewicht (w)	Score (s)
Objekt: Gebäudezustand (GZ)	Ordinal	1 bis 10	0,45	$s_1 = v_1 w_1$
Markt: Trend der Marktmiete (TM)	Ordinal	1 bis 10	0,35	$s_2 = v_2 w_2$
Standort: Verkehrsanbindung (VA)	Ordinal	1 bis 10	0,20	$s_3 = v_3 w_3$
Gesamtscore				$\sum_{a=1}^n w_a s_a$
Nicht berücksichtigte Attribute:				
Objekt: Instandhaltungsstau (IS)	Met- risch	EUR 0 bis $+\infty$	0,00	
Standort: Standortqualität (SQ)	Ordinal	1 bis 10	0,00	

Tab. 1: Fiktives Beispiel eines Immobilienscoringmodells

Unter der Annahme, dass die Person, die das Scoring vornimmt, Immobilien mit einem geringen Risiko bevorzugt und dass sich das Gesamtrisiko der Rendite aus Immobilien-, Markt- und Standortrisiko zusammensetzt, entspricht das Scoring der Präferenzfunktion des Investors (Bedingung I.). Die Risikofaktoren (Attribute) sind unabhängig (II.), weshalb der Instandhaltungsstau aufgrund seiner hohen Korrelation mit dem Gebäudezustand im Scoring nicht berücksichtigt wurde. (Außerdem hätte es die Gewichtung aufgrund einer unterschiedlichen Skala erschwert, da es in einem solchen Fall ratsam ist, die Variable zu transformieren). Ein niedrigerer Wert für einen Risikofaktor kann durch einen höheren Wert für einen anderen Faktor kompensiert werden, ohne die Bewertung zu verändern (III.). Bei allen Faktoren wird das Risiko reduziert, wenn mehr Objekte in das Portfolio aufgenommen werden (IV.). Die Subadditivität gilt aber auch für die Gewichtung und sollte 1 nicht überschreiten. Das fünfte Axiom ist für jedes Objekt erfüllt, dessen Zustand, Marktmiete und Verkehrsanbindung bewertet werden kann (V.). In einer (fiktiven) Regressionsanalyse eines großen Immobilienportfolios wurden alle drei Faktoren als signifikant befunden, was sich in der Gewichtung widerspiegelt (siehe Tab. 2, die die statistischen Ergebnisse der Regressionsanalyse zeigt, die für die Erstellung des Scoringmodells verwendet wurde). Der Standortqualitätsfaktor hat eine geringere Aussagekraft als der vergleichbare Verkehrsanbindungsfaktor und wurde daher – zusammen mit anderen unbedeutenden und unterdrückenden Variablen, die das Vorzeichen der Gewichtung willkürlich ändern – eliminiert (VI.). Die hohe Aussagekraft des Modells, gemessen als modifiziertes R^2 , belegt auch, dass keine signifikanten Faktoren ausgelassen wurden (VII.).

Variable	Koeffizient	T-Statistik	P-Wert	Signifikanz
Y-Achse	-403,4925	-6,0681	0,0000	***
GZ	61,4262	5,7599	0,0000	***
TM	42,3671	2,8690	0,0064	***
VA	35,9301	2,2991	0,0264	**
IS	-1,7448	-1,3271	0,1915	
SQ	4,1489	0,8146	0,4198	
MX	9,4248	1,6721	0,1018	
Modell: Adj. R ² = 0,88, F = 0,0000***				

Tab. 2: Fiktives Beispiel einer Regressionsanalyse als Grundlage für das Scoringmodell in Tab. 1

5 Dekomposition des Risikos

Auf der Grundlage dieser Axiome kann ein entsprechendes Risikoscoring konstruiert werden, das bestimmte Qualitätsanforderungen erfüllt. In einem umfassenden Vergleich von 136 Studien sowie einer Metaanalyse konnten sowohl Grove und Meehl (1996) als auch Grove et al. (2000) zeigen, dass "mechanische" Schätzungen der Faktorgewichtungen im Mittel besser und in jedem Fall mindestens ebenso gut sind wie die Expertenschätzungen. Dies bestätigt die früheren Erkenntnisse, dass statistische Schätzungen den Expertenschätzungen vorzuziehen sind (Meehl, 1954). Einhorn (1972) beschreibt jedoch als beste Strategie eine Kombination aus Expertenschätzungen zur Identifizierung von Inputparametern und statistischen Schätzungen zur Verbesserung der Validität. Ein häufiger Fehler in der Praxis ist die Annahme einer allgemeinen Validität der Methode – einmal aufgestellt, wird die Validität eines Scorings im Laufe der Zeit nicht mehr überprüft. Dies ist besonders wichtig bei Expertenschätzungen, da Erfahrung und damit Präzision in der Bewertung nur durch positives und negatives Feedback aufgebaut werden kann (Einhorn und Hogarth, 1978).

Ein naheliegender Einstiegspunkt ist Ross' (1976) Arbitrage Pricing Theory als Multi-Faktormodell, auch wenn die Prämissen aus Short-Selling und der Ausnutzung von Arbitrage für Immobilien unzutreffend sind. Der Aufbau ist jedoch konsistent mit dem Linsenmodell (Brunswick, 1952), das sich gut dazu eignet, die Gültigkeit zwischen Umwelt und kognitivem System des Entscheidungsträgers herzustellen (Karelaia und Hogarth, 2008). Im einfachsten Fall wird dazu eine lineare Regression verwendet

$$y = \sum_{i=1}^n \beta_i x_i + \varepsilon \quad (10),$$

so dass die Gewichte nach der Methode der kleinsten Quadrate bestimmt werden können. Auf diese Weise ist auch die Unabhängigkeit der Faktoren x_i gewährleistet, da sonst die Kovarianzmatrix nahe an der Singularität liegt und eine Schätzung unmöglich macht. Darüber hinaus ist $\sum \beta_i \leq 1$ bei

entsprechender Transformation der Faktoren x_i unerlässlich. Dies steht zwar im Konflikt zu den herkömmlichen Scoringmethoden, aber die Festlegung der Subadditivität bedeutet, dass die Summe der Gewichte kleiner 1 sein kann, ohne das Wesentlichkeitsaxiom zu verletzen. Außerdem muss jedes Gewicht $\beta_i > 0$ sein, was jedoch bereits durch die t-Statistik der Faktoren abgefangen wird. Negative Gewichte dürfen nicht berücksichtigt werden, auch nicht in der statistischen Schätzung. Dies ist ein Hinweis darauf, dass ein Faktor übergewichtet und dann durch die gleiche Information wieder reduziert wurde. Dies kann bei Kennzahlen geschehen, die sich im Erklärungsgehalt überschneiden und invers zueinander verlaufen. In einer Metastudie haben Karelaia und Hogarth (2008) gezeigt, dass Experten bei einer großen Anzahl von Faktoren wesentlich mehr "Bias" in die Schätzung einbringen, als bei einer sparsamen Anzahl von Faktoren. Dies lässt sich auch mit herkömmlichen statistischen Methoden, wie dem Informationskriterium von Akaike (1998), leicht lösen.

6 Zusammenfassung und Ausblick

Es konnte gezeigt werden, dass ein Risikoscoring mehr Qualitätsanforderungen erfüllen muss, als oft angenommen wird. Dies dürfte für alle Investoren, Vermögensverwalter, Portfoliomanager, Aufsichtsbehörden und andere Fachleute, die Scoringmethoden zur Messung des Immobilienrisikos anwenden, von Interesse sein. Darüber hinaus wurden statistische Methoden vorgeschlagen, um die Probleme des Immobilienscorings anzugehen. Die Modellierung eines Risikoscorings kann jedoch vielseitig sein, und in der Tat ist wenig über verschiedene Systeme bekannt, die entsprechende Qualitätskriterien angeben. Erste Schritte in diese Richtung wurden von Adair und Hutchison (2005) und anderen unternommen. Eine Richtung für zukünftige Forschungen ist die Identifizierung von Kennzahlen bzw. Faktoren mit hohem Erklärungsgehalt, sowohl auf Einzelobjektebene als auch auf Portfolioebene. Ein weiteres offenes Feld für die Forschung, das in diesem Artikel nicht berührt wurde, ist die Schulung und das De-Biasing der Nutzer (Hubbard und Evans, 2010).

Investoren, die ein Risikoscoring verwenden wollen, sollten sicherstellen, dass ein valides Instrument für Risikomessung und -berichterstattung verwendet wird. Wenn die oben genannten Kriterien und Axiome verletzt werden, sollte davon ausgegangen werden, dass eine Risikobewertung nicht valide ist. Dies ist insbesondere dann der Fall, wenn das Scoring keine harten Daten wie z.B. Renditen als Zielvariablen operationalisiert. Die Rendite ist aufgrund verschiedener theoretischer und praktischer Probleme nicht perfekt, aber sie ist die beste bekannte Größe. Umgehungslösungen mit Proxy-Variablen sollten nur dann akzeptiert werden, wenn sie messtechnische Vorteile gegenüber der Rendite bieten. In jedem Fall ist die Qualität kein geeigneter Proxy für das Risiko einer Immobilie – entgegen der Ansicht vieler Praktiker. Darüber hinaus sollten die Bewertungsfaktoren ausgewählt und die

Gewichte mit statistischen Methoden geschätzt werden – natürlich unter der Voraussetzung, dass die Datenmenge ausreichend ist. Obwohl Expertenwissen in vielerlei Hinsicht bei der Konstruktion eines Scoring hilfreich sein kann, sind Expertenschätzungen der Gewichte eher ungeeignet. Validierung und Rekalibrierung müssen regelmäßig stattfinden. Das bedeutet, dass die Prognosefähigkeit des Scorings mit Daten getestet wird, die außerhalb des Zeitraums und der Stichprobe liegen. Schließlich sollte ein Scoring nicht allein als alleiniges Instrument zur Risikobewertung einer Immobilie oder eines Immobilienportfolios verwendet werden. Scoring hat seine Vorzüge als ergänzendes Instrument zur Messung der Faktoren, die nicht leicht messbar sind.

7 Literaturverzeichnis

- Adair, A.; Hutchison, N. (2005): The reporting of risk in real estate appraisal property risk scoring. In: *Journal of Property Investment & Finance* 23 (3), S. 254–268.
- Akaike, Hirotugu (1998): Information Theory and an Extension of the Maximum Likelihood Principle. In: Emanuel Parzen, Kunio Tanabe und Genshiro Kitagawa (Hg.): *Selected Papers of Hirotugu Akaike*. Springer Series in Statistics (Perspectives in Statistics). New York: Springer, S. 199–213.
- Albrecht, P. (2003): Zur Messung von Finanzrisiken. Mannheimer Manuskripte zu Risikotheorie, Portfolio Management und Versicherungswirtschaft. Universität Mannheim. Online verfügbar unter <https://ub-madoc.bib.uni-mannheim.de/210/1/MAMA03.pdf>.
- Altman, Edward I. (1968): Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. In: *Journal of Finance* 23 (4), S. 589–609.
- Altman, Edward I. (2002): Revisiting Credit Scoring Models in a Basel 2 Environment. NYU Working Paper No. S-CDM-02-06. New York University, New York.
- Artzner, P.; Delbaen, F.; Eber, J.-M.; Heath, D. (1999): Coherent Measures of Risk. In: *Mathematical Finance* 9 (3), S. 203–228.
- Bechmann, Arnim (1978): Nutzwertanalyse, Bewertungstheorie und Planung. Bern, Stuttgart: Haupt (Beiträge zur Wirtschaftspolitik, 29).
- Blundell, Gerald F.; Fairchild, Simon; Goodchild, Robin N. (2005): Managing Portfolio Risk in Real Estate. In: *Journal of Property Research* 22 (2-3), S. 115–136. DOI: 10.1080/09599910500456759.
- Bond, S. A.; Hwang, S.; Marcato, G. (2012): Commercial Real Estate Returns: An Anatomy of Smoothing in Asset and Index Returns. In: *Real Estate Economics* 40 (4), S. 637–661.
- Brunswick, Egon (1952): The conceptional framework of psychology. Chicago: The University of Chicago Press (International Encyclopedia of Unified Sciences).
- Buck, John N. (1948): The H-T-P test. In: *Journal of Clinical Psychology* 4 (2), S. 151–159. DOI: 10.1002/1097-4679(194804)4:2<151::AID-JCLP2270040203>3.0.CO;2-O.
- Cheng, P.; Lin, Z.; Liu, Y. (2011): Heterogeneous Information and Appraisal Smoothing. In: *Journal of Real Estate Research* 33 (4), S. 443–469.
- Cheng, P.; Lin, Z.; Liu, Y. (2013): Liquidity Risk of Private Assets: Evidence from Real Estate Markets. In: *Financial Review* 48 (4), S. 671–696.

- Clamor, Tim; Haas, Heide; Voigtländer, Michael (2011): Büroleerstand – ein zunehmendes Problem des deutschen Immobilienmarktes. Institut der deutschen Wirtschaft Köln (IW-Trends, 4/2011).
- Clayton, J.; Geltner, D.; Hamilton, S. W. (2001): Smoothing in Commercial Property Valuations: Evidence from Individual Appraisals. In: *Real Estate Economics* 29 (3), S. 337–360.
- Diekstra, Frans (1992): How to determine quality in real estate. In: *Site Selection Europe* (12), S. 20–23.
- Dörwald, Natascha (2004): Umstrittene Bewertung. In: *portfolio INTERNATIONAL*, 01.02.2004 (Februar), S. 16. Online verfügbar unter https://www.wiso-net.de/document/PORT__965.
- Einhorn, H. J.; Hogarth, R. M. (1978): Confidence in judgment: Persistence of the illusion of validity. In: *Psychological Review* 85 (5), S. 395–416.
- Einhorn, Hillel J. (1972): Expert measurement and mechanical combination. In: *Organizational Behavior & Human Performance* 7 (1), S. 86–106.
- Emmer, Susanne; Kratz, Marie; Tasche, Dirk (2015): What is the best risk measure in practice? A comparison of standard measures. In: *Journal of risk* 18 (2), S. 31–60. Online verfügbar unter <http://arxiv.org/pdf/1312.1645v4>.
- Fines, K. D. (1968): Landscape evaluation: A research project in East Sussex. In: *Regional Studies* 2 (1), S. 41–55.
- Geltner, D.; MacGregor, B.; Schwann, G. (2003): Appraisal Smoothing and Price Discovery in Real Estate Markets. In: *Urban Studies* 40 (5-6), S. 1047–1064.
- Giliberto, S. M. (1988): A Note on the Use of Appraisal Data in Indexes of Performance Measurement. In: *Real Estate Economics* 16 (1), S. 77–83.
- Good, I. J. (1952): Rational Decisions. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 14 (1), S. 107–114. Online verfügbar unter <http://www.jstor.org/stable/2984087>.
- Grove, W. M.; Meehl, P. E. (1996): Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. In: *Psychology, Public Policy, and Law* 2 (2), S. 293–323.
- Grove, William M.; Zald, David H.; Lebow, Boyd S.; Snitz, Beth E.; Nelson, Chad (2000): Clinical versus mechanical prediction: A meta-analysis. In: *Psychological Assessment* 12 (1), S. 19–30.
- Hastie, R.; Dawes, R. M. (2010): Rational Choice in an Uncertain World. The Psychology of Judgment and Decision Making. 2. Aufl. Los Angeles...: Sage.

- Hubbard, Douglas; Evans, Dylan (2010): Problems with scoring methods and ordinal scales in risk assessment. In: *IBM Journal of Research and Development* 54 (3), S. 1–10.
- IPD (2012): Performancevergleich gewerblicher Immobilieninvestments in Deutschland nach Standort und Objekttypen. Wiesbaden, August 2012.
- Kaplan, Robert S.; Norton, David P. (1992): The Balanced Scorecard—Measures that Drive Performance. In: *Harvard Business Review* (January-February), S. 71–79.
- Karelaia, N.; Hogarth, R. M. (2008): Determinants of Linear Judgment: A Meta-Analysis of Lens Model Studies. In: *Psychological bulletin* 134 (3), S. 404–426.
- Kröll, F.; Lausberg, C. (2012): Measuring real estate risk with scoring models: Theoretical background, empirical evidence and requirements for future use. Edinburgh: 19th European Real Estate Society Conference, 2012. Online verfügbar unter <https://eres.architecturez.net/doc/oai-eres-id-eres2012-235>.
- Lai, T.-Y.; Wang, K. (1998): Appraisal Smoothing: The Other Side of the Story. In: *Real Estate Economics* 26 (3), S. 511–535.
- Lausberg, Carsten; Wiegner, Andreas (2009): Marktdaten im Immobilienrating. In: O. Everling und R. Slowik (Hg.): *Praxishandbuch Rating von Immobilienportfolios*. Köln: Immobilien-Manager-Verlag (ImmobilienFachwissen), S. 23–43.
- Lillich, L. (1992): *Nutzwertverfahren*. Heidelberg: Physica-Verlag (3).
- Lorenz, David; Trück, Stefan; Lützkendorf, Thomas (2006): Addressing risk and uncertainty in property valuations: a viewpoint from Germany. In: *Journal of Property Investment & Finance* 24 (5), S. 400–433. DOI: 10.1108/14635780610691904.
- Lusht, K. M. (1988): The Real Estate Pricing Puzzle. In: *Real Estate Economics* 16 (2), S. 95–104.
- Lux, Nicole (2012): *Assessing Real Estate Risk – Applied Models, Concepts, Methods*. London: Euro-money Books.
- Marans, R. W.; Spreckelmeyer, K. F. (1982): Measuring Overall Architectural Quality: A Component of Building Evaluation. In: *Environment and Behavior* 14 (6), S. 652–670. Online verfügbar unter <https://doi.org/10.1177/0013916582146002>.
- McGrath, James J. (1960): Improving credit evaluation with a weighted application blank. In: *Journal of Applied Psychology* 44 (5), S. 325–328. DOI: 10.1037/h0045044.
- Meehl, P. E. (1954): *Clinical versus Statistical Prediction. A Theoretical Analysis and a Review of the Evidence*. Minneapolis: University of Minnesota Press.

- Pedersen, C. S.; Satchell, S. E. (1998): An Extended Family of Financial-Risk Measures. In: *The Geneva Papers on Risk and Insurance Theory* 23 (2), S. 89–117.
- RICS (2017): Real estate risk management under the Alternative Investment Fund Managers Directive (AIFMD) RICS guidance note, Europe. London. Online verfügbar unter <https://www.rics.org/de/news-insight/latest-news/news-opinion/real-estate-risk-management-and-valuation-under-the-aifmd/>.
- Rockafellar, R. T.; Uryasev, S.; Zabarankin, M. (2002): Deviation Measures in Risk Analysis and Optimization. Department of Industrial and Systems Engineering. University of Florida, Gainesville.
- Ross, Stephen A. (1976): The arbitrage theory of capital asset pricing. In: *Journal of Economic Theory* 13 (3), S. 341–360.
- Saaty, Thomas L. (1980): The Analytic Hierarchy Process. Planning, Priority setting, Resource Allocation. New York...: McGraw-Hill.
- Schonemann, Peter H. (2005): Psychometrics of Intelligence. In: Kimberly Kempf-Leonard (Hg.): Encyclopedia of social measurement. Amsterdam.: Elsevier (Gale virtual reference library), S. 193–201.
- Schulte, Karl-Werner; Rottke, Nico; Pitschke, Christoph (2005): Transparency in the German real estate market. In: *Journal of Property Investment & Finance* 23 (1), S. 90–108. DOI: 10.1108/14635780510575111.
- Slovic, P.; Fischhoff, B.; Lichtenstein, S. (1977): Behavioral Decision Theory. In: *Annual Review of Psychology* 28, S. 1–39. DOI: 10.1146/annurev.ps.28.020177.000245.
- Sullivan, M. J.; Cassidy, S. M.; Ermer, C. M. (1991): A Note on the Effect of Transaction Costs on Real Estate Investment Returns. In: *Journal of Real Estate Research* 6 (1), S. 113–117.
- Teasdale, Graham; Jennett, Bryan (1974): Assessment of Coma and Impaired Consciousness: A Practical Scale. In: *The Lancet* 304 (7872), S. 81–84. DOI: 10.1016/S0140-6736(74)91639-0.
- TEGoVA (2003): European Property and Market Rating: A Valuer's Guide. London, 2003. Online verfügbar unter https://www.tegova.org/data/bin/a56efb621c7ae1_EPMR1.pdf.
- VÖB (2006): VÖB-ImmobilienAnalyse. Instrument zur Beurteilung des Chance-/Risikoprofils von Immobilien. Berlin.
- Wheaton, W. C.; Torto, R. G.; Sivitanides, P. S.; Southard, J. A.; Hopkins, R. E.; Costello, J. M. (2001): Real Estate Risk: A Forward Looking Approach. In: *Real Estate Finance* 18 (3), S. 20–28.

Zangemeister, Christof (1970): Nutzwertanalyse in der Systemtechnik: eine Methodik zur multidimensionalen Bewertung und Auswahl von Projektalternativen. München: Wittemann.

Zerbst, R. H. (1974): The determinants of single family residential property values with special reference to the measurement of housing quality and environmental factors. Doctoral dissertation. Ohio State University. Online verfügbar unter <https://etd.ohiolink.edu/>.