Stability of Parameters in Hedonic Urban Land Price Models

Wieser, Robert

August 2009

PARAMETERSTABILITÄT IN HEDONISCHEN BODENPREISMODELLEN

ROBERT WIESER

Robert.Wieser@housing-economist.eu

Version, August 2009

Abstrakt

Diese Arbeit untersucht die Wirkungen unterschiedlicher Spezifikationen und Schätzverfahren auf die Ergebnisse in einer hedonischen Bodenpreisanalyse für Wien. Untersucht werden Kaufpreistransaktionen mit unbebauten Wohnbauland-grundstücken und Abbruchobjekten im Zeitraum 1987 bis 2003. Es kommen unterschiedlich dichte Spezifikationen und unterschiedliche Schätzverfahren, wie OLS, Box-Cox, Cluster-Fixed Effects oder Verfahren der räumlichen Ökonometrie zur Anwendung. Die Ergebnisse werden anhand unterschiedlicher statistischer Tests bewertet. Die Datenstruktur weißt zugleich Querschnitts- und Zeitreihencharakter auf. In der Anwendung der räumlich ökonometrischen Modelle wird daher auch untersucht, ob durch Konstruktion geeigneter raumzeitlicher Matrizen bessere Ergebnisse erzielt werden können als mit den Standardformen nearest neighbour und Delaunay. Die Ergebnisse zeigen zum Teil sehr hohe Schwankungsbreiten der Parameter. Insbesondere für asymmetrisch im Raum verteilte Attribute erscheint eine adäquate Modellierung der räumlichen Interdependenzen bzw. fehlender räumlicher Informationen entscheidend.

1. Einleitung

Entwicklung und Bestimmungsfaktoren von Immobilienpreisen sind sowohl einzel- wie volkswirtschaftlich von Interesse. Informationen über Wertentwicklungen und Einflussfaktoren der Preise in verschiedenen Marktsegmenten bilden wichtige Entscheidungsgrundlagen für Investoren, Projektentwickler und Bauträger. Banken, Versicherungen und Investmentfonds nutzen Informationen zur Wertentwicklung im Rahmen der quantitativen Asset Allocation, d.h. der Aufteilung der Finanzmittel auf verschiedene Anlageklassen wie Aktien, Anleihen und Immobilien. Projektentwickler und Bauträger treffen Standortentscheidungen und bewerben ihre Projekte auf Basis von Informationen über Lagefaktoren, wie Freiraumqualität, Verkehrserschließung und soziale Infrastrukturen. Aber auch Staat und Zentralbanken benötigen zur Beurteilung, Planung und Kontrolle von Lenkungsmaßnahmen fundierte Informationen über die Entwicklung von Preisen und Einflussfaktoren auf den Immobilienmärkten (Maurer et al, 2001).

Hedonische Bodenpreismodelle haben sich in den letzten Jahrzehnten zu einem wertvollen Werkzeug der Analyse von Immobilienpreisen entwickelt. Ihr Anwendungsbereich wird ständig erweitert, was vor allem durch drei Entwicklungen getragen ist: Erstens durch die Steigerung der Rechenkapazität von Computern, was die Berechnungszeiten der sehr datenintensiven und zunehmend komplexeren Analysen erheblich reduziert. Zweitens durch den immer weiter verbreiteten Aufbau von entsprechenden Immobiliendatensätzen, und drittens durch die fortschreitende Entwicklung im Bereich der empirischen Methoden und die Bereitstellung entsprechender Software.

Eine vollständige hedonische Analyse vollzieht sich in zwei Schritten. Im ersten Schritt werden die impliziten Preise der Grundstückseigenschaften durch Regression der Eigenschaften auf den Grundstückspreis ermittelt. Die geschätzten impliziten Preise reichen aus, um beispielsweise Grundstücke näherungsweise zu bewerten oder Preisindizes zu erstellen. Unter bestimmten restriktiven Annahmen – vollkommen elastisches oder unelastisches Angebot - können aus den impliziten Preisen marginale Zahlungsbereitschaften für einzelne Grundstückseigenschaften abgeleitet werden. Für eine vollständige Kosten-Nutzen-Analyse beispielsweise von staatlichen bzw. planerischen Maßnahmen am Bodenmarkt benötigt man allerdings zusätzlich eine Schätzung der Angebots- und Nachfragestruktur. Im zweiten Schritt wird durch Regression von Nachfrage- und Angebotsfaktoren (demand and supplier shifters) auf die impliziten Preise die Nachfrage- und die Angebotsstruktur am Bodenmarkt geschätzt.

Die meisten hedonischen Analysen enden bereits nach dem ersten Schritt. Schon dieser erste Schritt, die Schätzung der hedonischen Preisfunktion und der impliziten Attributpreise, ist allerdings mit einer Reihe von Problemen verbunden. Diese Arbeit vergleicht die Ergebnisse unterschiedlicher hedonischer Modellansätze und verfolgt dabei drei Zielsetzungen. Zum ersten soll anhand statistischer Prüfmaße geklärt werden, welche Spezifikationen und Schätzverfahren „beste“ Ergebnisse liefern. Zum zweiten wird der Frage nachgegangen, wie robust die hedonische Analyse im Hinblick auf unterschiedliche Modellstrukturen und Schätzmethoden ist (Frage der Parametersensitivität). Zum dritten wird untersucht, ob durch Konstruktion einer geeigneten raum-zeitlichen Gewichtungsmatrix die Ergebnisse räumlich ökonometrischer Ansätze gegenüber der Anwendung von Standardmatrizen der nearest neighbour- und Delaunay-Form verbessert werden können. Grundlage der Analysen sind Transaktionen mit unbebauten Wohnbaulandgrundstücken und Abbruchobjekten auf dem Wiener Bodenmarkt, dokumentiert in der Wiener Kaufpreissammlung im Zeitraum von 1987 bis 2003.

2. *Spezifikation in hedonischen Bodenpreismodellen*

2.1 Grundlagen

Die Theorie der hedonischen Preise geht davon aus, dass die Marktkräfte auf (freien) Immobilienmärkten bewirken, dass unterschiedlich hohe Immobilienpreise auf unterschiedliche Zusammensetzungen verschiedener Merkmale (Größe, Lage, Nachbarschaft etc.) zurückzuführen sind. Auf funktionierenden Märkten kommt es zu Gleichgewichtspreisen für die unterschiedlichen Ausprägungen der einzelnen Attribute der Grundstücke (Rosen, 1974). Der beobachtbare Marktpreis eines Grundstückes wird durch eine hedonische Preisfunktion beschrieben:

$$(1) \qquad P = P(\mathbf{x})$$

wobei P den Marktpreis und $\mathbf{x} = (x_1, x_2, ..., x_k)$ einen Vektor von k Ausprägungen (Größe, Distanz zum Zentrum, Distanz zu Freiräumen usw.) der einzelnen Grundstücksmerkmale darstellt. Die partielle Ableitung der hedonischen Preisfunktion nach dem Merkmal x_i

$$(2) \qquad p_{x_i}(x_i; \mathbf{x}_{-i}) = \frac{\partial P(\mathbf{x})}{\partial x_i} \quad \text{(i=1 bis k)}$$

ergibt den marginalen impliziten Preis als Funktion wider, in der $\mathbf{x}_{-i}$ den Vektor aller anderen Grundstücksattribute darstellt. Der marginale implizite Preis gibt an, welchen zusätzlichen Betrag ein Nachfrager zahlen muss, wenn er (bei gleicher Ausstattung mit allen anderen Merkmalen - *ceteris paribus*) vom Merkmal x_i eine marginal veränderte Ausprägung (z.B. einen Quadratmeter mehr an Grundstücksfläche oder einen Quadratmeter mehr an Wohnfläche) haben möchte.

Für den Fall einer links-halblogarithmischen (semi-logarithmischen) Preisfunktion hat Gleichung (1) folgendes Aussehen:

$$\mathbf{P} = \exp^{\mathbf{xb+e}}$$

Mit exp als Basis des natürlichen Logarithmus. Nach Umformung erhält man

$$\ln \mathbf{P} = \mathbf{xb} + \mathbf{e}$$

Die Regression ergibt

$$\ln \mathbf{p} = \mathbf{x\beta} + \boldsymbol{\varepsilon}$$

wobei b und e die unbekannten, wahren und β und ε die geschätzten Parameter kennzeichnen. Der prognostizierte Preis eines Grundstücks ergibt sich als $p = \exp^{\mathbf{x\beta}}$ und der (implizite) Preis für die Eigenschaft x_1 bei der Ausprägung x_1^j (j = 1 …N; beispielsweise wenn N Grundstücke Grundstücksflächen von 150 – 15.000m^2 aufweisen) bei gegebenem Niveau der anderen k-1 Eigenschaften ($x_{i\neq 1}$) in monetären Einheiten ergibt sich als: $p_{x_1^j} = \exp^{x_1^j \beta_1}$.

Hier wird deutlich, dass der implizite Preis von der Menge (Ausprägung) der Eigenschaft abhängt. Diese Nichtlinearität der hedonischen Preise wirft besondere Probleme im Zusammenhang mit der Identifikation des Nachfragesystems in der zweiten Stufe der hedonischen Analyse auf (vgl. Malpezzi, 2002, Sheppard, 1999, Follain und Jimenez, 1985). Für die erste Stufe, die Schätzung der Preisfunktion, reicht hier im Moment aus, festzustellen, dass in den meisten Fällen die Eigenschaft beim Mittelwert ihrer Ausprägung bewertet wird, d.h.: $x_1^j = \mu(x_1)$, also beispielsweise bei der durchschnittlichen Grundstücksgröße.

Die Schwierigkeiten, die in dieser ersten Stufe der hedonischen Analyse auftreten, sind vielfältig (Vgl. Sheppard, 1999, Malpezzi, 2002): (1) Die Theorie gibt keinen Hinweis auf die

Form der Preisfunktion (Spezifikation). (2) Mit den zugänglichen Daten sind nicht alle relevanten, den Wert bestimmende Attribute ausreichend gut abgedeckt. Eine Reihe von Preis bestimmenden Faktoren bleibt aufgrund fehlender Informationen unberücksichtigt. (3) Der städtische Bodenmarkt setzt sich aus einer Reihe von Teilmärkten zusammen, die sowohl sachlich, als auch geographisch oder sozioökonomisch bestimmt sein können. Für die hedonische Preisanalyse bedeutet das, dass die gesuchten impliziten Preise der Grundstücksmerkmale je nach Teilmarkt variieren können. (4) Viele Bodenmärkte sind staatlich reguliert. Die möglichen Wirkungen staatlicher Interventionen sind in den Modellen zu berücksichtigen (vgl. z.B.: Cheshire und Sheppard 1995). Diese Arbeit konzentriert sich auf das Problem der Spezifikation und die Wahl einer adäquaten statistischen Methode.

Eine zentrale Schwäche der hedonischen Preisanalyse ist darin zu sehen, dass die Theorie wenige Anhaltspunkte dafür liefert, welche funktionale Form der Preisfunktion die richtige ist. Viele ökonomischen Theorien sind nicht so weit entwickelt, dass sie über einen allgemeinden Zusammenhang der Form $y = f(x_1, \ldots, x_n)$ hinaus auch noch die Funktionsform von f festlegen können. Auch im Fall der hedonischen Immobilienpreisanalyse gibt es keine stringente theoretische Basis für die Wahl der richtigen Funktionsform. Es ist Aufgabe des Empirikers, eine geeignete Spezifikation zu finden. Diese kann handwerklich nur anhand unterschiedlicher Kriterien und kontextabhängig angenähert werden. Dabei sollte man möglichst nach *a priori* Informationen suchen, welche sich aber in der Regel vorwiegend nur auf die Vorzeichen der Parameter richten. Über die Elastizitäten (Stärke der Abhängigkeit, Form der Elastizität) kann aufgrund der mitunter starken Unterschiedlichkeiten zwischen Boden(teil)märkten viel weniger ausgesagt werden.

In der hedonischen Bodenpreisanalyse wurden bisher je nach Untersuchungsziel und Datengrundlage unterschiedliche Standpunkte vertreten. Vor allem in frühen Arbeiten, in denen zumeist die Datengrundlage (Zahl der Beobachtungen bzw. Transaktionen am Bodenmarkt, Informationen über wertbestimmende Attribute) schwach war, hat man in der Regel, auch aus Gründen der Freiheitsgrade bzw. der Effizienz der Schätzung, sehr knappe Spezifikationen gewählt. Eine möglichst einfache Spezifikation („parsimoneus specification") wurde und wird in der Ökonometrie aus guten Gründen auch immer wieder gefordert (vgl. Assenmacher und Braun, 1981). Beispielsweise zeigt sich, daß knappe Spezifikationen in Prognosemodellen oft besser arbeiten als komplexere Funktionsformen. Im Fall der hedonischen Preisanalyse haben knappe Spezifikationen aber unter Umständen einen hohen „Preis". Fehlende Informationen über entscheidende Eigenschaften der Immobilien können je nach gewähltem Schätzansatz unterschiedlich stark verzerrend wirken. Wir untersuchen in der Folge die Wirkungen unterschiedlicher Modellspezifikationen (Funktionsformen und Variablen) im Zusammenhang mit unterschiedlichen Schätzverfahren.

2.2 Funktionsformen

Bei der Festlegung der Regressionsgleichung läßt man sich gewöhnlich weitgehend von Einfachheitsüberlegungen leiten (Assenmacher, 2004). Im Fall der hedonischen Analyse ist dies, wie bereits erwähnt, nicht immer der Fall. Wir vergleichen im Folgenden die Ergebnisse aus linearen, semi-logarithmischen, doppel-logarithmischen, gemischt semi- und doppellogarithmischen mit Box-Cox transformierten Formen, Group-level fixed effects - und räumlich-ökonometrischen Modellen. In den meisten hedonischen Analysen wurden lange Zeit lineare Formen (linear bis doppel-logarithmisch) favorisiert, was teilweise gestützt war durch die Untersuchung von Cropper et al (1988), die in einer Simulationsstudie nachgewiesen haben, daß zur Beurteilung der Zahlungsbereitschaft für ein Charakeristikum die einfachen linearen Formen den anderen überlegen sind, sobald bestimmte Eigenschaften nicht beobachtet werden können.

Seit der Untersuchung von Cropper et al. (1988) sind zwei Jahrzehnte vergangen. In dieser Zeit haben sich die ökonometrischen Techniken geändert und die Datensätze sind größer und zum Teil auch gehaltvoller geworden. Heute kommen neben räumlich-ökonometrischen parametrischen Modellen, immer häufiger auch semi- und nichtparametrische Verfahren und

Verfahren der Panelökonometrie zum Einsatz. Eine aktuelle Simulation von Kuminoff et al. (2008) kommt zu folgenden Ergebnissen: Erstens, der Einbezug von räumlichen fixed effects (lokalen Dummies) steigert die Performance (Prognosefähigkeit) der einfachen linearen Modelle, unabhängig davon, ob einzelne Attribute bekannt sind oder nicht. Zweitens, die Performance von komlexeren nicht-linearen Modellen (Box-Cox-Formen) steigt durch den Einbezug der lokalen Dummies nur, wenn Attributsinformationen fehlen, nicht wenn alle relevanten Attribute berücksichtigt sind. Drittens, die relative Performance unterschiedlicher Spezifikationen hängt von der Größe der Datensätze (Anzahl an Beobachtungen) ab.

Die Untersuchungen von Cropper et al. (1988) und von Kuminoff et al. (2008) bieten wertvolle Unterstützung bei der Suche nach der besten Spezifikation. Ihre Stärke und zugleich Schwäche liegt darin, daß die Simulationsergebnisse auf Annahmen über die Form der Nutzenfunktionen der Immobilienkäufer und auf der Annahme geräumter Märkte beruhen. Die Beurteilung erfolgt bei Kuminoff et al. (2008) allein auf Grundlage der Prognosefähigkeit im Hinblick auf den Gesamtpreis der Immobilie und nicht auf Grundlage der Schätzgüte für die impliziten Attributepreise. Die folgende Untersuchung wird unter anderem zeigen, daß der Einsatz räumliche Dummys (fixed effects) zwar die Prognosefähigkeit im Hinblick auf den Immobilienpreis erhöht, andererseits aber in vielen Fällen verzerrend auf einzelne Attributpreise einwirkt. Es ist daher immer eine Frage des Untersuchungszieles, ob man räumliche Dummys verwenden sollte oder nicht.

Beim linearen Modell geben die Parameter absolute (implizite) Preise für eine Einheit des Attributes an, bei semi-logarithmischen Modellen präsentieren die Parameter prozentuelle Aufschläge und bei doppel-logarithmischen Modellen repräsentieren sie Elastizitäten des Bodenwertes eine relative Veränderung der jeweiligen Merkmalsmenge[1]. Die Modelle haben folgende Formen, wobei hier zwischen dichotomen (D) und kontinuierlichen (C) erklärenden Variablen (**x**) unterschieden wird:

Linear (Variablen und Parameter linear)

$$\mathbf{p} = K + \sum_{i \in D} \beta_i x_i + \sum_{j \in C} \beta_j x_j + \boldsymbol{\varepsilon} \quad (4)$$

mit hedonischen Preisen β_i für dichotome Variable und $\beta_j \mu_j$ für kontinuierliche Variable, errechnet beim Mittelwert μ_j der erklärenden Variablen x_j.

semi-logarithmisch (erklärende Variablen und Parameter linear, abhängige Variable nicht linear)

$$\ln \mathbf{p} = K + \sum_{i \in D} \beta_i x_i + \sum_{j \in C} \beta_j x_j + \boldsymbol{\varepsilon} \qquad (5)$$

mit hedonischen Preisen $exp(\beta_i)-1$ für dichotome Variable und $exp(\beta_j \mu_j)$ für kontinuierliche Variable, errechnet beim Mittelwert μ_j der x_j.

doppel-logarithmisch (Variablen nicht linear, Parameter linear)

$$\ln \mathbf{p} = K + \sum_{i \in D} \beta_i x_i + \sum_{j \in C} \beta_j \ln x_j + \boldsymbol{\varepsilon} \qquad (6)$$

mit hedonischen Preisen $exp(\beta_i)-1$ für dichotome Variable und $\beta_j \mu_j$ für kontinuierliche Variable, errechnet beim Mittelwert μ_j der x_j.

Malpezzi (2002) listet eine Reihe von Vorteilen von semi-logarithmischen (bzw. gemischt semi und doppel-logarithmischen) Modellen auf: (1) Der monetäre Wert (implizite Preis) einer Eigenschaft hängt auch von den anderen Eigenschaften einer Immobilie ab. Im linearen Modell sind die impliziten Preise unabhängig von den anderen Eigenschaften. Beispielsweise unterstellt das lineare Modell einen gleich hohen Wertzuwachs durch ein zusätzliches Badezimmer für ein Haus mit einem und für ein Haus mit fünf Schlafzimmern. Bei semi-

---

[1] Da dichotome Variablen den Wert Null annehmen können und der Logarithmus von Null nicht definiert ist, können diese auch nicht logarithmiert werden. In diesem Fall geben die Parameter nicht Elastizitäten, sondern Semi-Elastizitäten an. Bei einer 0-1Variable (Dummy) zeigt der Parameter, wie sich der Bodenpreis relativ verändert, wenn das Merkmal vorhanden ist.

logarithmischen Funktionsformen werden diese Interaktion und die Nichtlinearität der Preise berücksichtigt. (2) Die Koeffzienten eines semi-logarithmischen Modells haben eine einfache Interpretation. Wenn beispielsweise in einer Gleichung zur Erklärung von Häuserpreisen der Koeffizient für eine Zentralheizung 0,15 beträgt, dann würde der Einbau einer Zentralheizung den Wert eines Hauses um ca. 15 Prozent erhöhen[2]. (3) Die Semi-logarithmische Form reduziert das Problem der Heteroskedastizität (unterschiedliche Varianz der Fehlerterme). (4) Die Modellergebnisse sind für den Preis einfach zu berechnen. (5) Diese Funktionsform ist relativ flexibel, d.h. man kann auf der rechten Seite ohne Probleme Variablen unterschiedlicher Form und Interaktionen zwischen Variablen einbauen.

Die Annahme, dass die lineare, die semi-logarithmische oder die doppel-logarithmische Funktionsform den Zusammenhang zwischen Bodenpreis und Preisbestimmenden Faktoren adäquat beschreibt, wurde allerdings schon früh kritisiert (Goodman, 1978; Halvorson/Pollakowski, 1981). Jede dieser drei Formen beinhaltet eine dem Modell aufgesetze Restriktion, weil in allen Fällen die Funktionsform *a priori* bzw. exogen festgelegt wird. Eine Möglichkeit, die Funktionsform endogen bestimmen zu lassen, bietet die Methode von Box und Cox (1964).

Box-Cox Transformation

$$\frac{p^{\psi}-1}{\psi} = K + \sum_{i \in D} \beta_i x_i + \sum_{j \in C} \beta_j \frac{x_j^{\lambda}-1}{\lambda} \qquad (7)$$

mit hedonischen Preisen $\beta_i/(p^{\psi-1})$ für dichotome Variable und $\beta_j \frac{\mu_j^{\lambda}}{p^{\psi-1}}$ für kontinuierliche Variable, errechnet beim Mittelwert μ_j der x_j (vgl. Cheshire und Sheppard, 1998, FN. 3 S. 363).

Die Skalarparameter ψ und λ, und der Parametervektor $\boldsymbol{\beta}$ lassen sich mittels Maximum Likelihood-Schätzung bestimmen, wobei die folgende Funktion maximiert wird:

$$L_{\max} = (\psi - 1)\sum_{l=1}^{n} \ln p_l - \frac{n}{2} \ln\left(\frac{SSR(\psi, \lambda, \boldsymbol{\beta})}{n}\right)$$

Die Funktionsform ändert sich mit den Werten für ψ und λ. Für den Fall $\psi = \lambda = 1$ liegt eine lineare Funktion vor, für $\psi = \lambda = 0$ eine doppel-logarithmische und für $\psi = 0$ und $\lambda = 1$ eine semi-logarithmische Funktion (siehe beispielsweise Pindyck und Rubinfeld 1991, S. 249ff).

Gegen die Verwendung der Box-Cox Transformationen werden verschiedene Aspekte vorgebracht (vgl. Cassel und Mendelsohn, 1985). Ergeben sich beispielsweise andere Werte als 0 und 1 für die Parameter ψ und λ, so besitzen die Parameterschätzer $\boldsymbol{\beta}$ keine leicht zugängliche Interpretation mehr. Richtig ist, daß die Interpretation durch die Transformation der Variablen erschwert wird, sie läßt sich aber durch entsprechende Umrechnungen finden. Schwerwiegender ist der Einwand, daß die Bestimmung der optimalen Transformationsparameter dominiert wird durch jene Variablen, welche den größten Teil der Variation in der abhängigen Variablen erklären. Dieser Umstand wird besonders dann zum Problem, wenn im Mittelpunkt der Untersuchung Einflussfaktoren stehen, die nicht den Hauptteil in der Variation erklären. Dann ist es möglich, dass der Effekt dieser weniger bedeutenden Faktoren durch die Transformation nicht korrekt berücksichtigt wird[3].

Die Einwände gegen die Box-Cox-Transformation verdeutlichen den Zielkonflikt zwischen einem optimalen „Fit" des Modells und der Ermittlung „richtiger" hedonischer Preise für die

---

[2] Halvorsen und Palmquist (1980) zeigen, dass die Prozentinterpretation nur eine Annäherung ist, die umso schlechter ist, je höher der Koeffizient. Eine bessere Approximation ist die Umformung exp(b)-1, wobei b der Koeffizient und exp die Basis des natürlichen Logarithmus ist (vgl. auch Kennedy, 1981).

[3] Wird beispielsweise der Gesamtkaufpreis einer Liegenschaft erklärt, dann wird die Grundstücksgröße allein oft mehr als 50 Prozent, Grundstücksgröße und Verkehrserreichbarkeiten oft mehr als 70% der Variation im Kaufpreis erklären, während andere Einflussfaktoren (Lärm, Freiraumzugang, Luftqualität) nur marginal zur Erklärung beitragen. Trotzdem sind auch diese Faktoren wertbestimmend und stehen oft im Zentrum von Untersuchungen.

Eigenschaften der Immobilien. Die ursprüngliche Intention von Box und Cox (1964) war, durch die Transformation Residuen zu erhalten, die stärker der Normalverteilungsannahme entsprechen und weniger heteroskedastisch sind als die Residuen der einfachen linearen Regression. Anzumerken ist, dass die BoxCox-Transformation Maximum-Likelihood-Schätzer unter der Annahme der Normalverteilung erhalten.

2.3 Räumliche und temporale Effekte

Ein weiterer Faktor in der Suche nach der richtigen Spezifikation ist die Auswahl der einzubeziehenden Variablen. In den meisten Fällen wird es notwendig sein, so viel wie möglich an Preisbestimmenden Informationen einzubauen. Die vorhandenen Informationen werden jedoch von unterschiedlicher Bedeutung für die Preisbildung sein, so dass in manchen Fällen die Nichtberücksichtigung weniger bedeutender Faktoren kaum eine Rolle spielen wird. Welcher Faktor welche Bedeutung hat, ist a priori aber nicht immer leicht zu sagen. Daher ist in der hedonischen Analyse in vielen Fällen wahrscheinlich ein General-to-specific (*Gets*) – Ansatz einem Specific-to-general (*Stge*) – Ansatz vorzuziehen (vgl. Munro und Angulo, 2009). Eine Ausnahme stellt die Berechnung von Preisindizes dar. Untersuchungen zeigen, daß knappe Spezifikationen hier sehr oft gute Ergebnisse erzielen, wenn die räumlichen und zeitlichen Interdependenzen zwischen den Transaktionen entsprechend berücksichtigt werden (vgl. z.B. Bourassa et al, 2007).

Cluster-Effekte (lokale/regionale Teilmärkte)

Auch in einem *Gets*-Ansatz steht man vor dem Problem, daß nicht alle preisrelevanten Faktoren meßbar sind bzw. nicht über alle Faktoren Informationen vorliegen. Man kann aber für nicht beobachtbare Faktoren auf regionaler oder lokaler Ebene mit sog. „Group-Level-Fixed Effects“ (GLFE) kontrollieren. Darunter versteht man den Einbezug von Indikatorvariablen (Dummies) für die einzelnen Cluster (Teilmärkte) im untersuchten Bodenmarkt. Bourossa et al. (2007) sind der Auffassung bzw. erhalten das Ergebnis, dass GLFE-Schätzer anderen Verfahren der räumlich-ökonometrischen Analyse (Lattice-Modelle und geostatistische Modelle) in der Prognose von Bodenpreisen zumindest nicht unterlegen, teilweise sogar überlegen sind.

Im Grunde kann jede der oben angeführten Funktionsformen mit Teilmarktdummies versehen werden. Im Fall der doppel-logarithmischen Form ändert sich die Funktion dann zu

$$\ln \mathbf{p} = K + \sum_{i \in D} \beta_i x_i + \sum_{j \in C} \beta_j \ln x_j + \sum_{l \in L} \beta_l x_l + \boldsymbol{\varepsilon} \qquad (8)$$

wobei die x_l die lokalen Cluster-Effekte (Group Level Dummies) darstellen. Der Einsatz von lokalen Teilmarktdummies ist vor allem dann vorteilhaft, wenn nur wenige oder keine räumlichen Informationen zu den Daten vorliegen. Ist der Datensatz ausreichend groß, dann kann man durch eine kleingliedrige Strukturierung des Untersuchungsraumes die lokalen Clustereffekte gut abfangen und dadurch Verzerrungen in den Strukturvariablen vermeiden. Voraussetzung dafür ist aber, daß die Cluster-Effekte nicht zu stark mit den Strukturvariablen korrelieren. In dem Fall entsteht das Problem der Multikollinearität und die Parameter verlieren an Aussagekraft.

Das Modell in (8) kann je nach Annahmen über die Eigenschaften der Residuen mit unterschiedlichen Methoden geschätzt werden. Es können mit dieser Datenstruktur Fixed-Effects oder Random-Effects-Modelle geschätzt werden, wenn man davon ausgeht, dass jedes Grundstück zu einem wohldefinierten Cluster (Teilmarkt) gehört. (Wooldridge, 2002, S. 454f). Fixed-Effects kommen zur Anwendung, wenn wir davon ausgehen, dass unbeobachtete Teilmarkteffekte mit einer oder mehreren der erklärenden Variablen korreliert sind. Random Effects dagegen unterstellen, dass keine Korrelation zwischen den unbeobachteten Teilmarkteffekten und den erklärenden Variablen existiert. Ein Hausman-Test kommt zur Anwendung um zu prüfen, ob Korrelation vorliegt oder nicht. Wenn andererseits die Teilmarkteffekte korreliert sind oder Heteroskedastizität der Cluster vorliegt, müssen die Standardfehler der OLS-Schätzung und alle gebräulichen Teststatistiken korrigiert werden.

Die Anwendung von Fixed Effects Modellen in diesem Zusammenhang ist jedoch nicht unproblematisch. Wenn das Ziel der Analyse darin besteht, den Wert bestimmter räumlicher Charakteristika in den Bodenpreisen zu ermitteln, wird der Einsatz von Teilmarktdummies oder Fixed Effects problematisch. Viele der räumlichen Charakteristika sind im Raum nicht gleichverteilt sondern konzentrieren sich in wenigen Teilräumen (etwa Fluglärm, bestimmte Freiräume oder Freiraumtypen, gemischte Bodennutzungen, ethnische Konzentration etc.). Dann ist aber davon auszugehen, daß die Teilmarktdummies mit solchen Charakteristika hoch korreliert sind. Sie müssen es geradezu sein, weil diese Charakteristika mitunter konstitutiv für das Entstehen von unterschiedlichen räumlichen Teilmärkten sind. In dem Fall würde also der Einbezug von Group Level Fixed Effects eine Schätzung der Werte der Charakteristika zerstören[4].

Nachbarschaftliche Spillovers

Die Cluster- bzw. Teilmarkteffekte basieren auf lokalen bzw. regionalen Faktoren, die sich auf eine größere Gruppe von Grundstücken gleichermaßen auswirken. Beispielsweise hat ein hoher Grünraumanteil in einem Bezirk Auswirkungen auf alle Grundstücke in einem gewissen Buffer dort. Davon zu unterscheiden sind Effekte, die auf noch kleinräumigerer Ebene auf die einzelnen Grundstücke einwirken. Die Stadt Wien beispielsweise ist bau-struturell in mehr als 10.000 Baublöcke und administrativ in mehr als 1.600 Zählsprengel unterteilt. Bei einer Datengrundlage von weniger als 4.000 Grundstückstransaktionen zwischen 1987 und 2003 können Clustereffekte auf Baublockebene aus Gründen der Freiheitsgrade nicht geschätzt werden. Auch eine Schätzung von Clustereffekten auf Zählsprengelebene könnte schon problematisch sein. In dieser Arbeit werden die Teilmarkteffekte durch Dummies für die 82 Katastralgemeinden berücksichtigt. Dadurch werden aber kleinräumigere Effekte nicht abgefangen und es ist davon auszugehen, daß trotz Einbezug einer Reihe von Lagevariablen die Residuen der Grundstücke nicht unabhängig voneinander sind. Ein weiterer Grund für kleinräumig korrelierte Residuen (räumliche Autokorrelation) könnte in der Praxis der Vergleichswertermittlung im Rahmen von Kaufpreisgutachten liegen. Sehr oft werden nur wenige (3 bis 5) strukturell, zeitlich und räumlich ähnliche Grundstücke zur Vergleichswertermittlung herangezogen. Unter der Voraussetzung, daß die Gutachten tatsächlich wesentlich zur Preisermittlung herangezogen werden, sind die Störgrößen abhängig bzw. räumlich autokorreliert.

Regressionsmodelle der räumlichen Ökonometrie beziehen die Nachbarschaftseffekte in den Grundstückspreisen in die Schätzung mit ein. Wir betrachten hier zwei Typen von räumlichen Regressionsmodellen: das Spatial Lag- bzw. das mixed regressive-spatial autoregressive Model (SAR) und das conditional autoregressive model bzw. general spatial autoregressive model (CAR) (vgl. LeSage 1998, Kap. 3)[5]. Basis aller Lattice-Modelle ist die Definition einer räumlichen Gewichtungsmatrix W, welche die Abhängigkeiten zwischen den Grundstücken modelliert. Die Gewichtungsmatrix wird exogen festgelegt, weshalb unterschiedlichste Formen von solchen Matrizen diskutiert werden.

Das SAR-Modell ist für eine semi-logarithmische Funktionsform wie folgt formuliert:

$$\ln \mathbf{p} = \rho \mathbf{W}_1 \ln \mathbf{p} + \sum_{i \in D} \beta_i x_i + \sum_{j \in C} \beta_j \ln x_j + \boldsymbol{\varepsilon} \qquad (9)$$

Der Spatial-Lag in der abhängigen Variablen lnp stellt einen Durchschnittswert von ln**p** der benachbarten Grundstücken dar. Das CAR-Modell ist eine Verallgemeinerung bzw. Erweiterung folgender Form:

---

[4] Vgl. auch Beck und Katz (2001), die die Anwendung von Fixed-Effects in einem anderen Zusammenhang kritisieren.

[5] In der Literatur gibt es unterschiedliche Bezeichnungen für die einzelnen Typen der Lattice-Modelle.

$$\ln \mathbf{p} = \rho \mathbf{W}_1 \ln \mathbf{p} + \sum_{i \in D} \beta_i x_i + \sum_{j \in C} \beta_j \ln x_j + \mathbf{u} \qquad (10)$$

$$\mathbf{u} = \lambda \mathbf{W}_2 \mathbf{u} + \boldsymbol{\varepsilon}$$

Dieses Modell hat zusätzlich zum Spatial-Lag der abhängigen Variablen einen Spatial-Lag im Fehlerterm. Dieser läßt sich durch die Nichtberücksichtigung relevanter Einflussfaktoren, insbesondere in Form von Spatial-Lags exogener Variablen, begründen.

2.4 Analyse mit IPCS- (independently-pooled-cross-section) Daten

Die Beobachtungen von Transaktionen auf dem Bodenmarkt lassen sich auf drei verschiedenen Ebenen gewinnen. Als Querschnittdaten, als Längsschnittdaten (Zeitreihendaten) oder als eine Kombination von beidem. Bei Querschnittdaten werden Beobachtungen für die Variablen an unterschiedlichen Grundstücken erhoben, die alle zu einem gewissen Zeitpunkt oder innerhalb einer gewissen Periode transferiert wurden. In der Praxis reicht ein Zeitpunkt (grob gesprochen ein Tag) zumeist nicht aus, da zuwenige Beobachtungen vorliegen würden. Die meisten Querschnittschätzungen in der Vergangenheit basieren daher auf Transaktionen, die innerhalb einer Periode von nicht mehr als 1 bis 1,5 Jahren statt gefunden haben. In vielen dieser Studien wurde die Außerachtlassung der Zeitdimension damit begründet, daß es sich beim untersuchten Datensample um eine Periode stabiler Preise gehandelt hat.

Von Längsschnittdaten spricht man, wenn sich die Beobachtungen für aufeinanderfolgende, meist gleich lange Perioden (z.B. ein Jahr) auf dieselben Grundstücke beziehen. Reine Zeitreihendaten in dieser Form sind für den Bodenmarkt nicht gegeben, da dieselben Grundstücke niemals regelmäßig in Jahresabständen verkauft werden. Hätte man beispielsweise einen Grundstücksdatensatz mit N Grundstücken, die regelmäßig über einen längeren Zeitraum T alle Jahre transferiert werden, dann läge ein sog. balancierter Paneldatensatz mit N*T Beobachtungen vor. Solche Datensätze gibt es in der Realität für Merkmale von Staaten, Kantonen, Bundesländern, Gemeinden, Unternehmen, Haushalten usw., leider aber nicht für den Bodenmarkt.

In der ökonometrischen Literatur werden Datensätze, die keine Paneldaten, sondern unabhängig gepoolte Daten mit Querschnitts- und Zeitreihencharakter darstellen, als independently-pooled-cross-section-Daten (hier abgekürzt als IPCS-Daten) bezeichnet (Wooldridge 2002, Kap.13). Ein offensichtlicher Grund, IPCS-Daten zu verwenden, die dieselbe Grundgesamtheit (hier den Bodenmarkt Wien) repräsentieren, ist, die Datenmenge zu erhöhen. Mit größeren Datensätzen erhält man präzisere Schätzwerte und sie ermöglichen machtvollere statistische Tests. Eine Einschränkung dabei ist, daß die funktionalen Zusammenhänge zwischen Grundstückspreis und zumindest einigen Charakteristika über den Untersuchungszeitraum konstant bleiben müssen.

Abgesehen davon weißt die Analyse von IPCS-Daten gegenüber einer reinen Querschnittsdatenanalyse nur wenige zusätzliche Komplikationen auf, die zudem teilweise leicht behebbar sind. Beispielsweise können unterschiedliche Verteilungseigenschaften der abhängigen Variablen zu verschiedenen Zeitpunkten (z.B. in verschiedenen Jahren) relativ einfach durch Einbezug von Zeitdummys abgefangen werden. Zeitdummys sind Indikatorvariablen (0,1-Variablen) für jeden einzelnen Zeitpunkt. Typischerweise werden Jahresdummys oder Quartalsdummys, zuweilen auch Monatsdummys gewählt, je nach Fragestellung und Datensatzgröße. Der früheste Zeitpunkt im Datensatz (erstes Jahr, erstes Quartal, erster Monat) wird gewöhnlich als Referenzpunkt herangezogen.

Auch das Problem sich verändernder Zusammenhänge über die Zeit kann mit Hilfe von Dummyvariablen abgefangen werden. Will man beispielsweise untersuchen, ob sich der Einfluss der Erreichbarkeit im öffentlichen Verkehr auf die Grundstückspreise im Zeitablauf geändert hat, dann kann man dies durch eine Interaktion von Erreichbarkeitsvariable und Zeitdummys überprüfen. Generell können mit Hilfe von Chow-Tests strukturelle Veränderungen über die Zeit überprüft werden (Wooldridge, 2002, S. 431 ff).

Ein weiterer Vorteil von IPCS-Daten liegt darin, daß auch Methoden der Panel-Ökonometrie anwendbar sind. Wenn man beispielsweise annimmt, daß jedes Grundstück zu einem bestimmten räumlichen Cluser (Teilmarkt) gehört, der über die Zeit konstant ist, lassen sich auch Fixed-Effects oder Random-Effects Methoden anwenden. Da die Grundstückspreise innerhalb von wohl definierten räumlichen Clustern sehr wahrscheinlich (positiv) korreliert sind, kann es entscheidend sein, unbeobachtete Effekte auf Clusterebene zu berücksichtigen. Fixed Effects-Methoden kommen zur Anwendung, wenn wir annehmen, daß die unbeobachteten Cluster-Effekte mit einer oder mehreren der unabhängigen Variablen korreliert sind. Da die Cluster (Teilmärkte) zumeist unterschiedlich groß sind, kommen Fixed Effects-Methoden für unbalancierte Panel zur Anwendung. Random Effects Methoden sind nur dann angebracht, wenn wir davon ausgehen, daß die Cluster-Effekte nicht mit den erklärenden Variablen korrelieren (deshalb „random“ effects).

Zur Konstruktion einer räumlich-temporalen Gewichtungsmatrix W

Kern aller Lattice-Modelle ist die Konstruktion einer räumlichen Gewichtungsmatrix W. Die bekanntesten Formen basieren auf einer Anzahl räumlich engster Nachbarn (nearest neighbour matrices) oder auf den (euklidischen) Distanzen zwischen benachbarten Grundstücken. In empirischen Arbeiten stellt sich sehr häufig die Matrix der vier nächsten Nachbarn als beste Annäherung für die räumlich relevanten Effekte heraus. Auch die Delaunay Triangulation (variable Anzahl mit maximal 6 Nachbarn) wird häufig verwendet.

Das Problem mit den Standardmatrizen, die auch häufig in ökonometrischer Software implementiert sind, ist, daß sie nur die räumliche, nicht aber die zeitliche Dimension betrachten. Sie sind daher für reine Querschnittstudien gedacht, und nicht optimal geeignet für Daten, die sowohl räumlichen als auch ausgeprägten zeitlichen Charakter haben. Der hier untersuchte Datensatz reicht von 1987 bis 2003 und hat daher eine ausgeprägte zeitliche Dimension. Bei Verwendung einer Standard-nearest neighbour Matrix stellt sich daher ein theoretisches Problem. Im Extremfall kann das bedeuten, daß der Preis eines Grundstücks, das im Jahr 1987 verkauft wurde durch den Preis eines benachbarten Grundstück, das im Jahr 2003 verkauft wurde mit erklärt wird. Obwohl der Bodenmarkt ein langes räumliches Gedächtnis hat (Quartiere behalten oft sehr lange ihren definitorischen Charakter) wäre diese Lösung theoretisch nicht sauber (manche Quartiere verändern eben ihren Charakter in relativ kurzer Zeit). Es ist daher notwendig eine Matrix zu konstruieren, die neben der räumlichen Nähe auch den zeitlichen Abstand zwischen Verkaufszeitpunkten berücksichtigt, in der Form, daß in der Konstruktion der Gewichtungsmatrix nur benachbarte Grundstücke berücksichtigt werden, die innerhalb eines Zeitraums vorher verkauft wurden. Das Hauptproblem, das sich dabei stellt, ist, dass für die ersten Verkäufe im Sample keine entsprechenden Nachbarn vorliegen. Dem sollte beim Aufbau der Endmatrizen durch Elimination von Beobachtungen Rechnung getragen werden (Pace et al. 1998).

Eine der ersten Anwendungen zeit-räumlicher Matrizen in der hedonischen Analyse findet sich in Pace et al. (1998) mit über 70.000 Haustransaktionen im Zeitraum 1969 bis 1991 in Fairfax County Virginia. Ausgangspunkt ist die Konstruktion einer räumlichen Matrix S und einer zeitlichen Matrix T, deren Reihensummen 1 betragen:

$$\left(\underset{(n \quad x \quad n)}{S} \underset{(n \quad x \quad 1)}{[1]} = \underset{(n \quad x \quad 1)}{[1]}, \underset{(n \quad x \quad n)}{T} \underset{(n \quad x \quad 1)}{[1]} = \underset{(n \quad x \quad 1)}{[1]} \right)$$

Die reihennormalisierten Matrizen S und T bezeichnet man auch als standardisierte, bzw. reihen-stochastische Matrizen. Man kann S und T als lineare Filter betrachten. Die Matrizen enthalten entlang der Hauptdiagonale Nullen (keine Beobachtung wird durch sich selbst erklärt) und sind triangularisiert. Beispielsweise könnte S für fünf Grundstücke, die nach dem Verkaufszeitpunkt geordnet sind, wie folgt aussehen:

$$S = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \end{bmatrix}, \; T = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Das bedeutet, für Grundstück 2 ist aus räumlicher Hinsicht nur Grundstück 1 erklärungsrelevant, für Grundstück 3 sind die Grundstücke 1 und 2 jeweils zu 50% erklärungsrelevant, für Grundstück 5 sind die Grundstücke 1 und 3 jeweils zu 50% erklärungsrelevant. Entscheidend ist, daß die Reihensummen (mit Ausnahme der Reihe 1) jeweils 1 betragen. Darüber hinaus haben wir implizit angenommen, daß nur zwei Nachbarn relevant sind. In der Matrix T wird angenommen, daß nur das zeitlich nächste vorher verkaufte Grundstück bewertungsrelevant ist. Deshalb besteht die Matrix nur aus Einsen und Nullen.

Angenommen, wir filtern zuerst nach der räumlichen Distanz zwischen den Grundstücken, dann ergibt sich folgende Interaktionsmatrix ST

$$ST = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \end{bmatrix}$$

Wenn wir aber zuerst zeitlich filtern, erhalten wir die Matrix TS

$$TS = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \end{bmatrix}$$

Selbst nach einer erneuten Normalisierung (Standardisierung) sind die Matrizen ST und TS verschieden. Das heißt, wir erhalten unterschiedliche Ergebnisse, je nachdem, nach welcher Dimension zuerst gefiltert wird. Die Konstruktion einer zeit-räumlichen Matrix erfordert also mehrere *a priori*-Festlegungen. Zum ersten ist zu entscheiden, wie viele Nachbarn in räumlicher und zeitlicher Hinsicht einbezogen werden sollen[6]. Alternativ könnte man auch fragen, wieweit die räumliche und zeitliche Dimension aufgespannt werden sollen. Im Handbuch Immobilienbewertung Österreich (Bienert und Funk 2007) wird im Abschnitt 2.2.2 beispielsweise folgendes ausgeführt:

„Aufgrund statistischer Erfahrungswerte sollten rund 7 bis 8 Vergleichswerte vorliegen“ (S.166). Weniger scharf wird der zeitliche Rahmen zum Bewertungsstichtag umrissen: „Allgemein gilt jedoch, daß der zeitliche Abstand zum Wertermittlungsstichtag – der durchaus Jahre betragen kann – den Rahmen, welcher die verläßliche Herstellung einer Preisrelation noch zuläßt, nicht überschreiten darf.“ (S. 169).

Persönliche Kommunikation mit Gutachtern hat ergeben, daß durchaus auch nur drei strukturell ähnliche Grundstücke als Vergleichswerte als ausreichend angesehen werden. Im Folgenden wurde mit räumlich 3 bis 12 Nachbarn und mit zeitlichen Verkaufsabständen von 1 und 5 Jahren experimentiert. Als Entscheidungskriterium wurde der Wert der Log-Likelihoodfunktion (LL) bzw. das Akaike-Informationskritierium (AIC) und das Bayes-

---

6 Tu et al. (2004) zeigen, wie die optimale Anzahl an zeitlichen Nachbarn p und räumlichen Nachbarn q mit Geographically Weighted Regressions (GWR) endogen bestimmt werden kann.

Schwarz-Informationskriterium (BIC) herangezogen[7]. Es zeigt sich, daß die besten Ergebnisse mit 3 räumlichen Nachbarn im Zeitrahmen von 1 Jahr erzielt werden. Verglichen werden die Ergebnisse auch mit den Ergebnissen aus der Anwendung von Standardmatrizen (nearest neighours, Delaunay). Dabei erweist sich, daß die zeit-räumlichen Matrizen zu deutlich besseren Ergebnissen führen.

3. Datengrundlagen und Schätzergebnisse

3.1 Daten und Spezifikationen

Gegenstand der Analyse sind 3.710 Transaktionen mit unbebauten Grundstücken und Abbruchobjekten mit der Hauptwidmung „Wohnbauland/gemischte Baugebiete" in den Bauklassen I bis VI im Untersuchungszeitraum von 1987 bis 2003 in Wien. Die Grundstücksdatenbasis liefert die Kaufpreissammlung der Stadt Wien. Für die daraus ausgewählten Grundstücke werden die Preisfunktionen aus den Gleichungen 4 bis 10 mittels Regressionsanalyse geschätzt. Dabei werden die mit dem Verbraucherpreisindex deflationierten Kaufpreise (reale Kaufpreise) auf die strukturellen Eigenschaften und die Lageattribute der Grundstücke regressiert. Die Lageattribute wurden den Transaktionen wie folgt zugeteilt. Für Transaktionen der Jahre 1987 bis 1995 gelten die Informationen aus der Großzählung 1991, aus den Realnutzungskatierungen 1991 und aus den Erreichbarkeiten im öffentlichen Verkehr 1991. Für Transaktionen ab 1996 wurden die entsprechenden Daten für das Jahr 2001 zugewiesen. Die folgende Aufstellung gibt einen Überblick über die wichtigsten erfassten Merkmale der Grundstücke und der Lagefaktoren.

- Strukturelle Merkmale der Grundstücke: Grundstücksfläche, Flächenform, Bauklasse, Prozentanteil der Hauptwidmung, Erwerbertypen, Veräußerertypen, Erwerbsdatum, Hangneigung, Ausrichtung – Himmelsrichtung, Parzellierung
- Merkmale der Mikro- und Makrolage: Soziodemographische Zusammensetzung der Nachbarschaft (Altersgruppen, nationale Herkunft, Akademikeranteile), Bauliche Dichte, Lärmbelastungen Strasse, Fluglärm, Ungewichtete Erreichbarkeiten im öffentlichen Verkehr nach ÖIR-Berechnungen, Entfernung zu Strassen unterschiedlicher Kategorien, Anzahl der Volksschulen im Umkreis

Die Datenbasis der Untersuchung stützt sich auf unterschiedliche Quellen: die Kaufpreissammlung der Grundstückstransaktionen in Wien (1987 bis 2003) – MA69; die Großzählungen für Wien der Statistik Austria (1991 und 2001) – bearbeitet von der MA18; Gebäude und Wohnungszählung für Wien der Statistik Austria (2001) – bearbeitet von der MA18; die Realnutzungen (1991 und 2001) – MA18; generalisierte Flächenwidmungen (1986, 1990, 1996, 2001, 2005) – MA 21A/B; Erreichbarkeiten im öffentlichen Verkehr (1991 und 2001) – ÖIR; Verortete Erreichbarkeiten, Distanzen und Externalitäten: Freiraumdistanzen, Gebäudehöhen, Lärmquellen, Strassen usw. – IFIP. Tabelle A1 im Anhang zeigt deskriptive Statistiken und Korrelationen der Variablen im Gesamtdatensatz.

Zur Überprüfung der Funktionsformen, die in Kapitel 3 beschrieben wurden, werden die Gleichungen sukzessive dichter spezifiziert, wobei die Ausgangsspezifikation nur aus den strukturellen Eigenschaften (Grundstücksfläche, Grundstücksform, Anteil Hauptwidmung, Parzellierungsdummy, Abbruchobjektdummy, Bauklassendummies, Verkäufer- und Käufertypendummies) inklussive den Verkaufsjahresdummies besteht. Folgende, der Informationsdichte nach aufsteigende Spezifikationen wurden untersucht:

- Strukturelle Faktoren mit Transaktionsjahr (***Strukturell***)
- (*Strukturell*) plus Hauptlagevariable – Zentrumsdistanz, Transdanubiendummy, Nobelbezirksdummy, ÖFFI-Erreichbarkeit (***StrukMainLage***)

[7] Die Anwendbarkeit der Likelihood ratio hängt allerdings, ebenso wie die daraus abgeleiteten Kriterien AIC und BIC, von der Annahme normalverteilter Residuen ab.

- (*StrukMainLage*) plus Sozio-ökonomische Faktoren – Familien, Oldies, Yuppies, Neubeginner, Bildung (***StruSozio***)
- (*StruSozio*) plus sekundäre Lagefaktoren - Lärm Flugzeug, Distanz zu Straßen unterschiedlicher Kategorie, Luftschadstoffproduzenten, Industriell/gewerbliche Nutzungsanteile, strukturierte Freiräume, sonstige Freiräume (***StruLowLage***)
- (*Strukturell*) plus alle Einflussfaktoren (***Full Model***)

Alle Standardformen, BoxCox-Transformationen und Cluster-Modelle wurden mit STATA 9.0 geschätzt. Die OLS-Modelle wurden mit den Optionen *Robust* und *Cluster* zur Korrektur von Heteroskedastizität mit entsprechender Korrektur der Standardfehler geschätzt. Die räumlichen Lattice-Modelle wurden mit den Programmen von LeSage (1998) in MATLAB 9.0 geschätzt. Für die Konstruktion der ST-Matrizen wurde eigens ein Programm in MATLAB erstellt.

3.2 Box-Cox-Modelle

Wegen der Unsicherheit bezüglich der funktionalen Form der hedonischen Preisfunktion wurden zunächst für die fünf Spezifikationen Box-Cox-Regressionen mit unterschiedlichen Annahmen über die Transformationsparameter geschätzt. Das STATA-Programm *boxcox* lässt vier unterschiedliche Modellvarianten zu, wobei folgende Optionen gewählt werden können: Mit der Option LHS wird nur die abhängige Variable transformiert, mit der Option RHS werden nur die (kontinuierlichen) erklärenden Variablen, mit der Option LAMBDA werden die abhängige und die (kontinuierlichen) erklärenden Variablen mit dem selben Parameter transformiert und mit der Option THETA erhalten wir zwei unterschiedliche Transformationsparameter, jeweils einen für die linke und einen für die rechte Seite der Gleichung. Die Option THETA ist die allgemeinste im STATA-Programm. Theoretisch sind noch allgemeinere vorstellbar, wobei auf der rechten Seite mehrere unterschiedliche Transformationsparameter zur Anwendung kommen könnten. Ein entsprechendes Programm war aber nicht verfügbar. Das STATA-Modul *boxcox* lässt mit der Option NOTRANS auch zu, dass auf der rechten Seite dichotome erklärende Variable (Dummies) nicht transformiert werden, was sinnvoll ist und hier auch genutzt wurde. Alle zu tranformierenden kontinuierlichen Variablen müssen andererseits überall strikt positive Realisationen aufweisen. Um negative Werte oder Nullen als Realisationen auszuschließen, mussten vorab entsprechende Transformationen durchgeführt werden.

Tabelle 1 fast die Ergebnisse für die 5 Spezifikationen (STRUKTURELL; STRUMAIN, STRUSOZIO, STRULOW und FULL) in den vier Modellvarianten (LHS; RHS, LAMDA und THETA) zusammen. In der Basisspezifikation, welche nur die struktuellen Variablen inklusive den Kaufjahresdummies enthält, konnte mit der Option THETA keine Konvergenz der Maximum-Likelihood Funktion erreicht werden. In allen anderen Varianten konnte Konvergenz erreicht werden.

Insgesamt zeigt sich, dass die Box-Cox-Variationen im STATA-Programm *boxcox* nicht ausreichen, um eine optimale Spezifikation zu finden. Der Grund liegt darin, dass die Spezifikationen zu viele Variablen (Merkmale, Charakteristika) enthalten, die unterschiedliche lineare bzw. nicht-lineare Beziehungen zum Verkaufspreis der Grundstücke aufweisen. Einige allgemeine Aussagen lassen sich dennoch treffen. Die Full-Spezifikation erhält in allen Modellvarianten die höchsten Log-Likelihood-Werte, wobei das Modell THETA den allerhöchsten Wert (-48073,326) erzielt. Von den untersuchten Funktionsformen, wird nur die Doppel-logarithmische (Theta=Lambda=0) bzw. die Semi-logarithmische (Theta=0) in einigen Fällen nicht abgelehnt. Die stärkste Unterstützung erhält die Doppel-logarithmische Form in der Modellvariante LAMBDA, d.h. in der Variante, in der beide Seiten der Gleichung mit demselben Transformationsparameter transformiert werden. Dieser Parameter ist entweder nicht von Null verschieden (Full, StrukMain, StrukLow) oder nahe bei Null (Strukturell, StrukSozio). Stark abgelehnt wird in allen Varianten die lineare Funktionsform (Theta=Lambda=1). Nur als Ergänzung wurde auch die invers-multiplikative Form (Theta=Lambda=-1) angeführt, die ebenfalls abgelehnt wird.

Als Hauptergebnisse der Box-Cox-Regressionen kann man daher festhalten, dass innerhalb der hier verglichenen Modellvarianten ein möglichst voll spezifiziertes Modell mit logarithmierten Variablen auf beiden Seiten die besten Ergebnisse erzielen wird. Die dazugehörige theoretische Ausgangsfunktion hat daher multiplikativen Charakter der Form $\mathbf{P} = K\sum_{i\in D}\exp^{b_i x_i}\sum_{j\in C} x_j^{\,b_j}\exp^{e}$ und die dazugehörige linearisierte Schätzgleichung die Form aus Gleichung (6).

Tabelle 1 Prüfergebnisse der Box-Cox-Modelle

	LHS	RHS	Lambda	Theta
Strukturell	-49873,123	-55277,924	-49162,938	keine Konvergenz
Lamda	-	0,821	0,019	
Theta	n.s.	-		
chi2 (Prob. > chi2) (H0):				
Theta=Lambda=-1	5512,24 (0,000)	1744,61 (0,000)	6154,22 (0,000)	
Theta=Lambda=0	0,56 (0,453)	938,66 (0,000)	3,87 (0,049)	
Theta=Lambda=1	10856,28 (0,000)	46,68 (0,000)	12276,65 (0,000)	
	LHS	RHS	Lambda	Theta
StrukMainLage				
Log likelihood	-49378,383	-55231,869	-48346,419	-48345,678
Lamda	-	0,816	n.s.	n.s.
Theta	n.s.	-		n.s.
chi2 (Prob. > chi2) (H0):				
Theta=Lambda=-1	5829,07 (0,000)	1808,84 (0,000)	6981,21 (0,000)	6982,69 (0,000)
Theta=Lambda=0	2,49 (0,115)	975,37 (0,000)	0,01 (0,908)	1,50 (0,221)
Theta=Lambda=1	11758,97 (0,000)	52,00 (0,000)	13822,90 (0,000)	13824,38 (0,000)
	LHS	RHS	Lambda	Theta
StrukSozio				
Log likelihood	-49310,164	-55219,341	-48483,647	-48448,834
Lamda	-	0,824	0,043	0,155
Theta	-0,023	-		0,021
chi2 (Prob. > chi2) (H0):				
Theta=Lambda=-1	5865,84 (0,000)	1775,00 (0,000)	7471,00 (0,000)	7540,63 (0,000)
Theta=Lambda=0	5,09 (0,024)	985,75 (0,000)	29,27 (0,000)	98.89 (0,000)
Theta=Lambda=1	11866,07 (0,000)	47,71 (0,000)	13519,10 (0,000)	13588,72 (0,000)
	LHS	RHS	Lambda	Theta
StrukLowLage				
Log likelihood	-49332,08	-55219,765	-48362,809	-48359,454
Lamda	-	0,820	n.s.	0,037
Theta	-0,035	-		n.s.
chi2 (Prob. > chi2) (H0):				
Theta=Lambda=-1	5696,34 (0,000)	1809,25 (0,000)	6827,63 (0,000)	6834,34 (0,000)
Theta=Lambda=0	11,20 (0,001)	984,86 (0,000)	0,50 (0,479)	7,21 (0,001)
Theta=Lambda=1	11825,17 (0,000)	49,80 (0,000)	13763,72 (0,000)	13770,42 (0,000)
	LHS	RHS	Lambda	Theta
Full Model				
Log likelihood	-49176,726	-55189,949	-48075,825	-48073,326
Lamda	-	0,821	n.s.	0,030
Theta	-0,032	-		n.s.
chi2 (Prob. > chi2) (H0):				
Theta=Lambda=-1	5884,38 (0,000)	1775,82 (0,000)	7258,02 (0,000)	7263,02 (0,000)
Theta=Lambda=0	9,47 (0,002)	983,39 (0,000)	0,15 (0,698)	5,15 (0,023)
Theta=Lambda=1	12073,91 (0,000)	47,46 (0,000)	14275,71 (0,000)	14280,71 (0,000)

Quelle: eigene Berechnungen

3.3 Lineare Modelle und Modelle mit Cluster-Effekten

Obwohl die Box-Cox-Modelle die lineare Funktionsform mit linearer Ausgangsfunktion bereits abgelehnt haben, werden hier noch einmal alle linearen Modelle unter etwas anderem Blickwinkel untersucht. Die Modellergebnisse werden dabei anhand unterschiedlicher Bewertungskriterien geprüft. Der *Ramsey Reset* Test prüft auf fehlende Variablen im Modell (Modellspezifikation). Bis auf die Full-Modelle der gemischt semi-doppellogarithmischen Form und der Cluster-Fixed Effects besteht kein Modell den Test auf dem 10%-Niveau. Das bedeutet, dass alle anderen Varianten nach diesem Test ein Spezifikationsproblem aufweisen. Das betrifft entweder fehlende Variablen oder nicht adäquat berücksichtigte Nichtlinearitäten.

Multikollinearität wurde mit Hilfe der *Varianz-Inflation-Factors (VIFs)* geprüft. Die Standard-Funktionsformen weisen kein dramatisches Problem der Multikollinearität auf. Die durchschnittlchen VIFs liegen nie höher als bei 3,4, was angesichts der Stichprobengröße von 3.710 Beobachtungen nicht problematisch erscheint. Darüber hinaus zeigen sich die meisten Variablen signifikant auf zumindest 10%-Niveau mit den theoretisch erwarteten Vorzeichen (siehe Tabelle 3). Anders liegt der Fall in den Cluster-Fixed Effects Modellen, in denen der Einbezug der Makrolagevariablen (Zentrumsdistanz, Öffi-Erreichbarkeit, Nobelbezirksdummy, Transdanubiendummy) die durchschnittlichen VIFs in die Höhe schnellen lässt. Die Korrelationen der Makrolagevariablen mit Cluster Effekten sind zum Teil sehr hoch, wodurch die Signifikanz der Parameter eingeschränkt wird.

Zum Vergleich der Prognosefähigkeit der Modelle werden die Kriterien *LL*, *AIC* und *BIC* bzw. der Root Mean Square Error (RMSE) und das adj. R^2 herangezogen[8]. Die Log-Likelihood Ziffer berücksichtigt nicht den Verlust an Freiheitsgraden und ist daher nicht für den Vergleich von unterschiedlich dicht spezifizierten Modellen geeignet. Das Akaike-Informationskriterium und das Bayes-Informationskriterium bestrafen für den Verlust an Freihheitsgraden in reicher spezifizierten Modellen. Das BIC bestraft stärker als das AIC und wird gewöhnlich bei widersprüchlichen Ergebnissen der beiden Kriterien stärker berücksichtigt. Generell gilt, je kleiner der AIC oder BIC-Wert, desto besser das Modell. Das AIC ist definiert als $-2*lnLL + 2*p$, das BIC als $-2*lnLL + p*lnN$, wobei p die Anzahl der Parameter und N die Anzahl der Beobachtungen widergeben. Einschränkend muss festgehalten werden, dass LL, AIC und BIC Normalverteilung der Residuen unterstellen, eine Annahme, die in hedonischen Bodenpreisanalysen sehr oft nicht zutrifft.

Die schlechtesten Ergebnisse gemessen an den AIC und BIC-Werten liefern die *linearen Spezifikationen mit linearer Ausgangsfunktion*. Damit werden die Box-Cox-Ergebnisse im vorigen Abschnitt bestätigt. Während das adj. R^2 von 0,6509 im struktuellen Modell auf 0,6647 im Full-Model leicht ansteigt, steigt auch der BIC-Wert, d.h. nach dem Bayes-Kriterium sollte das strukturelle Modell dem Full-Model in der einfachen linearen Spezifikation vorgezogen werden. Ein Resultat der vielen Nichtlinearitäten, die in den einfach-linearen Spezifikationen in Niveaus der Variablen nicht adäquat berücksichtigt werden.

Bei allen anderen Standardfunktionsformen präferieren das adj. R^2 und der BIC-Wert die Full-Modelle gegenüber allen anderen Spezifikationen. Die besten Ergebnisse liefert das Full Model der gemischt semi-doppellogarithmischen Form. Diese weicht nur in zwei Variablen von der doppel-logarithmischen Form ab (Distanz zu Straßen der Kategorie 2, Distanz zum Wienerwald). Die gemischt semi-doppel-logarithmische Form wird nach dem BIC-Wert allerdings dem Cluster-Fixed Effects Modell nicht vorgezogen[9]. Auch erreicht das Cluster-Fixed Effects Modell den kleinsten *Root Mean Square Error* (0,4173 gegenüber 0,4256). Hier wird deutlich, warum den Cluster Effekten in der Prognose von Häuserpreisen gegen-

---

[8] Zu berücksichtigen ist, daß beim adjustierten R^2 die Interpretierbarkeit des Bestimmheitsmaßes R^2 als Anteil der erklärten Varianz in den Preisen verloren geht. Wegen der durchdachten Spezifikationen liegt das adj. R^2 immer sehr nahe am R^2. Die Korrelation zwischen den beiden Maßen über alles Spezifikationen liegt bei 100%.

[9] Würde man die Parameter der KGs explizit in einem Dummy-Variablen-Modell schätzen, dann führte das zu einem Verlust von weiteren 81 Freiheitsgraden. In dem Fall läge der BIC-Wert im Full Modell über dem BIC-Wert der gemischt semi-doppellogarithmischen Form. Entscheidender ist daher hier, dass der RMSE des Fixed Effects Modells unter dem RMSE der gemischt semi-doppellogarithmsichen Form liegt.

über Standardfunktionsformen der Vorzug gegeben wird (Bourasse et al. 2007, Kuminoff et al. 2008). Die Cluster-Effekte bringen bereits mit wenigen zusätzlichen Lagevariablen bessere Ergebnisse (Root MSE) als ein Full Modell ohne Cluster Effekte. Will man also Grundstückspreise prognostizieren, dann reicht der Einbezug von Cluster Effekten bereits aus, um nahe an die Prognosekraft eines viel aufwändigeren, weil voll spezifizierten Modells heranzureichen. Allerdings darf nicht übersehen werden, dass Informationen über einzelne Lagecharakteristika die Progonosekraft auch in Modellen mit Cluster Effekten weiter erhöhen. Je dichter die Spezifikation, desto kleiner wird der RMSE.

Tabelle 2 – Prüfergebnisse der Standardfunktionsformen und Cluster-Effekte

Modell	Obs	**Log Likelihood**	df	**AIC**	**BIC**	**Adj. R^2**	**Root MSE**	**∅ VIF**	**Ramsey Reset**	**Ramsey R. (P-Value)**
linear/Ausgangsfunktion: linear										
Strukturell	3710	-55301,3	43	110688,5	110955,9	0,6509	7,24e+0,5	1,50	75,26	0,0000
StrukMainLage	3710	-55257,9	47	110609,7	110902,0	0,6586	7,16e+0,5	1,62	92,47	0,0000
StrukSozio	3710	-55243,2	48	110582,4	110880,9	0,6612	7,13e+0,5	1,54	90,41	0,0000
StrukLowLage	3710	-55244,7	59	110607,3	110974,2	0,6599	7,15e+0,5	2,00	103,37	0,0000
Full Model	3710	-55213,7	68	110563,4	110986,2	0,6647	7,10e+0,5	3,41	111,75	0,0000
Modell	Obs	**Log Likelihood**	df	**AIC**	**BIC**	**Adj. R^2**	**Root MSE**	**∅ VIF**	**Ramsey Reset**	**Ramsey R. (P-Value)**
semi-logarithmisch/Ausgangsfunktion: exponentiell										
Strukturell	3710	-3884,6	43	7855,2	8122,6	0,6588	0,6935	1,50	173,58	0,0000
StrukMainLage	3710	-3390,8	47	6875,7	7168,0	0,7383	0,6074	1,62	243,87	0,0000
StrukSozio	3710	-3323,9	48	6743,8	7042,3	0,7475	0,5966	1,54	255,66	0,0000
StrukLowLage	3710	-3348,9	59	6815,8	7182,7	0,7433	0,6015	2,00	259,04	0,0000
Full Model	3710	-3192,7	68	6521,3	6944,2	0,7634	0,5775	3,41	270,70	0,0000
Modell	Obs	**Log Likelihood**	df	**AIC**	**BIC**	**Adj. R^2**	**Root MSE**	**∅ VIF**	**Ramsey Reset**	**Ramsey R. (P-Value)**
doppel-logarithmisch/Ausgangsfunktion: multiplikativ										
Strukturell	3710	-3175,5	42	6435,0	6696,2	0,7672	0,5728	1,52	20,50	0,0000
StrukMainLage	3710	-2356,7	47	4807,4	5099,6	0,8501	0,4596	1,66	7,42	0,0001
StrukSozio	3710	-2351,0	48	4797,9	5096,4	0,8505	0,4590	1,56	3,19	0,0228
StrukLowLage	3710	-2361,8	59	4841,6	5208,5	0,8492	0,4610	1,65	10,56	0,0000
Full Model	3710	-2066,5	68	4268,9	4691,8	0,8711	0,4263	2,18	2,32	0,0737
Modell	Obs	**Log Likelihood**	df	**AIC**	**BIC**	**Adj. R^2**	**Root MSE**	**∅ VIF**	**Ramsey Reset**	**Ramsey R. (P-Value)**
Gemischt semi-doppel-logarithmisch/Ausgangsfunktion: multiplikativ-exponentiell										
Strukturell	3710	-3175,5	43	6437,0	6704,4	0,7672	0,5728	2,18	2,32	0,0737
StrukMainLage	3710	-2356,7	47	4807,4	5099,6	0,8501	0,4596	1,66	7,42	0,0001
StrukSozio	3710	-2351,0	48	4797,9	5096,4	0,8505	0,4590	1,56	3,19	0,0228
StrukLowLage	3710	-2327,5	59	4773,1	5140,0	0,8520	0,4568	1,73	9,54	0,0000
Full Model	3710	-2060,5	68	4257,0	4679,8	0,8715	0,4256	2,30	2,04	0,1064
Model	Obs	**Log Likelihood**	df	**AIC**	**BIC**	**Adj. R^2 (within)**	**Root MSE**	**zentrierte ∅ VIF**	**Ramsey Reset**	**Ramsey R. (P-Value)**
Group-Level-Fixed Effects (ohne Katastralgemeinde-Parameter)										
Strukturell	3710	-2047,3	43	4180,6	4448,0	0,8047	0,4274	3,73		
StrukMainLage	3710	-2013,5	47	4121,0	4413,3	0,8080	0,4237	54,32		
StrukSozio	3710	-2020,3	48	4136,7	4435,2	0,8073	0,4246	3,56		
StrukLowLage	3710	-1986,4	59	4090,9	4457,8	0,8102	0,4214	5,68		
Full Model	3710	-1946,5	68	4029,1	4451,9	0,8137	0,4173	67,80		
Model	Obs	**Log Likelihood**	df	**AIC**	**BIC**	**Wald Chi2**		**∅ VIF**		
Group-Level-FLGS-Panel(hetero) (ohne Katastralgemeinde-Parameter)										
Strukturell	3710	-2937,5	43	5961,0	6228,4	14685,98		3,73		
StrukMainLage	3710	-2122,4	47	4338,8	4631,1	25063,76		54,32		
StrukSozio	3710	-2148,0	48	4392,0	4690,5	24579,85		3,56		
StrukLowLage	3710	-2126,3	59	4370,6	4737,5	24205,35		5,68		
Full Model	3710	-1833,2	68	3802,4	4225,3	29152,80		67,80		

Quelle: eigene Berechnungen

3.4 Korrektur für Heteroskedastitzität

Eine zentrale Annahme des OLS-Modells betreffend die (konditionale) Varianz der Fehlerterme lautet $\sum_{\varepsilon} = \sigma^2 I_N$, d.h. die Varianz der Fehlerterme ist konstant (homoskedastisch). Wenn dies nicht der Fall ist, d.h. wenn gilt: $\sum_{\varepsilon} \neq \sigma^2 I_N$, dann ist der OLS-Schätzer der Parameter $\boldsymbol{\beta}$ zwar weiterhin unverzerrt, konsistent und (asymptotisch) normalverteilt, jedoch nicht länger effizient (vgl. Baum, 2006, Kap. 6). Im Zusammenhang mit dem Bodenpreismodell sind zwei Fälle von Heteroskedastizität von Interesse: Heteroskedastizität quer über alle Beobachtungen und Heteroskedastizität quer über die Beobachtungen innerhalb von vordefinierten Clustern. Beispielsweise kann es sein, dass die Fehlerterme systematisch mit dem Kaufpreisniveau korrelieren (erster Fall) oder dass die Fehlerterme sich systematisch nach räumlichen Clustern unterscheiden, weil wichtige Informationen im Modell fehlen (zweiter Fall, unechte Heroskedasizität).

Zur Korrektur des ersten Falls wird der Huber-White-Sandwich-Schätzer herangezogen (Option *robust* in STATA), im zweiten Fall kommt der Cluster-robuste-VCE-Schätzer zur Anwendung (Option *cluster* in STATA). Der Cluster-robuste Schätzer ist gewöhnlich konservativer als der Huber-White Schätzer, d.h. er führt zu höheren Standardfehlern bzw. kleineren T-Werten. Beide Schätzer verändern nur die Varianz der Schätzungen, nicht aber die Punktschätzung, d.h. die Koeffizienten bleiben unverändert.

Ausgangspunkt ist hier das Full-Modell der gemischt semi-doppellogarithmischen Form. Heteroskedastizität des einfachen OLS-Schätzers wird mit dem Breuch-Pagan-Test untersucht. Die Standardversion des Tests in Stata (*hettest*) verwendet die Schätzwerte der abhängigen Variablen ($\hat{y}_i$, fitted values) und regressiert diese auf die quadrierten Residuen der Regression. Unter der Annahme der Homoskedastizität, $\sum_{\varepsilon} = \sigma^2 I_N$, sollte diese Regression keinen Erklärungswert haben (kleines R^2, kleine F-Statistik). Die Teststatistik ist Chi^2 verteilt. Der Test zeigt für das Basismodell signifikante Anzeichen von Heteroskedastizität: $Chi^2(1) = 69{,}03$ ($Prob>chi^2 = 0{,}0000$).

Die Ergebnisse in Tabelle 3 zeigen, dass der Cluster-Schätzer die konservativsten Standardfehler bzw. P-Werte produziert. In allen, bis auf einen Fall, verschlechtert sich die Signifikanz der Koeffizienten. Ausnahme ist die Flächenform, deren Koeffizient erst durch die Korrektur signifikant wird. Nicht mehr oder weiterhin nicht signifikant auf dem 10%-Niveau sind der Kaufpreisindex 1988, die Dummy für Abbruchobjekte, die Öffi-Erreichbarkeit, die sozio-demographischen Faktoren Neubeginner und Oldies, die Distanz zur Lobau und die Dummy für Nähe zu Luftschadstoffproduzenten. Der Cluster-Schätzer ist jedoch nicht unabhängig davon, ob man Katastralgemeinden oder Zählsprengel als Cluster definiert. Der Datensatz umfasst Grundstücke in 705 Zählsprengeln, die sich wiederum auf 82 Katastralgemeinden aufteilen. Die Besetzung nach Zählsprengeln variert von 1 bis 70, jene nach Katastralgemeinden von 1 bis 319 Beobachtungen. Die Zählsprengelebene geht also deutlich tiefer. Dennoch gibt es nur wenige merkbare Abweichungen zwischen den beiden Cluster-Schätzern. Lobau und ÖFFI-Erreichbarkeit sind im ZSP-Cluster signifikant, die Donauinseldummy verliert an Signifikanz. Das Clustern auf kleinräumigerer Ebene hat also nicht nur Wirkungen in eine Richtung.

Tabelle 3 Korrektur für Heteroskedastizität

Variablen	Koeffizienten	P-Werte der Koeffizienten (OLS)			
		OLS-nonrobust	OLS-robust	OLS-ClusterKG	OLS-ClusterZSP
N	3710	3710	3710	3710	3710
Bauklassen (Ds)					
II	0,175	0,0000	0,0000	0,0000	0,0000
III	0,436	0,0000	0,0000	0,0000	0,0000
IV	0,739	0,0000	0,0000	0,0000	0,0000
V	1,173	0,0000	0,0000	0,0000	0,0000
VI	0,780	0,0000	0,0000	0,0008	0,0000
Kaufpreisindex (Ds)					
1988	0,055	**0,1486**	**0,1062**	**0,1088**	**0,1582**
1989	0,141	0,0005	0,0003	0,0021	0,0012
1990	0,267	0,0000	0,0000	0,0000	0,0000
1991	0,441	0,0000	0,0000	0,0000	0,0000
1992	0,590	0,0000	0,0000	0,0000	0,0000
1993	0,680	0,0000	0,0000	0,0000	0,0000
1994	0,682	0,0000	0,0000	0,0000	0,0000
1995	0,735	0,0000	0,0000	0,0000	0,0000
1996	0,709	0,0000	0,0000	0,0000	0,0000
1997	0,648	0,0000	0,0000	0,0000	0,0000
1998	0,656	0,0000	0,0000	0,0000	0,0000
1999	0,699	0,0000	0,0000	0,0000	0,0000
2000	0,635	0,0000	0,0000	0,0000	0,0000
2001	0,646	0,0000	0,0000	0,0000	0,0000
2002	0,676	0,0000	0,0000	0,0000	0,0000
2003	0,625	0,0000	0,0000	0,0000	0,0000
Strukturell					
Grundstücksfläche (E)	0,884	0,0000	0,0000	0,0000	0,0000
Flächenform (SE)	-0,032	**0,2347**	0,0183	**0,0199**	0,0206
Hauptwidmung (SE)	0,002	0,0012	0,0095	0,0248	0,0140
Abbruchobjekt (D)	-0,026	**0,2741**	**0,3170**	**0,3312**	**0,3522**
Parzelliert (D)	0,086	0,0000	0,0001	0,0008	0,0001
Hauptlage					
Transdanubien (D)	-0,111	0,0151	0,0163	0,0801	**0,1250**
Nobelgebiete (D)	0,240	0,0000	0,0000	0,0000	0,0000
Öffi-Erreichbarkeit (E)	-0,206	0,0226	0,0259	**0,1047**	0,0870
Zentrumsdistanz (E)	-0,312	0,0000	0,0000	0,0000	0,0000
Sozio-ökonomische Faktoren (SE)					
Fneubeginn	-0,029	0,0400	0,0418	**0,1592**	**0,1398**
Fbildung	-0,117	0,0000	0,0000	0,0000	0,0000
Ffamilien	-0,027	0,0093	0,0138	0,0368	0,0397
Fyuppies	0,044	0,0001	0,0002	0,0020	0,0034
Foldies	-0,016	**0,1448**	**0,1625**	**0,3528**	**0,3012**
Sonstige Lagefaktoren					
Distanz Wienerwald (SE)	-0,013	0,0039	0,0038	0,0616	0,0594
Distanz Lobau (E)	-0,035	0,0143	0,0037	**0,2110**	0,0802
Prater (D)	0,136	0,0000	0,0000	0,0055	0,0006
Donauinsel (D)	0,044	0,0323	0,0346	0,0888	**0,1143**
Schönbrunn (D)	0,140	0,0031	0,0042	0,0562	0,0916
Straßen Kat2 (SE)	0,022	0,0020	0,0010	0,0472	0,0247
Straßen Kat3 (E)	0,026	0,0003	0,0005	0,0024	0,0049
Fluglärm 1 (D)	-0,153	0,0000	0,0000	0,0001	0,0014
Fluglärm 2 (D)	-0,103	0,0031	0,0052	0,0052	0,0178
Luftverpester (D)	-0,059	**0,1212**	**0,1167**	**0,1200**	**0,1408**
Einkauf (D)	0,124	0,0000	0,0000	0,0185	0,0008
Schulen (SE)	0,019	0,0188	0,0459	0,0666	0,0936
TypB (D)	0,119	0,0000	0,0000	0,0003	0,0004
Betriebs- und Industriegebiete (SE)	-0,001	0,0017	0,0005	0,0037	0,0066
Höherwertige Freiräume (E)	-0,031	0,0001	0,0001	0,0142	0,0125
Sonstige Freiräume (E)	0,025	0,0007	0,0003	0,0047	0,0033

Anmerkung: OLS-robust verwendet den Huber-White-sandwich Schätzer, OLS-Cluster berücksichtigt Heteroskedastizität innerhalb von Clustern.

Quelle: eigene Berechnungen

3.5 Lattice-Modelle

Die Lattice-Modelle berücksichtigen räumliche Interdependenzen in anderer Form als die Clustereffekt-Modelle. Die räumlichen Interdependenzen zwischen den Kaufpreisen von Grundstücken werden allein durch Nähe bzw. Distanz, unabhängig von der Zugehörigkeit zu vordefinierten räumlichen Clustern bestimmt. Diese Vorgehensweise erscheint auf den ersten Blick theoretisch konsistenter, da vordefinierte Cluster zumeist auf verwaltungstechnisch räumlichen Einteilungen (Volkszählung, ehemalige Verwaltungseinheiten) beruhen, die zumeist aufgrund der Datenlage sehr groß ausfallen müssen und zudem nicht unbedingt einen Bezug zur Kleinräumigkeit der Bodenmärkte haben. Andererseits haben auch Lattice-Modelle den Nachteil, dass sie die rasch wechselnde Qualität zwischen Quartieren „um die Ecke“ nicht adäquat abbilden können, da sie nur auf (euklidischen) Distanzen beruhen.

In Tabelle 4 sind die Ergebnisse zu den räumlich-ökonometrischen Modellen zusammengefasst. Basisspezifikation bildet das OLS-Full Modell der gemischt semi-doppel-logarithmischen Form. Dieses wurde angereichert durch Gewichtungsmatrizen der Standardform (Nearest Neighbour, Delaunay) und mit unterschiedlichen Raum-Zeit Matrizen (ST-Modelle). Die ST-Modelle unterscheiden sich nach der maximalen Anzahl an räumlichen Nachbarn (3, 4, 8 oder 12) und der Zeitdimension (1 oder 5 Jahre). Beispielsweise wurden in der ersten Zeile nur 3 Grundstücke berücksichtigt, die innerhalb eines Jahres vor dem betreffenden Grundstück verkauft wurden. Für diese 3 Nachbarn gilt allerdings keine räumliche Restriktion. In den meisten Fällen liegt die maximale Distanz nicht über 2 km.

Der Likelihood-Ratio Test zeigt für alle Gewichtungsmatrizen an, dass die Residuen der OLS-Regression räumlich autokorreliert sind. Der Korrelationsparameter ρ ist auch in allen SAR-Modellen signifikant von Null verschieden. Allerdings zeigt sich in allen Fällen, dass eine SAC-Spezifikation besser arbeitet, weshalb hier nur SAC-Spezifikationen angeführt sind (Bestes Modell). Bei der Auswahl des besten Modells wurde darauf geachtet, welches Modell den höchsten Loglikelihood-Wert hat und ob die Parameter ρ (SAR) bzw. ρ und λ (SAC) signifikant von Null verschieden sind und dabei das richtige Vorzeichen aufweisen[10]. Im Falle der SAC-Modelle wurden drei verschiedene Spezifikationen getestet. Beim Modell SAC(W, W) wird dieselbe Gewichtungsmatrix für den Spatial Lag und den Spatial Error verwendet, beim Modell SAC(W, W2) wird die Matrix W für den Spatial Lag und das Produkt aus der Matrix W (W2 = W*W) für den Spatial Error verwendet. Umgekehrt liegt der Fall bei der Spezifikation SAC(W2, W). Das Produkt W2 ist Ausdruck von Zweitrundeneffekten, d.h. weitergehenden räumlichen (Wechsel-)Wirkungen (siehe Lesage, 1998).

In den meisten Fällen erhält man mit der Spezifikation SAC(W, W2) die besten Ergebnisse, d.h. der Spatial Lag ist durch die Matrix W festgelegt und im Spatial Error zeigen sich Zweitrundeneffekte. Das beste Ergebnis, gemessen am Log Likelihood-Wert zeigt das Modell mit 3 Nachbarn und einer Zeitrestriktion von 1 Jahr. Der Spatial Lag (ρ) ist signifikant, aber mit 0,013 sehr gering. Das bedeutet, dass die erklärenden Variablen im OLS-Basismodell die räumliche Autokorrelation sehr gut abfangen und nur mehr ein Restwert von zwar signifikanter, aber geringer Bedeutung bleibt. Der Spatial Error dagegen ist mit 0,29 relativ groß, was wiederum dahingehend interpretiert werden kann, dass auf kleinsträumiger Ebene nicht beobachtete Faktoren auf die Preise benachbarter Grundstücke einwirken. Das Modell erklärt mehr als 90% der Varianz in den Baulandpreisen und erzielt einen Mean Absolute Prediction Error (MAPE) von 0,275, d.h. der Durchschnitt der absoluten Differenz zwischen Prognose und realem Preis beträgt 0,275. Bezogen auf den Logarithmus des durchschnittlichen Kaufpreises von 13,238 beträgt der MAPE nur rund 2,1%. Das Modell erzielt aber auch den niedrigsten RMSE aller Spezifikationen und wird daher hier als das beste Modell bezeichnet. Im Vergleich zum OLS-Modell verbessern sich der MAPE und der RMSE um jeweils 13,3%. Die Standard-Matrizen-Modelle (NN und Delaunay) erzielen dagegen nur Verbesserungen von maximal etwa 2%.

---

[10] Ein negativer Wert für ρ würde bedeuten, daß der Preis eines Grundstücks negativ mit dem Preis umliegender Grundstücke korreliert ist, was theoretisch nicht haltbar ist.

Tabelle 4 Prüfergebnisse für räumliche Modelle mit unterschiedlichen Gewichtungsmatrizen (N = 3710)

Matrix			LR Value der OLS Residuen	Bestes Modell	Log-Likelihood	ρ (Spatial Lag)	λ (Spatial Error)
max. Anzahl Nachbarn	Zeitliche Restriktion	Räumliche Restriktion					
ST-Modelle							
3	0>t<365 (1 Jahr)	keine	37,5	SAC(W,W2)	1.737,8	0,013	0,289
3	0>t<1.825 (5 Jahre)	keine	44,6	SAC(W,W2)	1.639,9	0,025	0,225
4	0>t<365	keine	45,7	SAC(W,W2)	1.698,3	0,015	0,318
4	0>t<1.825	keine	53,6	SAC(W,W2)	1.621,4	0,016	0,252
8	0>t<365	keine	63,2	SAC(W2,W)	1.373,6	0,003	0,154
8	0>t<1.825	keine	69,4	SAC(W,W2)	1.562,7	0,039	0,290
12	0>t<365	keine	66,3	SAC(W2,W)	1.375,0	0,003	0,191
12	0>t<1.825	keine	82,0	SAC(W,W2)	1.532,2	0,061	0,298
NN-Modelle							
3	keine	keine	58,7	SAC(W2,W)	1.378,8	0,003	0,136
4	keine	keine	57,0	SAC(W2,W)	1.381,7	0,005	0,158
8	keine	keine	54,9	SAC(W2,W)	1.379,3	0,006	0,206
12	keine	keine	43,2	SAC(W2,W)	1.371,0	0,005	0,229
Delaunay	keine	keine	44,3	SAC(W,W)	1.370,1	0,006	0,169
OLS	keine	keine		OLS	-2.060,5		
						-	-

Matrix			Adj. R2	Mean Absolute Prediction Error (MAPE)	Root Mean Square Error (RMSE)	MAPE-Verbesserung geg. OLS in %	RMSE-Verbesserung geg. OLS in %
max. Anzahl Nachbarn	Zeitliche Restriktion	Räumliche Restriktion					
ST-Modelle							
3	0>t<365 (1 Jahr)	keine	0,901	0,275	0,369	13,3	13,3
3	0>t<1.825 (5 Jahre)	keine	0,894	0,285	0,382	10,4	10,1
4	0>t<365	keine	0,899	0,279	0,375	12,2	11,9
4	0>t<1.825	keine	0,893	0,287	0,385	9,7	9,5
8	0>t<365	keine	0,874	0,314	0,418	1,2	1,9
8	0>t<1.825	keine	0,888	0,293	0,393	7,6	7,6
12	0>t<365	keine	0,874	0,313	0,417	1,4	1,9
12	0>t<1.825	keine	0,886	0,297	0,398	6,5	6,5
NN-Modelle							
3	keine	keine	0,874	0,313	0,417	1,6	2,0
4	keine	keine	0,875	0,312	0,417	1,8	2,1
8	keine	keine	0,874	0,313	0,417	1,5	2,0
12	keine	keine	0,874	0,314	0,418	1,3	1,7
Delaunay	keine	keine	0,874	0,314	0,418	1,3	1,8
OLS	keine	keine	0,871	0,318	0,426		

Anmerkung: NN-Modelle basieren auf nearest neighbour matrizen ohne Berücksichtigung der zeitlichen Dimension; ST-Modelle berücksichtigen nur Nachbargrundstücke, die vorher verkauft wurden (räumliche und zeitliche Dimension)

Quelle: eigene Berechnungen

3.6 Parametersensitivitäten

Die geschätzten Parameter in den unterschiedlichen Modellen sind nicht immer direkt vergleichbar, da ihre Dimension im Fall kontinuierlicher Variablen davon abhängt, wie die abhängige Variable und die erklärende Variable selbst gemessen sind, d.h., hier insbesondere, ob sie logarithmiert wurden oder nicht. In der Tabelle 3 sind nur Parameter aufgeführt, die der Dimension nach vergleichbar sind. Auf die Darstellung der Ergebnisse aus den einfach-linearen Modellen wurde verzichtet. Die Sensitivität der Parameter im Hinblick auf die Spezifikation und das Schätzverfahren wird hier durch den Variationskoeffizienten gemessen. Dieser erlaubt allerdings nur einen groben Vergleich der Größenordnung nach. Als grobe Klassifikation könnte man hier einführen: sehr stabil (<0,1), stabil (0,1-0,2), sensitiv (0,2-0,5), sehr sensitiv (>0,5). Diese Einteilung ist natürlich ad hoc und könnte auch anders gewählt werden.

Tabelle 5 Variation der Parameter in den OLS- und Cluster-Spezifikationen

Variablentypen	Variablen	OLS - Semi-logarithmisch	OLS - Doppel-logarithmisch	OLS - Gemischt Semi-Doppel-logarithmisch	Gemischt Semi-Doppel-logarithmisch - FGLS Panel (hetero)	Gemischt Semi-Doppel-logarithmisch - Fixed Effects	Absoluter Variations-koeffizient (inkl. Fixed Effects Modell)	Absoluter Variations-koeffizient (exkl. Fixed Effects Modell)
Strukturell	Grundstücksfläche (E)		0,8812	0,8837	0,8885	0,8874	0,00	0,00
	Flächenform (SE)		-0,0351	-0,0317	-0,0329	-0,0379	0,08	0,05
	Hauptwidmung (SE)	-0,0027	0,0017	0,0018	0,0016	0,0018	2,36	3,67
	Abbruchobjekt (D)	-0,0339	-0,0255	-0,0260	-0,0124	-0,0268	0,31	0,36
	Parzelliert (D)	0,0295	0,0838	0,0864	0,1118	0,0849	0,38	0,44
	vc1	-0,3477	-0,2954	-0,3018	-0,2877	-0,2956	0,08	0,09
	vc2	0,1022	-0,1671	-0,1699	-0,2117	-0,1989	1,01	1,29
	vc3	-0,1487	0,0420	0,0478	-0,0036	0,0582	**98,95**	5,86
	vc4	0,2018	-0,1580	-0,1491	-0,1908	-0,1923	1,73	2,50
	vc5	0,1465	-0,3045	-0,3083	-0,3304	-0,3349	0,92	1,16
	vc6	0,2860	-0,0687	-0,0580	-0,0790	-0,0632	**46,36**	8,85
	vc7	-0,3998	-1,0965	-1,1131	-1,2352	-1,1702	0,34	0,39
	vc8	0,1581	0,0545	0,0521	0,0547	0,0633	0,60	0,65
	ec1	-0,3273	-0,3457	-0,3460	-0,1702	-0,2998	0,25	0,29
	ec2	0,6873	1,1824	1,1954	1,2259	1,2416	0,21	0,24
	ec3	0,7140	0,2388	0,2347	0,2603	0,2300	0,63	0,65
	ec4	0,4838	0,2430	0,2297	0,1905	0,1780	0,47	0,46
	ec5	0,9005	0,3097	0,2899	0,3223	0,1742	0,72	0,65
	ec6	0,7942	0,1354	0,1309	0,1476	0,0997	1,14	1,09
	ec7	0,1838	-0,5964	-0,6019	-0,2517	-0,5635	0,93	1,17
	ec8	0,6451	0,3282	0,3267	0,3233	0,3239	0,37	0,39
	Bauklasse II (D)	0,2041	0,1628	0,1750	0,1665	0,1726	0,09	0,11
	Bauklasse III (D)	0,3662	0,4347	0,4356	0,4332	0,4071	0,07	0,08
	Bauklasse IV (D)	0,7947	0,7423	0,7394	0,7126	0,7013	0,05	0,05
	Bauklasse V (D)	1,3546	1,1882	1,1731	1,2022	1,1537	0,07	0,07
	Bauklasse VI (D)	1,2842	0,7899	0,7802	0,8466	0,6672	0,27	0,26
	1988 (D)	0,0608	0,0552	0,0552	0,0513	0,0631	0,08	0,07
	1989 (D)	0,1870	0,1409	0,1408	0,1352	0,1488	0,14	0,16
	1990 (D)	0,2483	0,2708	0,2672	0,2511	0,2940	0,07	0,04
	1991 (D)	0,3605	0,4464	0,4409	0,4146	0,4455	0,09	0,09
	1992 (D)	0,5801	0,5938	0,5902	0,5424	0,5895	0,04	0,04
	1993 (D)	0,6474	0,6805	0,6796	0,6508	0,6841	0,03	0,03
	1994 (D)	0,6705	0,6802	0,6823	0,6564	0,6915	0,02	0,02
	1995 (D)	0,7694	0,7342	0,7349	0,7022	0,7237	0,03	0,04
	1996 (D)	0,6894	0,7085	0,7085	0,7206	0,7090	0,02	0,02
	1997 (D)	0,6531	0,6436	0,6484	0,6614	0,6516	0,01	0,01
	1998 (D)	0,5884	0,6566	0,6558	0,6633	0,6548	0,05	0,05
	1999 (D)	0,6792	0,6937	0,6989	0,7026	0,6898	0,01	0,01
	2000 (D)	0,6164	0,6266	0,6348	0,6571	0,6445	0,02	0,03
	2001 (D)	0,6102	0,6486	0,6461	0,6622	0,6609	0,03	0,03
	2002 (D)	0,6257	0,6763	0,6760	0,7111	0,7047	0,05	0,05
	2003 (D)	0,6333	0,6209	0,6251	0,6375	0,6214	0,01	0,01
MainLage	Transdanubien (D)	-0,2099	-0,0942	-0,1114	-0,1313	0,2821	3,64	0,37
	Nobelgebiete (D)	0,2042	0,2388	0,2398	0,2781	0,2372	0,11	0,13
	Öffi-Erreichbarkeit (E)		-0,2717	-0,2056	-0,1435	-0,2637	0,27	0,31
	Zentrumsdistanz (E)		-0,2596	-0,3120	-0,3616	-0,3649	0,15	0,16
Sozio	Fneubeginn (SE)	-0,0636	-0,0329	-0,0286	-0,0248	-0,0111	0,60	0,47
	Fbildung (SE)	-0,1711	-0,1388	-0,1170	-0,1090	-0,0798	0,28	0,21
	Ffamilien (SE)	-0,0341	-0,0254	-0,0267	-0,0358	-0,0206	0,22	0,17
	Fyuppies (SE)	0,0591	0,0520	0,0438	0,0478	0,0275	0,26	0,13
	Foldies (SE)	-0,0415	-0,0189	-0,0165	-0,0090	-0,0206	0,57	0,65
LowLage	Prater (D)	0,0463	0,1448	0,1355	0,1040	0,1560	0,38	0,41
	Donauinsel (D)	0,0895	0,0393	0,0441	0,0438	-0,0182	0,96	0,44
	Schönbrunn (D)	0,1994	0,1618	0,1399	0,0988	0,0915	0,32	0,28
	Fluglärm 1 (D)	-0,1984	-0,1991	-0,1525	-0,1485	-0,0991	0,26	0,16
	Fluglärm 2 (D)	-0,0660	-0,1263	-0,1027	-0,0773	-0,0421	0,39	0,29
	Luftverpester (D)	0,0199	-0,0446	-0,0593	-0,0654	-0,0328	0,93	1,05
	Einkauf (D)	0,1855	0,1556	0,1239	0,1263	0,1375	0,17	0,20
	Schulen (SE)	-0,0005	0,0213	0,0189	0,0108	0,0077	0,76	0,78
	TypB (D)	0,0562	0,1136	0,1187	0,0999	0,0454	0,39	0,29
	Betriebs- und Industriegebiete (SE)	-0,0007	-0,0013	-0,0011	-0,0011	-0,0010	0,21	0,24
	Distanz Lobau (E)		-0,0100	-0,0346	-0,0309	-0,0232	0,44	0,53
	Distanz Straßen Kat3 (E)		0,0279	0,0256	0,0211	0,0183	0,19	0,14
	Distanz Höherwertige Freiräume (E)		-0,0311	-0,0313	-0,0259	-0,0276	0,09	0,10
	Distanz Sonstige Freiräume (E)		0,0268	0,0248	0,0245	0,0231	0,06	0,05
	Distanz Straßen Kat2 (SE)	0,0227		0,0220	0,0214	-0,0073	1,00	0,03
	Distanz Wienerwald (SE)	-0,0349		-0,0130	-0,0163	0,0003	0,91	0,55

Anmerkung: D = Dummy, E = Elastizität, SE = Semi-Elastizität

Quelle: eigene Berechnungen

Zusätzlich zum Fixed Effects Modell werden auch die Ergebnisse eines Paneldatenschätzers (XTGLS) angeführt, der die Heterogenität der Cluster (Katastralgemeinden) berücksichtigt. Die Ergebnisse dieses Schätzverfahrens liegen in den meisten Fällen näher an den OLS-Schätzern als das Fixed Effects Modell. Das Fixed Effects-Modell produziert vor allem bei jenen Variablen abweichende Werte, deren Ausprägungen stark asymmetrisch im Raum verteilt sind. Als Beispiel seien hier die Transdaubien Dummy und die Fluglärmdummys genannt. Im Fixed-Effects Modell erhalten die Grundstücke im 21. und 22. Bezirk (Transdanubinen) nach Berücksichtigung aller anderen Einflussfaktoren und der Fixed Effects einen Aufschlag gegenüber dem Durchschnitt in der Stadt von etwa 30%. Ein für Kenner des Wiener Bodenmarktes völlig unhaltbares Ergebnis, das als Resultat der Multikollinearität zu interpretieren ist. Das XTGLS-Verfahren dagegen errechnet einen Abschlag von rund 13%, was nahe an den Ergebnissen der OLS-Schätzungen liegt und wesentlich plausibler erscheint. Auch die Abschläge für die Fluglärmzonen liegen im XTGLS-Verfahren näher an den OLS-Werten als im Fixed Effects-Schätzer. Anders liegt der Fall bei den Nobelbezirksgebieten oder bei der einwohnergewichteten ÖFFI-Erreichbarkeit. Die Aufschläge in den Nobelgebieten werden durch XTGLS höher, die Abschläge einer schlechteren Erreichbarkeit niedriger bewertet.

Ohne Berücksichtigung der Fixed Effects-Schätzer müssten die Parameter folgender Charakteristika als sehr sensitiv angesehen werden: Anteil der Hauptwidmung, Faktor Oldies, Luftverpester, Schulen, Distanz zur Lobau und Distanz zum Wienerwald. Bei einer Untersuchung der Wirkungen von sehr hochrangigen Freiräumen, von Schulen und von Schadstoffemittenten auf die Grundstückspreise ist daher der Spezifikation und des gewählten Schätzverfahrens besondere Aufmerksamkeit zu widmen. Durch entsprechende Modellierung bzw-Feinadjustierung der Nichtlinearitäten, insbesondere mit Hilfer besserer bzw. genauerer Daten, sollten die Abweichungen aber kleiner werden.

3.7 OLS vs. Lattice-Modell

Im Folgenden wird die präferierte OLS-Spezifikation (gemischt semi-doppel-logarithmisch) dem Ergebnis des besten räumlich-ökonometrischen Modells gegenübergestellt und der Frage nachgegangen, in wieweit die Koeffizienten der OLS-Schätzung durch die Nichtberücksichtigung der räumlichen Interdependenzen und kleinräumiger Variablen (Spatial Lag und Spatial Error) verzerrt sind (Tabelle 6).

Der Mittelwert der absoluten Abweichung der Parameterwerte zwischen den Full-Modellen aus OLS und STM-Schätzung beträgt 5,6%. In Einzelfällen *überschätzt* das OLS-Modell die Parameter um mehr als 30% (Distanz Wienerwald, Donauinsel). Am stärksten *unterschätzt* werden die negativen Bruttoeinflüsse der Straßen der Kategorie 2 (mehr Lärm, Feinstaub, Abgase gegenüber besserer Erreichbarkeiten), eine ungünstige Flächenform der Grundstücke (Verhältnis von Umfang zu Fläche) und der Aufschlag für parzellierte gegenüber nicht parzellierten Grundstücken. Die Bauklasseneffekte werden im OLS-Modell generell überschätzt, während beim Kaufpreisindex Über- und Unterschätzungen in den einzelnen Jahren vorliegen. Insgesamt weist das STM-Modell über den gesamten Zeitraum 1987 bis 2003 eine stärkere Preissteigerung aus als das OLS-Modell. Im Jahr 2003 liegt der Index mit 190,2 um rund 3,5 Prozentpunkte höher als im OLS-Modell (186,8). Der Parameter der Grundstücksfläche bleibt unverändert, jener der Flächenform steigt im STM-Modell, ist aber nicht mehr auf 10%-Niveau signifikant. Dagegen steigt die Signifikanz der Hauptwidmung und der Parzellierung. Die Dummy für Abbruchobjekte bleibt insignifikant auf 10%-Niveau. Die Signifikanz der Hauptlagevariablen steigt im STM-Modell generell, die Parameter unterscheiden sich jedoch um weniger als 10% von den OLS-Parametern.

Tabelle 6 Vergleich der Parameterwerte aus OLS-Full und STM-SAC(W,W2) Full (3NN, 365)

Variablen	Koeffizienten		Werte (Auf- und Abschläge geg. Referenz in %)		Abweichung OLS-STM in %	P-Werte der Koeffizienten	
	OLS-Full	STM-Full	OLS-Full	STM-Full		OLS-robust	STM-Full
N	3710	3710	3710	3710	3710	3710	
Bauklassen (Ds)							
II	0,175	0,164	19,1	17,8	-6,8	0,000	0,000
III	0,436	0,409	54,7	50,5	-7,5	0,000	0,000
IV	0,739	0,712	109,4	103,8	-5,1	0,000	0,000
V	1,173	1,126	223,2	208,3	-6,6	0,000	0,000
VI	0,780	0,783	118,1	118,8	0,6	0,000	0,000
Kaufpreisindex (Ds)							
1988	0,055	0,074	105,7	107,7	1,9	**0,106**	**0,111**
1989	0,141	0,161	115,1	117,4	2,0	0,000	0,001
1990	0,267	0,293	130,6	134,0	2,6	0,000	0,000
1991	0,441	0,445	155,4	156,1	0,4	0,000	0,000
1992	0,590	0,589	180,4	180,3	-0,1	0,000	0,000
1993	0,680	0,698	197,3	200,9	1,8	0,000	0,000
1994	0,682	0,677	197,8	196,8	-0,5	0,000	0,000
1995	0,735	0,743	208,5	210,2	0,8	0,000	0,000
1996	0,709	0,705	203,1	202,4	-0,4	0,000	0,000
1997	0,648	0,652	191,3	192,0	0,4	0,000	0,000
1998	0,656	0,647	192,7	191,0	-0,8	0,000	0,000
1999	0,699	0,705	201,2	202,3	0,6	0,000	0,000
2000	0,635	0,628	188,7	187,3	-0,7	0,000	0,000
2001	0,646	0,666	190,8	194,6	2,0	0,000	0,000
2002	0,676	0,686	196,6	198,6	1,0	0,000	0,000
2003	0,625	0,643	186,8	190,2	1,8	0,000	0,000
Strukturell							
Grundstücksfläche (E)	0,884	0,880			-0,5	0,000	0,000
Flächenform (SE)	-0,032	-0,038			18,5	0,018	0,124
Hauptwidmung (SE)	0,002	0,002			0,1	0,010	0,001
Abbruchobjekt (D)	-0,026	-0,026	-2,6	-2,6	0,5	**0,317**	**0,264**
Parzelliert (D)	0,086	0,098	9,0	10,3	14,2	0,000	0,000
Hauptlage							
Transdanubien (D)	-0,111	-0,120	-10,5	-11,3	7,5	0,016	0,012
Nobelgebiete (D)	0,240	0,225	27,1	25,2	-7,0	0,000	0,000
Öffi-Erreichbarkeit (E)	-0,206	-0,208			1,3	0,026	0,004
Zentrumsdistanz (E)	-0,312	-0,338			8,2	0,000	0,000
Sozio-ökonomische Faktoren (SE)							
Fneubeginn	-0,029	-0,021			-25,1	0,042	0,126
Fbildung	-0,117	-0,117			-0,4	0,000	0,000
Ffamilien	-0,027	-0,028			6,8	0,014	0,005
Fyuppies	0,044	0,048			9,5	0,000	0,000
Foldies	-0,016	-0,015			-10,6	**0,163**	**0,194**
Sonstige Lagefaktoren							
Distanz Wienerwald (SE)	-0,013	-0,009			-31,2	0,004	0,068
Distanz Lobau (E)	-0,035	-0,028			-18,7	0,004	0,067
Prater (D)	0,136	0,151	14,5	16,3	12,2	0,000	0,000
Donauinsel (D)	0,044	0,030	4,5	3,0	-33,0	0,035	0,179
Schönbrunn (D)	0,140	0,137	15,0	14,7	-2,2	0,004	0,004
Distanz Straßen Kat2 (SE)	0,022	0,027			24,9	0,001	0,000
Distanz Straßen Kat3 (E)	0,026	0,024			-7,5	0,001	0,001
Fluglärm 1 (D)	-0,153	-0,169	-14,1	-15,6	10,0	0,000	0,000
Fluglärm 2 (D)	-0,103	-0,099	-9,8	-9,4	-3,3	0,005	0,005
Luftverpester (D)	-0,059	-0,058	-5,8	-5,6	-2,6	**0,117**	**0,141**
Einkauf (D)	0,124	0,129	13,2	13,8	4,5	0,000	0,000
Schulen (SE)	0,019	0,019			-1,3	0,046	0,022
TypB (D)	0,119	0,111	12,6	11,7	-7,3	0,000	0,000
Betriebs- und Industriegebiete (SE)	-0,001	-0,001			-19,3	0,001	0,017
Distanz Höherwertige Freiräume (E)	-0,031	-0,032			2,0	0,000	0,000
Distanz Sonstige Freiräume (E)	0,025	0,023			-7,5	0,000	0,002

Anmerkung: OLS-robust verwendet den Huber-White-Sandwhich Schätzer; D = Dummy, Ds = Dummies, E = Elastizität, SE = Semi-Elastizität

Quelle: eigene Berechnungen

4. *Schlussfolgerungen*

Die hedonische Bodenpreisanalyse wird zunehmend zu einem wichtigen Instrument der Raumplanung und der Wirtschaftspolitik. Im Rahmen der Raumplanung könnten implizite hedonische Preise als Grundlage für Kompensationseffekte in einem Kosten-Nutzen-Ansatz raumplanerischer Maßnahmen herangezogen werden. Jede raumplanerische Entscheidung ist mit Nutzen und Kosten verbunden, jede Entscheidung kennt Gewinner und Verlierer im Raum. Die hedonische Analyse von Bodenpreisen hilft, Kosten und Nutzen von Maßnahmen einzugrenzen und damit mögliche allokative Effekte und Verteilungswirkungen zu umreisen. Aus wirtschaftspolitischer Sicht ist vor allem die Preisentwicklung am Bodenmarkt von Interesse. Die starke Heterogenität von Grundstücken und Immobilien bewirkt, dass die Betrachtung von Durchschnittspreisen keine gesicherten Anhaltspunkte über Preisentwicklungen liefern. Entscheidend ist die Entwicklung des Preisniveaus bei Konstanthaltung der Qualitäten von Lage und Struktur der Grundstücke. Diese Informationen können durch die hedonische Preisanalyse gewonnen werden.

Dieses Papier hatte zum Ziel, die Wirkungen unterschiedlicher Spezifikationen und Schätzverfahren auf die hedonischen Preise zu ermitteln. Dabei zeigt sich, dass der Preisindex weniger sensitiv auf Spezifikation und Verfahren reagiert als die impliziten Preise einzelner Attribute. Lagefaktoren reagieren umso stärker auf die Wahl von Spezifikation und Schätzverfahren, je asymmetrischer die Verteilung der Faktoren im Raum ist. Als Beispiel sei hier der Koeffizient des Fluglärmkorridors 1 angeführt. Dieser Koeffizient misst den Abschlag, den der Fluglärm in Lärmzone 1 bei den Grundstückspreisen verursacht. Der Koeffizient schwankt je nach Schätzverfahren zwischen -0,1 und -0,2. Umgerechnet auf die belasteten Grundstücke in Wien lassen sich (aus diesem zugegeben groben Indikator) Vermögensschäden in Höhe von etwa 300 bis 600 Millionen Euro schätzen. Wollte man die Fluglärmgeschädigten kompensieren, dann wäre allerdings von entscheidender Bedeutung festzustellen, welcher Koeffizient nun die realen Verhältnisse am ehesten widerspiegelt. Das plausibelste Modell (STM) ermittelt einen Koeffizienten in Höhe von -0,17, also an der oberen Grenze der Schwankungsbreite.

Das Beispiel Fluglärm zeigt, wie wichtig die Wahl der richtigen Spezifikation und des Schätzverfahrens in der hedonischen Bodenpreisanalyse sein kann. In anderen Fällen, insbesondere in der Bewertung der höherrangigen Grünräume, der Nachbarschaft von Betriebs- und Industriegebieten und der Nähe höherrangiger Straßen zeigen sich noch größere Schwankungsbreiten. Diese Ergebnisse ermahnen zu einer rigorosen Modellierung der räumlichen Interdependenzen über adäquate Raum-zeitliche Matrizen und zu einer Auswahl von Schätzverfahren, welche das Fehlen wichtiger räumlicher Daten adäquat abbilden können.

Literatur

Assenmacher, W. (2004), Klassiches lineares Einzelgleichungsregressionsmodell, in: Gaab, W., Heilemann, U. und Wolters, J. (Hrsg.), Arbeiten mit ökonometrischen Modellen, Physica-Verlag, Heidelberg, 2004.

Assenmacher, W. und Braun, G.E. (1981), Das Einfachheitspostulat in Wissenschaftstheorie und Ökonometrie, Statistische Hefte, S. 152-175.

Baum, C.F. (2006), An Introduction to Modern Econometrics Using Stata, STATA Press, College Station, Texas.

Beck, N. (2001), Time-Series-Cross-Section Data: What have we learned in the past few years?, Annual Review of Political Science, Vol. 4, S. 271-293.

Beck, N. und Katz, J.N. (2001), Throwing Out the Baby with the Bath Water: A Comment on Green, Kim, and Yoon, International Organization, Vol. 55 (2), S. 487-495.

Bienert, S. und Funk, M. (2007, Hrsg.), Immobilienbewertung Österreich, Österreichischer Verband der Immobilientreuhänder (ÖVI), Wien.

Bourassa, S.C., Cantoni, E. und Hoesli, M. (2007), Spatial Dependence, Housing Submarkets, and House Price Prediction, Journal of Real Estate Finance and Economics, Vol. 35, S. 143-160.

Box, G.E.P. und Cox, D.R. (1964), An Analysis of Transformations, Journal of the Royal Statistical Society, Series B, Vol. 26, No. 2, S. 211-252.

Cassel, E. und Mendelson, R. (1985), The choice of functional forms for hedonic price equations: comment, Journal of Urban Economics, Vol. 18, S. 135-142.

Cheshire, P. und Sheppard, S. (1995), On the Price of Land and the Value of Amenities, Economica, Vol. 62, S. 247-267.

Cheshire, P. und Sheppard, S. (1998), Estimating the demand for housing, land, and neighbourhood characteristics, Oxford Bulletin of Economics and Statistics, Vol. 60, Nr. 3, S. 357-382.

Cropper, M.L., Deck, L.B. und McConnell, K.E. (1988), On the Choice of Functional Form for Hedonic Price Functions, Review of Economics and Statistics, Vol. 70(4), S. 668-675.

Follain, J.R. und Jimenez, E. (1985), Estimating the demand for housing characteristics – A Survey and Critique, Regional Science and Urban Economics 15, S. 77-107.

Goodman, A.C. (1978), Hedonic price, price indices and housing markets, Journal of Urban Economics, Vol. 5, S. 471-484.

Greene, W. (2000), Econometric Analysis, Upper Saddle River, NJ: Prentice Hall.

Halvorsen, R. und Palmquist, R. (1980), The Interpretation of Dummy Variables in Semilogarithmic Regressions, American Ecnomic Review, Vol. 70, S. 474-475.

Halvorsen, R. und Pollakowski, H.O., (1981), Choice of Functional Form for Hedonic Price Equations, Journal of Urban Economics, Vol. 10, S. 37-49.

Kennedy, P.E. (1981), Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations, American Economic Review, Vol. 71, S. 801.

Kuminoff. N.V., Parmeter, C.F. und Pope, J.C. (2008), Hedonic Price Functions: Guidance On Empirical Specification, Paper presented at the American Agrucultural Economics Association Annual Meeting: Orlando, Florida, July 27-29, 2008.

Lancaster, K. (1966), A new approach to consumer theory, Journal of Political Economy, Vol. 74, S. 132-157.

LeSage, J.P. (1998), Spatial Econometrics, Department of Economics, University of Toledo, www.spatial-econometrics.com/html/wbook.pdf

LeSage, J.P. (1999), The Theory and Practice of Spatial Econometrics, Department of Economics, University of Toledo. http://citeseerx.ist.psu.edu/viewdoc/ download?doi=10.1.1.111.4233&rep=rep1&type=pdf.

Malpezzi, S. (2002), Hedonic Pricing Models: A Selective and Applied Review, Center of Urban Land Economics Research, University of Wisconsin.

Maurer, R., Pitzer, M. und Sebastian, S. (2001), Konstruktion transaktionsbasierter Immobilienindizes: Theoretische Grundlagen und empirische Umsetzung für den Wohnungsmarkt in Paris, Manuskript.

Militino, A.F., Ugarte, M.D. und Carcia-REinaldos, L. (2004), Alternative Models for Describing Spatial Dependence among Dwelling Selling Prices, Journal of Real Estate Pricing and Economics, Vol. 29 (2), S. 193-209.

Munro, J. und Angulo, A. (2009), Model selection strategies in a spatial setting: Some additional results, Regional Science and Urban Economics, Vol. 39, S. 200-213.

Pace, R.K., Barry, R., Clapp, J.M. und Rodriquez, M. (1998), Spatiotemporal Autoregressive Models of Neighborhood Effects, Journal of Real Estate Finance and Economics, Vol. 17 (1), S. 15-33.

Pace, R.K., Barry, R., Gilley, O.W. und Sirmans, C.F. (2000), A method for spatial-temporal forecasting with an application to real estate prices, International Journal of Forecasting, Vol. 16, S. 229-246.

Palmquist, R.B. (1988), Welfare Measurement for environmental improvements using the hedonic model: The case of nonparametric marginal prices, Journal of Environmental Economics and Management, Vol. 15, S. 297-312.

Pindyck, R. S. und Rubinfeld, D.L. (1991), Econometric Models & Economic Forecasts, McGraw-Hill, 3rd Ed. New York u.a.

Rosen, S. (1974), Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, Journal of Political Economy, Vo. 82, S. 34-55.

Sheppard, St. (1999), Hedonic Analysis of Housing Markets, Handbook of Regional and Urban Economics, Vol. 3, Kapitel 41, S. 1595-1635.

Tu, Y., Yu, S-M., und Sun, H. (2004), Transaction-Based Office Price Indexes: A Spatiotemporal Modelling Approach, Real Estate Economics, Vo. 32, S. 297-328.

Wieser, R. (2006), Hedonic Prices on Vienna´s Urban residential land markets, IFIP-Working Paper 2/2006. Download: http://www.ifip.tuwien.ac.at/publikationen.htm.

Wieser, R. (2008), Marktbewertung struktureller Eigenschaften der Wiener Wohnbaulandgrundstücke, Mietwohnhäuser und Eigenheime, IFIP-Working Paper 2/2008. Download: http://www.ifip.tuwien.ac.at/publikationen.htm.

Wieser, R. (2009), Zahlungsbereitschaften der Wiener Wohnbevölkerung für (öffentliche) Freiräume, IFIP-Working Paper 1/2009. Download: http://www.ifip.tuwien.ac.at/publikationen.htm.

Wooldridge, J.M. (2002), Introductory Economics – A Modern Approach, 2nd Ed. Thomson, South-Western, Mason, Ohio.

ANHANG

Tabelle A 1 Deskriptive Statistiken und Bivariate Korrelationen wesentlicher Variablen im Modell

	Deskriptive Statistik		Bivariate Korrelationen					
	Mittelwert	Standardabweichung	Grundstücks-fläche	Flächen-form	Abbruch-objekt	Parzelliert	Schulen	ÖFFI-Erreichbarkeit
Preis (€)	561.291,7	1.224.186,0	0,68	-0,13	-0,10	-0,04	0,19	-0,26
Grundstücksfläche (m2)	1.494,2	2.933,6	**1,00**	-0,15	-0,01	-0,09	0,05	-0,03
Flächenform (Umfang/Fläche)	0,2	0,3	-0,15	**1,00**	0,00	0,03	-0,03	0,01
Hauptwidmung (in%)	95,6	14,7	-0,23	0,07	0,10	0,07	-0,24	0,26
Abbruchobjekt (Dummy)	0,9	0,3	-0,01	0,00	**1,00**	0,03	-0,16	0,22
Parzelliert (Dummy)	0,8	0,4	-0,09	0,03	0,03	**1,00**	0,01	0,03
Transdanubien (Dummy)	0,4	0,5	0,04	-0,02	0,12	0,01	-0,20	0,50
Baustruktur B (Dummy)	0,1	0,4	0,01	-0,03	-0,13	-0,05	0,33	-0,40
Einkauf (Dummy)	0,2	0,4	-0,04	0,01	0,13	0,04	-0,22	0,61
Schulen (Anzahl Umkreis 500m)	0,9	1,2	0,05	-0,03	-0,16	0,01	**1,00**	-0,47
Wienerwald (Distanz in m)	4.750,0	4.319,3	0,11	0,01	0,01	0,06	-0,02	0,43
Lobau (Distanz in m)	9.874,0	5.093,5	-0,12	0,01	0,00	-0,05	-0,03	-0,30
Prater (Dummy)	0,9	0,3	-0,05	-0,03	0,02	-0,04	-0,08	0,11
Donauinsel (Dummy)	0,5	0,5	0,04	-0,01	-0,08	-0,05	0,27	0,09
Schönbrunn (Dummy)	0,6	0,5	-0,07	0,02	-0,13	0,00	0,17	-0,54
Straßen Kat 2 (Distanz in m)	2.432,1	1.634,7	-0,03	0,00	0,15	0,00	-0,29	0,68
Straßen Kat 3 (Distanz in m)	640,4	561,3	-0,05	0,01	0,03	0,01	-0,24	0,30
Fluglärmzone 1 (Dummy)	0,1	0,3	0,01	0,00	-0,02	0,04	0,00	0,41
Fluglärmzone 2 (Dummy)	0,1	0,2	0,00	-0,01	-0,03	0,04	0,10	0,00
Luftverpester (Dummy)	0,0	0,2	0,02	-0,02	-0,02	-0,01	0,04	-0,12
∅ ÖFFI-Erreichbark. (in Minuten)	53,7	11,6	-0,03	0,01	0,22	0,03	-0,47	**1,00**
Distanz Stadtzentrum (in m)	7.659,4	2.770,3	-0,03	0,02	0,21	0,01	-0,49	0,81
FNeubeginner	0,1	0,8	-0,08	0,04	0,16	0,05	-0,49	0,62
F Bildung	-0,1	1,0	0,14	-0,01	0,00	0,03	0,05	0,14
F Familien	-0,1	0,8	-0,01	-0,01	-0,07	-0,04	0,17	-0,30
F Yuppies	0,1	0,9	-0,04	-0,01	-0,12	-0,01	0,27	-0,41
F Oldies	0,1	0,9	-0,04	-0,02	-0,13	-0,02	0,48	-0,41
Betriebsgebiete (FL-Ant. Umkreis 1 Tsd m)	8,2	15,1	0,09	-0,02	-0,05	-0,02	0,05	-0,20
Industriegebiete (FL-Ant. Umkreis 1 Tsd m)	6,8	16,9	0,04	0,07	0,05	0,01	-0,10	0,17
Parks und Weingärten (Distanz in m)	482,8	532,2	-0,01	0,00	0,12	0,01	-0,31	0,48
Sonstige Freiflächen (Distanz in m)	246,1	302,2	-0,08	-0,02	-0,17	0,02	0,40	-0,51

	Bivariate Korrelationen									
	Distanz Zentrum	Betriebs-gebiete	Industrie-gebiete	Parks od. Weingärten	Sonstige Freiräume	BKL I	BKL II	BKL III	BKL IV	BKL V
Preis (€)	-0,29	0,07	-0,05	-0,14	0,14	-0,36	0,05	0,10	0,29	0,35
Grundstücksfläche (m2)	-0,03	0,09	0,04	-0,01	-0,08	-0,17	0,11	0,02	0,08	0,15
Flächenform (Umfang/Fläche)	0,02	-0,02	0,07	0,00	-0,02	0,05	-0,01	-0,03	-0,03	-0,03
Hauptwidmung (in%)	0,29	-0,02	0,06	0,11	-0,18	0,36	-0,07	-0,17	-0,25	-0,18
Abbruchobjekt (Dummy)	0,21	-0,05	0,05	0,12	-0,17	0,21	-0,01	-0,13	-0,17	-0,02
Parzelliert (Dummy)	0,01	-0,02	0,01	0,01	0,02	0,02	-0,07	0,00	0,04	0,01
Transdanubien (Dummy)	0,37	0,09	0,25	0,35	-0,41	0,27	-0,04	-0,17	-0,18	-0,06
Baustruktur B (Dummy)	-0,40	0,17	-0,06	-0,19	0,37	-0,51	0,17	0,48	0,11	-0,01
Einkauf (Dummy)	0,52	-0,22	0,01	0,22	-0,18	0,27	-0,11	-0,15	-0,13	-0,04
Schulen (Anzahl Umkreis 500m)	-0,49	0,05	-0,10	-0,31	0,40	-0,55	0,07	0,40	0,37	0,07
Wienerwald (Distanz in m)	0,14	0,13	0,22	0,31	-0,23	0,02	-0,06	-0,02	0,06	0,00
Lobau (Distanz in m)	0,07	-0,14	-0,23	-0,21	0,15	0,06	0,03	-0,02	-0,10	-0,02
Prater (Dummy)	0,23	-0,01	-0,03	-0,15	0,02	0,16	0,02	-0,05	-0,20	-0,05
Donauinsel (Dummy)	-0,17	0,06	0,09	-0,07	0,16	-0,14	-0,06	0,10	0,17	0,01
Schönbrunn (Dummy)	-0,42	-0,11	-0,26	-0,37	0,42	-0,27	0,04	0,17	0,18	0,06
Straßen Kat 2 (Distanz in m)	0,63	-0,31	-0,02	0,25	-0,33	0,43	-0,07	-0,27	-0,29	-0,09
Straßen Kat 3 (Distanz in m)	0,29	-0,17	-0,01	0,40	-0,23	0,25	-0,07	-0,16	-0,13	-0,04
Fluglärmzone 1 (Dummy)	0,24	-0,07	0,15	0,23	-0,10	0,06	-0,07	-0,02	0,01	-0,02
Fluglärmzone 2 (Dummy)	-0,06	-0,02	0,00	0,02	0,02	-0,11	-0,04	0,08	0,14	-0,02
Luftverpester (Dummy)	-0,17	0,00	-0,07	-0,10	0,27	-0,13	0,04	0,09	0,07	0,02
∅ ÖFFI-Erreichbark. (in Minuten)	0,81	-0,20	0,17	0,48	-0,51	0,61	-0,14	-0,39	-0,38	-0,11
Distanz Stadtzentrum (in m)	**1,00**	-0,14	0,13	0,44	-0,58	0,64	-0,10	-0,40	-0,45	-0,13
FNeubeginner	0,57	-0,17	0,14	0,38	-0,48	0,59	-0,16	-0,41	-0,30	-0,09
F Bildung	0,06	0,31	0,29	0,24	-0,17	-0,13	-0,04	0,17	0,06	0,01
F Familien	-0,36	-0,01	-0,07	-0,10	0,19	-0,24	-0,01	0,15	0,20	0,04
F Yuppies	-0,41	-0,07	-0,13	-0,28	0,38	-0,37	0,13	0,26	0,19	0,03
F Oldies	-0,39	0,17	-0,03	-0,17	0,34	-0,41	0,06	0,29	0,30	-0,01
Betriebsgebiete (FL-Ant. Umkreis 1 Tsd m)	-0,14	**1,00**	0,19	-0,10	-0,06	-0,12	0,04	0,08	0,03	0,02
Industriegebiete (FL-Ant. Umkreis 1 Tsd m)	0,13	0,19	**1,00**	0,21	-0,19	0,10	-0,02	-0,04	-0,09	-0,02
Parks und Weingärten (Distanz in m)	0,44	-0,10	0,21	**1,00**	-0,30	0,31	-0,11	-0,19	-0,15	-0,05
Sonstige Freiflächen (Distanz in m)	-0,58	-0,06	-0,19	-0,30	**1,00**	-0,51	0,07	0,39	0,32	0,05