
Wohnimmobilienpreisindizes für Österreich GFR Februar 2015, Innsbruck

Wolfgang A. Brunauer
Statistical Model Development
wolfgang.brunauer@realevalue.at

Ronald S. Weberndorfer
Statistical Analysis & Geoinformatics
ronald.weberndorfer@realevalue.at

- Herausforderungen für die Indexkonstruktion
- Immobilienpreisindizes in Österreich
- Ausblick

Herausforderungen für die Indexkonstruktion

- **Vergleichbarkeit:** Berücksichtigung der verschiedenen Qualitätsstufen (z.B.: Standort, Größe, Alter, technische Ausstattung)
- **Zeitliche Effekte:** Auch wenn sich die preisbestimmenden Merkmale nur allmählich ändern, darf auf lange Sicht nicht davon ausgegangen werden, dass diese starr sind.
- **Änderung von Standards:** Änderungen von Qualitätsstandards im Zeitverlauf
- **Repräsentativität:** Was ist die Grundgesamtheit, auf die man Rückschlüsse ziehen möchte?

■ Repeated Sales Methode:

- Geringer Anteil von mehrfach verkauft Objekten, insbesondere für Einfamilienhäuser (EFH)
- Matching-Problem im Zeitverlauf, insbesondere für Wohnungen
- Mögliche Änderungen der Gebäudestruktur im Zeitverlauf

■ Stratifikationsmethode

- Asymmetrisches / heterogenes Transaktionsverteilung in Österreich
- Kleine Stichprobengrößen je nach Segmentdefinition

■ Hedonische Regression

- Überwindet die Probleme der Repeated Sales- bzw. der Stratifikationsmethode
- Für qualitätsbereinigte Vergleiche muss der Basisdatensatz alle relevanten Attribute enthalten (omitted variable bias).

Herausforderungen – Theorie der hedonischen Preise

- Immobilien bestehen aus einer Vielzahl von Eigenschaften, die jedes Objekt einzigartig machen
- Implizite Auswahl an vielen verschiedenen Waren und Dienstleistungen:
 - Haushalt kauft Objekt
 - Dadurch implizit Auswahl von Werten für alle K Eigenschaften $z = (z_1, z_2, \dots, z_K)$.
- Der explizite Preis für das Haus /die Wohnung ist eine Funktion des ganzen Bündels von Eigenschaften, also $P = P(z)$ (hedonische Preisfunktion), mit funktionalen Zusammenhängen welche möglicherweise über Zeit und Raum variieren.
- Die impliziten Preise der einzelnen Eigenschaften können beispielsweise mit einem multiplen Regressionsmodell ermittelt werden.

Herausforderungen – Datenquellen in Österreich

1. Kaufpreissammlung aus dem Grundbuch:

- + Repräsentativ (im Bezug auf den Stichprobenumfang)
- Nur Adresse, Kaufpreis und teilw. Flächen vorhanden kaum zusätzliche Eigenschaften
- Wahrscheinliche Verzerrung aufgrund besonderer Umstände (z.B.: Kauf von Familienmitgliedern) oder Steuerflucht
- sehr teurer Datensatz

2. Datensammlungen von Finanzinstituten:

- + Oft sehr detailliert
- Kleine Anzahl an Beobachtungen
- Wahrscheinlich nicht für den gesamten Markt repräsentativ

3. Angebotspreise von Immobilien-Makler-Plattformen:

- + Große Anzahl von Beobachtungen
- + Die Meisten relevanten Variablen vorhanden
- Möglicher "Angebotsaufschlag" – variiert dieser über die Zeit?
- Mögliche Verzerrungen durch Erfassung - insbesondere bei schlechter Qualität

→ Entscheidung für die Angebotspreisdaten (3), weil diese flächendeckend in ausreichenden Anzahl ("Länge") und mit geeigneten Objekt-Eigenschaften ("Breite") vorhanden sind.

Immobilienpreisindizes in Österreich

Indizes in Österreich – Überblick

- Preisdaten von der österreichischen Wirtschaftskammer (WKÖ Immobilienpreisspiegel – nicht in dieser Präsentation)
- Zwei Wohnimmobilienpreisindizes (RPPI -Residential Property Price Indices for Austria) veröffentlicht auf der Webseite der OeNB (Quelle: TU Wien):
 - Zeit Dummy Index: Transaktions-Perspektive (s. Anhang)
 - Räumlicher Imputations Index: Schwerpunkt auf Immobilienbestand
- Ab 3q2014: RPPI des österreichischen Statistikamtes (Statistik Österreich)
- Für Gewerbeimmobilien gibt es derzeit keine verlässlichen Daten

Indizes in Österreich – Räumlicher Imputations Index - Motivation

- Vergleichbarkeit:
 - Die Lageeigenschaften werden durch Lagevariablen auf Zählsprenzel- und auf Gemeindeebene berücksichtigt, sowie durch eine hierarchische Struktur von Land und Bezirk ("unbeobachtet" räumliche Heterogenität)
 - Modellierung von nichtlinearen Effekten in Objekt- und Lagevariablen mit semiparametrischer Modellspezifikationen

- Zeiteffekte:
 - Basismodell (7 Jahre für EFH, 9 Jahre für ETW)
 - Wandernde Modellperioden von zwei Jahre für das "Vergleichsmodell"

- Änderung des Standards:
 - Repräsentative Merkmale werden für jedes Jahresmodell ermittelt
 - Kombination in einen Fisher-Index

- Repräsentativität:
 - Imputation in jeden Zählsprenzel mit unterschiedlichen Eigenschaften
 - Gewichtete Aggregation mit den Anteilen des jeweiligen Bundeslandes an der Gesamtmenge

Indizes in Österreich – Räumlicher Imputations Index – Vorgehensweise (I)

1. Definition von Teilstichproben für den Modellbau
 - Basismodell: 2007Q1 - 2014Q1 (ETW) und 2005Q1 - 2014Q1 (EFH), mit Basisperiode 2014Q1
 - Vergleichsmodell: Bewegliches Fenster von zwei Jahren, mit Vergleich des letzten verfügbaren Quartals
2. Bestimmung der "durchschnittlichen Eigenschaften" für das letzte Jahr, bis zur letzten Basis bzw. bis zum Vergleichsmodell.
3. Regressionsmodell für Basis- und Vergleichszeitraum
4. Räumliche Imputation:
 - Simulation eines Objekts mit "durchschnittlichen Eigenschaften" für jeden Zählsprengel im Basis- und Vergleichszeitraum.
 - Ausnutzung der Modellierung durch
 - a) Lagevariablen und
 - b) unerklärter räumlicher Heterogenität zur Simulation von Marktwerten in unbeobachteten räumlichen Einheiten.

Indizes in Österreich – Räumlicher Imputations Index – Vorgehensweise (II)

5. Indexberechnung auf Zählsprengenebene:

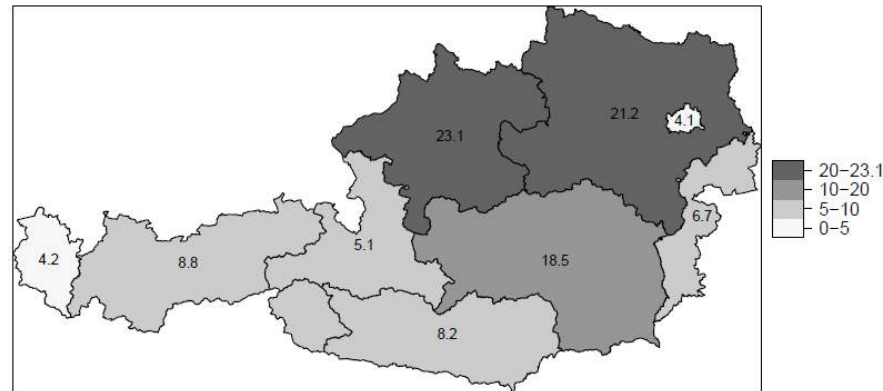
- Laspeyres-Index: Durchschnittliche Merkmale der Basisperiode
- Paasche-Index: Durchschnittliche Merkmale der laufenden Periode
- Fisher-Index: Geometrisches Mittel der Laspeyres und Paasche
- Neu skalieren, um 2007Q1 = 100 mit Dummy-Effekte aus Basismodell

6. Gewichtungsschema:

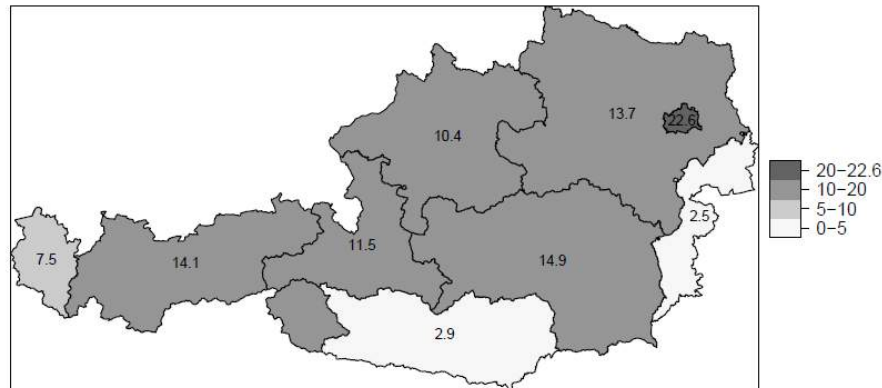
- Für jedes Teilmodell und Land: Mittelwertbildung über Zählbezirke pro Bundesland und Typ (Ausnahme: Für Bundesland Wien / Typ EFH werden die "innere" Bezirke nicht einbezogen)
- Gewichtung der Bundesland-Indizes mittels Anteil des Bundeslands und Typ am Gesamtbestand (basierend auf der *Household Finance and Consumption Survey*, HFCS 2010)

Indizes in Österreich – Räumlicher Imputations Index – Gewichtungsschema (basieren auf HFCS)

Gewichte EFH



Gewichte ETW



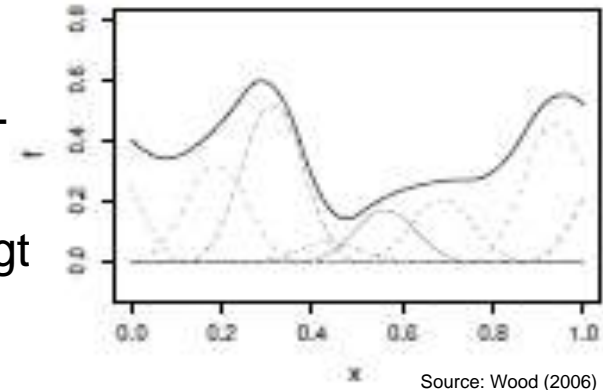
State	Verteilung pro Objekttyp und Bundesland		Verteilung pro Bundesland und Objekttyp		Verteilung am Gesamtbestand	
	EFH	ETW	EFH	ETW	EFH	ETW
B	90.6%	9.4%	6.7%	2.5%	5.2%	0.5%
K	90.8%	9.2%	8.2%	2.9%	6.4%	0.6%
N	84.4%	15.6%	21.2%	13.7%	16.5%	3.1%
O	88.6%	11.4%	23.1%	10.4%	17.9%	2.3%
Sa	61.0%	39.0%	5.1%	11.5%	4.0%	2.6%
St	81.3%	18.7%	18.5%	14.9%	14.4%	3.3%
T	68.7%	31.3%	8.8%	14.1%	6.9%	3.1%
V	66.1%	33.9%	4.2%	7.5%	3.2%	1.7%
W	39.1%	60.9%	4.1%	22.6%	3.2%	5.0%
AUT	77.8%	22.2%	100.0%	100.0%	77.8%	22.2%

Indizes in Österreich – Räumlicher Imputations Index – Methode

- Generalisierte Additive Modelle (GAM; Wood, 2006): Basisstrukturelle- und Verteilungsannahmen wie bei Generalized Linear Models

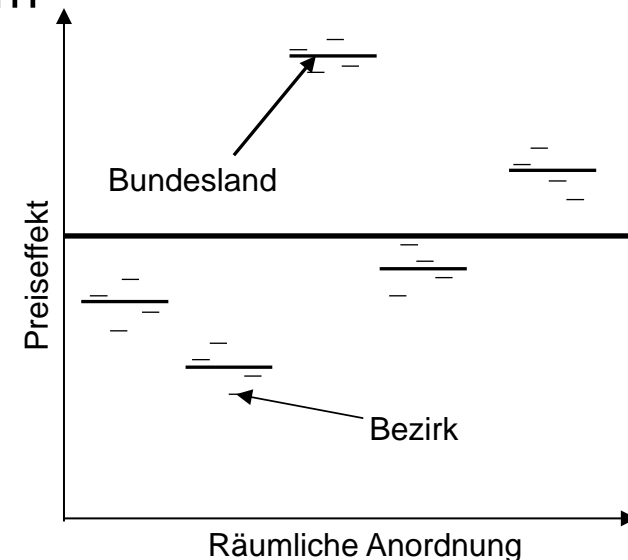
- $E(y_i | \mathbf{z}_i, \mathbf{x}_i) = h(\eta_i)$ mit additiver Prädiktor $\eta_i = f_1(\mathbf{z}_{i1}) + \dots + f_q(\mathbf{z}_{iq}) + \mathbf{x}_i' \boldsymbol{\gamma}$

- $\mathbf{x}_i' \boldsymbol{\gamma}$ ist die parametrische Teil der Prädiktor
- \mathbf{z}_j ist eine kontinuierliche Kovariable, Zeitskala, orts- oder Cluster-Index
- f_j sind ein- oder höherdimensionalen, nicht unbedingt stetige Funktionen



- Automatisierte Auswahl von Glättungsparametern unter Verwendung des Generalisiertem Kreuzvalidierung (GCV) Kriterium

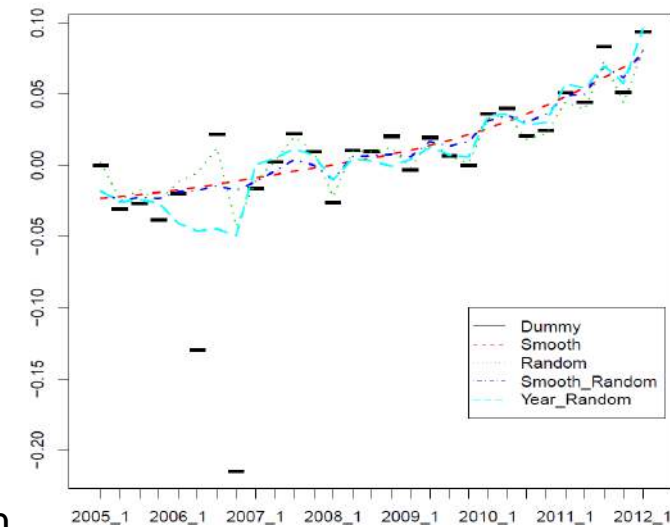
- Kontinuierliche Effekte werden mit penalisierten Regressionssplines (sog. Thin Plate Regression Splines) modelliert
- Räumliche (Bezirke) und zeitliche (Quartal) Indizes werden über von Random Effects modelliert



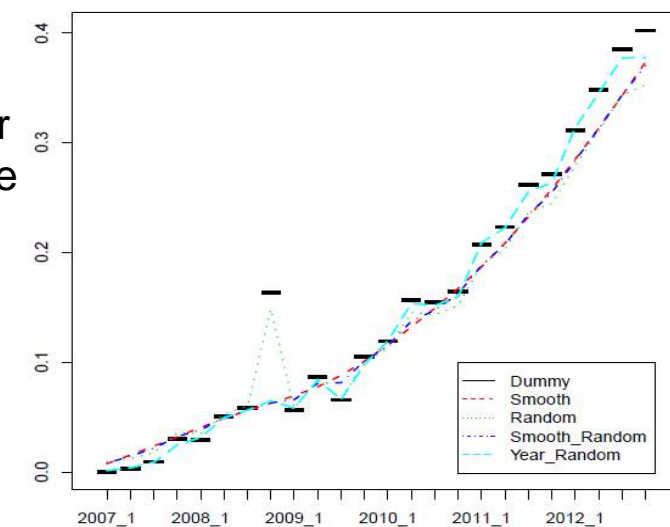
Indizes in Österreich – Räumlicher Imputations Index – Trendmodellierung

- Für die Modellierung von Trendeffekten wurden folgende Alternativen geprüft:
 - **Dummy-Ansatz:** Quartalsweiser Zeit-Dummy-Effekt → volatile Schätzergebnisse
 - **Smooth-Trend-Ansatz** mit Verwendung von Regressionssplines → unterschätzt abrupte Sprünge
 - **Random-Zeiteffekt** → Effekt wird unterschätzt, wenn der Markt nach oben oder im Zeitverlauf nach unten geht
 - **Smooth-Random-Effekt-Modell:** Kombination aus einem nicht-linearen Zeittrend und Zufallseffekten → Ergebnisse scheinen durch die nichtlineare Entwicklung dominiert.
 - **Year-Random-Zeiteffekt:** Zeiteffekte in einer hierarchischen Struktur, d.h. Basiseffekt wird als jährlicher Effekt geschätzt, Abweichungen von diesem werden in die Zufallseffekte übernommen → robuster Kompromiss zwischen Datentreue und Bestrafung
- Entscheidung für hierarchische Zeiteffekte (Jahres-Zufallseffekt-Modell).

Various trend models – SFH



Various trend models – Flat



Indizes in Österreich – Räumlicher Imputations Index – Modellergebnis (Basismodell)

EFH

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	12.307679	0.035462	347.066	< 2e-16	***
year2006	-0.021249	0.020817	-1.021	0.307396	
year2007	0.031625	0.012357	2.559	0.010505	*
year2008	0.025220	0.012185	2.070	0.038503	*
year2009	0.031708	0.011852	2.675	0.007477	**
year2010	0.053941	0.011473	4.702	2.61e-06	***
year2011	0.074080	0.016927	4.376	1.22e-05	***

...
[Additional attributes, condition, state effects]

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:

	edf	Ref.df	F	p-value	
s(y_q)	2.978	5.286	0.236	0.953	
s(lnarea)	5.757	6.922	436.682	< 2e-16	***
s(lnarea_plot)	4.438	5.516	183.393	< 2e-16	***
s(age)	6.698	7.582	106.075	< 2e-16	***
s(ln_census_educ)	5.525	6.732	40.821	< 2e-16	***
s(ln_census_plot_price)	6.308	7.287	37.446	< 2e-16	***
s(mun_no_nights)	4.020	4.797	11.871	4.22e-11	***
s(mun_pot)	5.209	6.391	9.742	2.76e-11	***
s(mun_pp_growth_04_09)	6.338	7.268	8.753	4.24e-11	***
s(ln_distr_wko_sfh)	7.880	8.117	4.076	6.65e-05	***
s(district)	65.086	81.406	2.725	9.70e-15	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.769 Deviance explained = 77.3%
REML score = 2242.8 Scale est. = 0.084245 n = 10757

ETW

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	11.373231	0.112406	101.180	< 2e-16	***
year2008	0.042663	0.012949	3.295	0.000988	***
year2009	0.070320	0.012371	5.684	1.34e-08	***
year2010	0.131047	0.012215	10.728	< 2e-16	***
year2011	0.179753	0.019087	9.417	< 2e-16	***

...

[Additional attributes, condition, state effects]

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

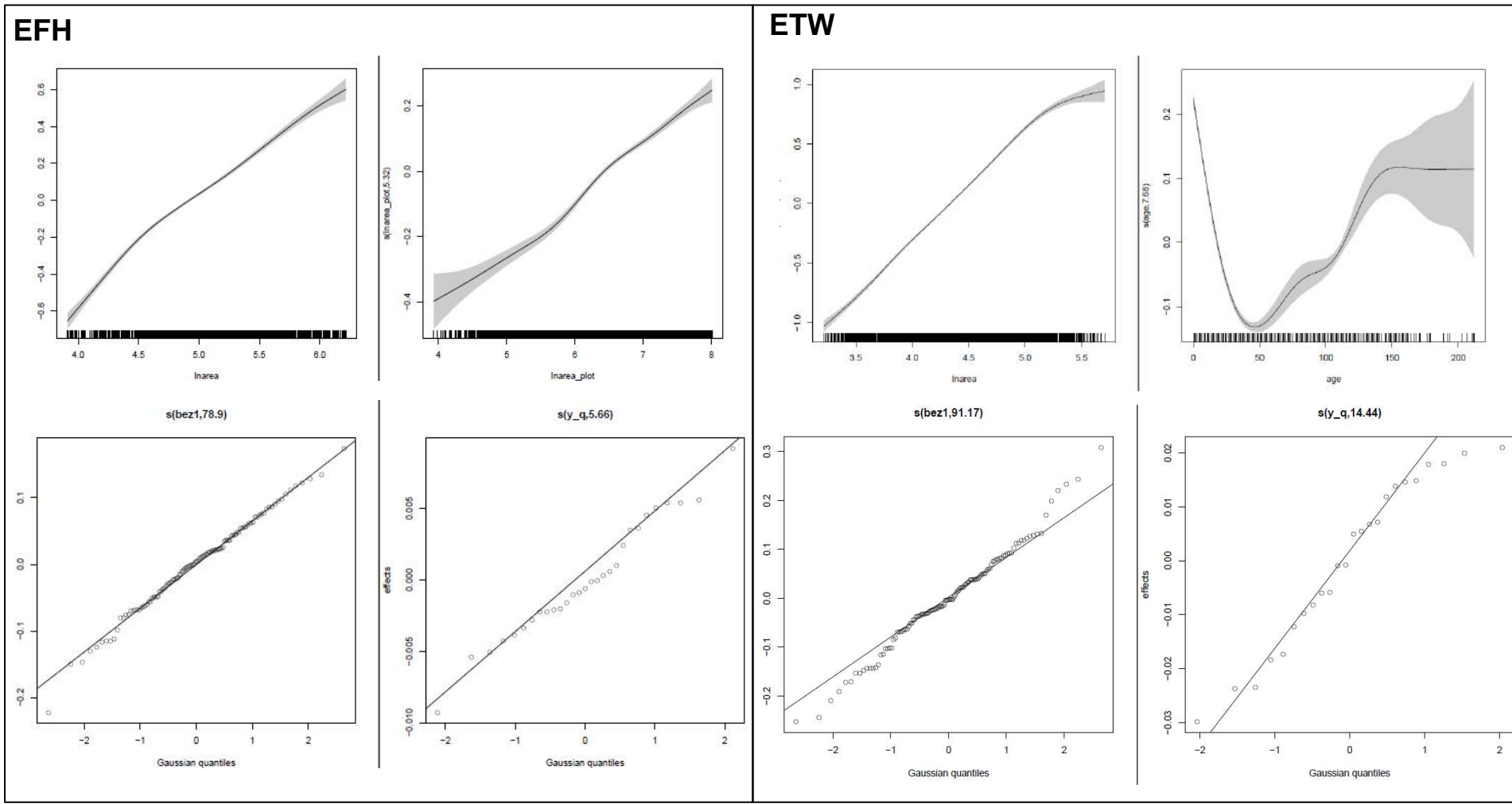
Approximate significance of smooth terms:

	edf	Ref.df	F	p-value	
s(y_q)	8.667	10.641	2.431	0.00564	**
s(lnarea)	7.481	8.447	907.079	< 2e-16	***
s(age)	7.537	8.282	276.515	< 2e-16	***
s(ln_census_educ)	5.971	7.201	33.040	< 2e-16	***
s(ln_census_plot_price)	1.703	2.119	101.291	< 2e-16	***
s(mun_no_nights)	7.587	8.191	11.357	< 2e-16	***
s(mun_pot)	4.556	5.601	5.670	1.25e-05	***
s(ln_mun_dens)	2.136	2.545	8.243	6.03e-05	***
s(district)	90.015	98.000	19.312	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

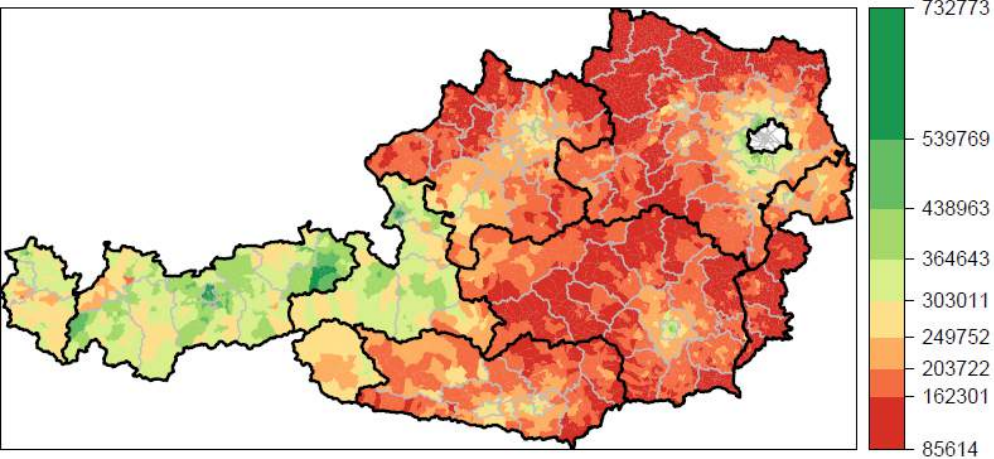
R-sq.(adj) = 0.883 Deviance explained = 88.5%
REML score = -124.74 Scale est. = 0.054618 n = 15311

Indizes in Österreich – Räumlicher Imputations Index – Modellergebnis (Basismodell)

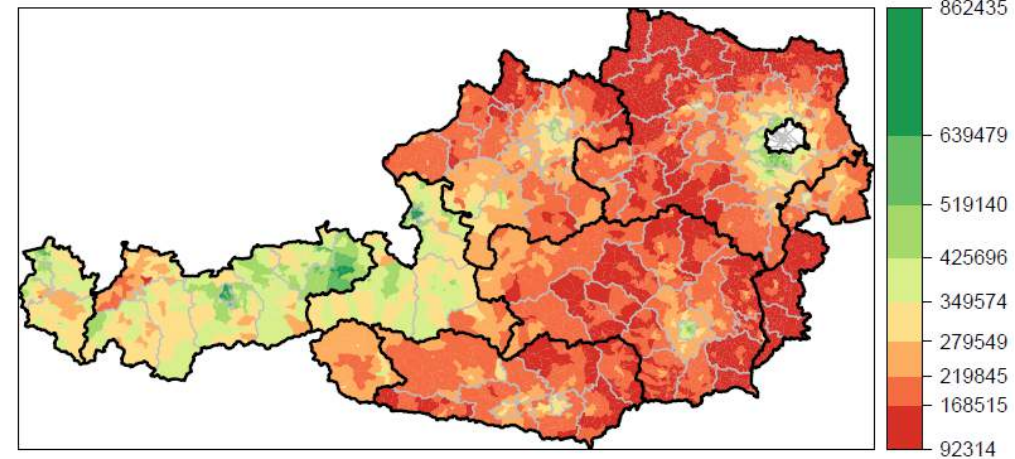


Indizes in Österreich – Räumlicher Imputations Index – EFH (ohne Wien)

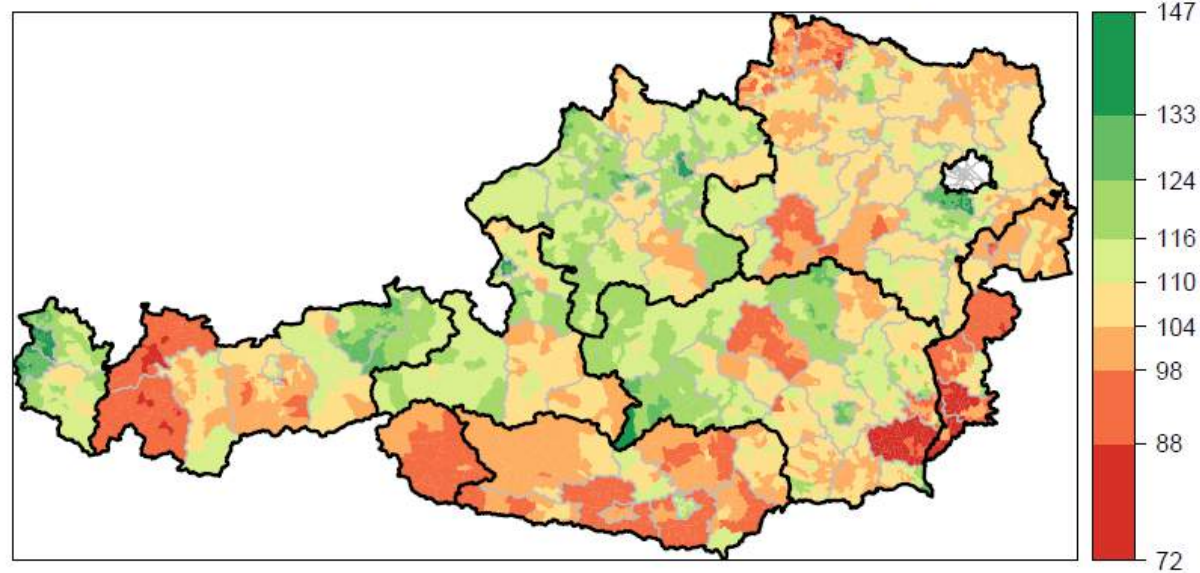
MV SFH, 1q2010 (w/o W)



MV SFH, 2q2014 (w/o W)

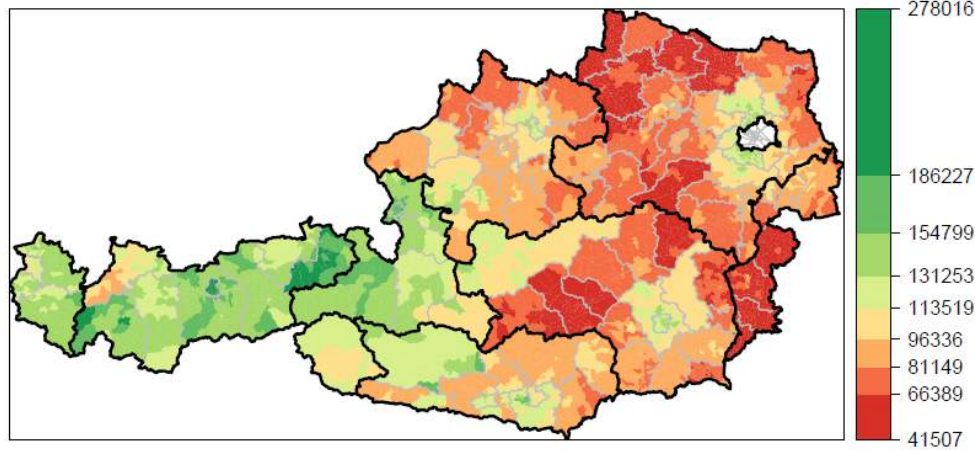


Index 2q2014 SFH, base 1q2010 (w/o W)

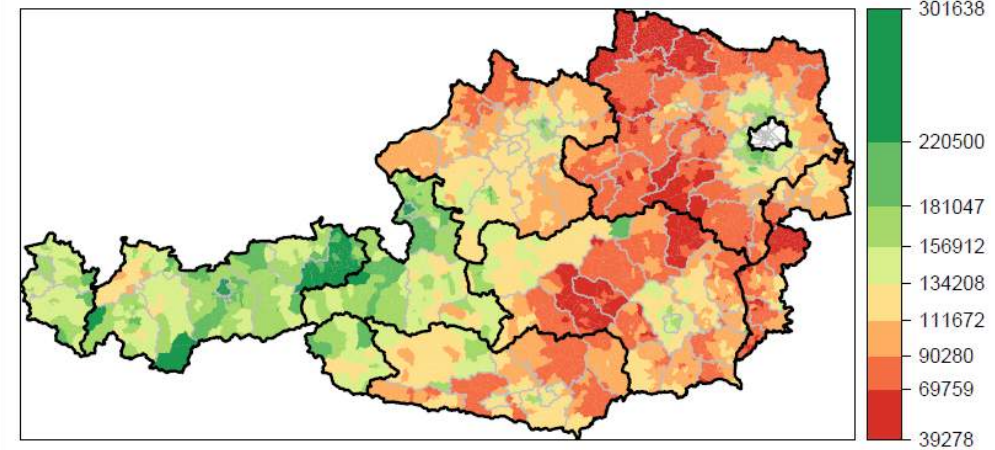


Indizes in Österreich – Räumlicher Imputations Index – ETW / Österreich (ohne Wien)

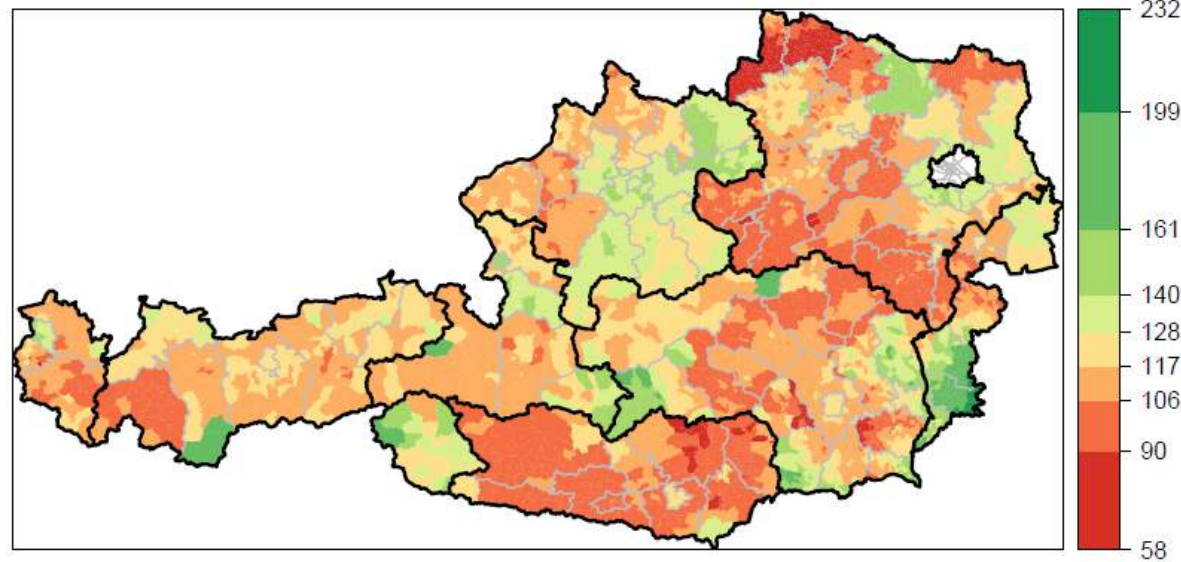
MV apartments, 1q2010 (w/o W)



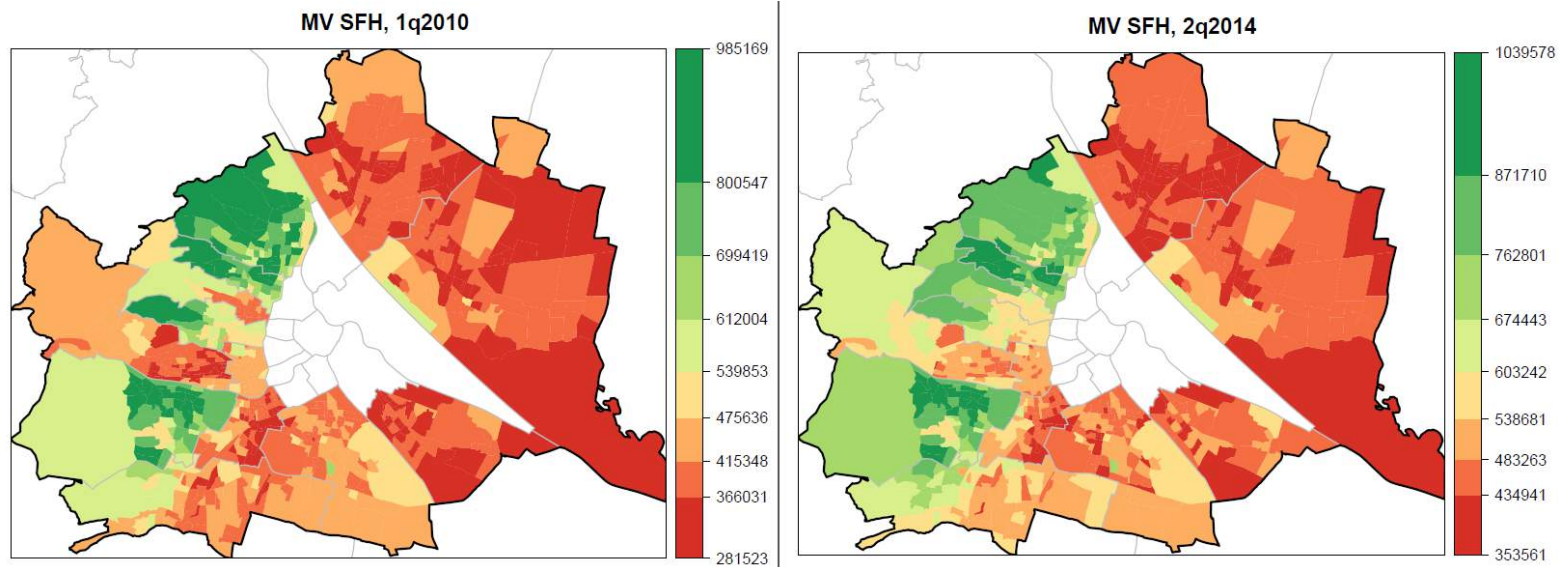
MV apartments, 2q2014 (w/o W)



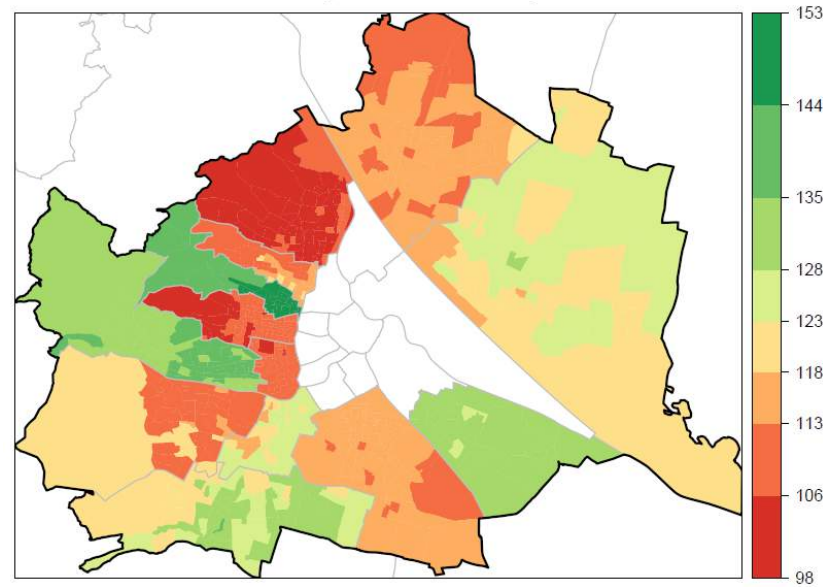
Index 2q2014 apartments, base 1q2010 (w/o W)



Indizes in Österreich – Räumlicher Imputations Index – EFH/ Wien

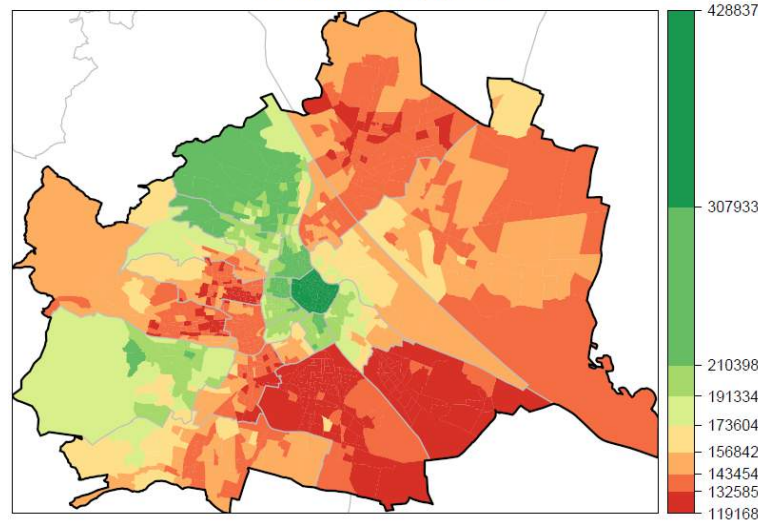


Index 2q2014 SFH, base 1q2010

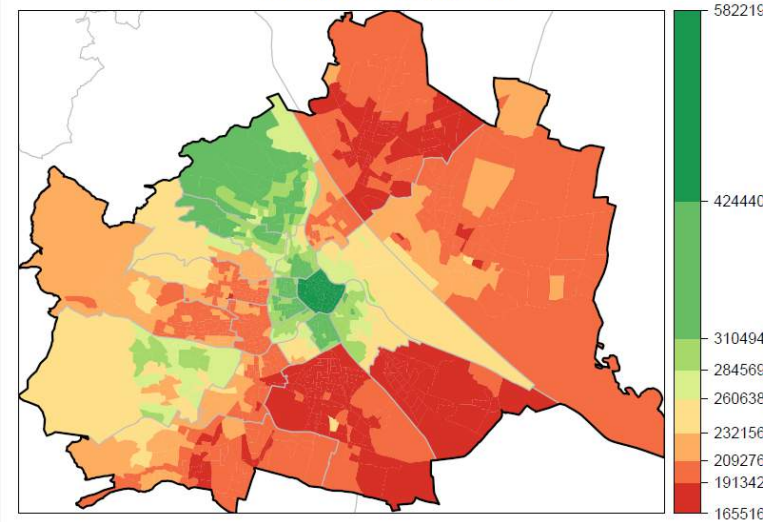


Indizes in Österreich – Räumlicher Imputations Index – ETW / Wien

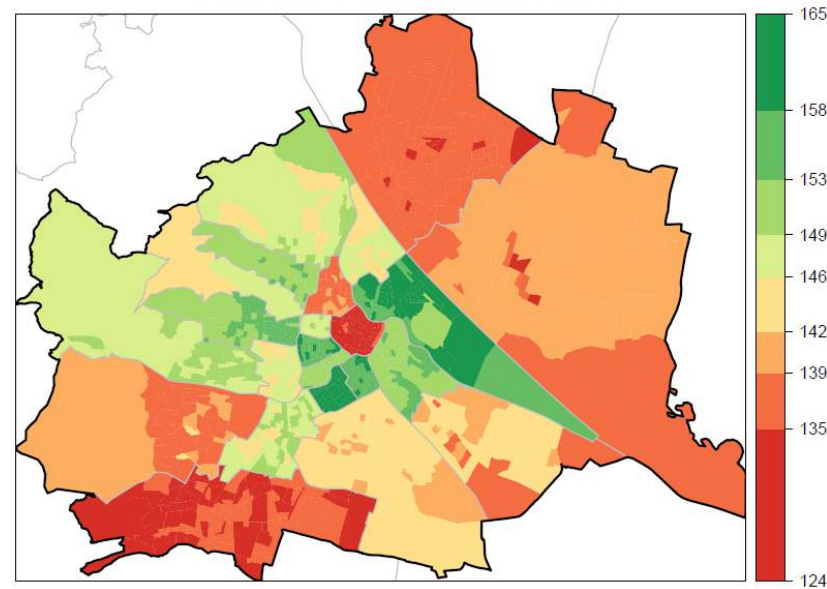
MV apartments, 1q2010



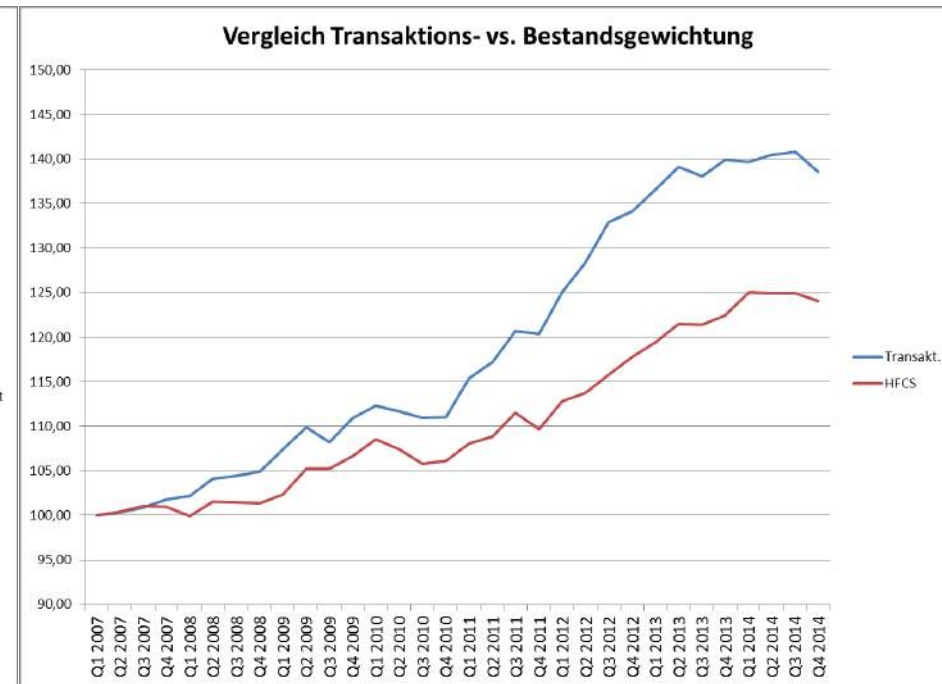
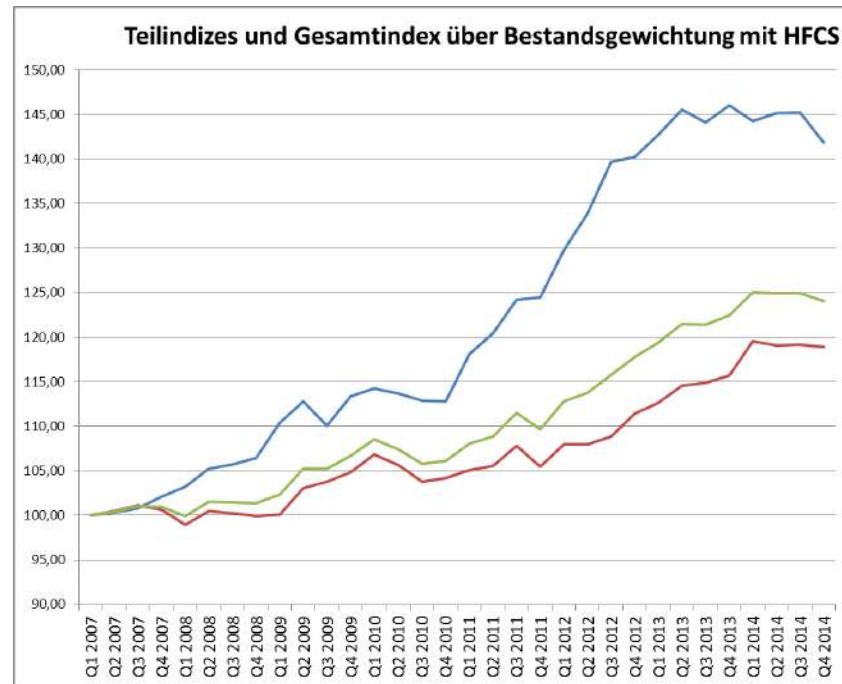
MV apartments, 2q2014



Index 2q2014 apartments, base 1q2010



Indizes in Österreich – Räumlicher Imputations Index – Overall results



Links:
Über Gesamtbestand gewichtet (HFCS)

Right:
Vergleich: Gewichtung über Gesamtbestand (HFCS) vs. Anzahl der Transaktionen

- Mögliche Weiterentwicklungen
 - Gewichtung: Detailliertere Gewichtungsschemata. Beispielsweise bezirksweise Verteilung des Gesamtbestand
 - Durchschnittliche Qualität: Bestimmung der durchschnittlicher Qualität auf einer detaillierteren räumlichen Skala (z.B.: Bundesland oder Bezirk)
 - Modell:
 - Räumlich variierenden Effekte für Objekteigenschaften
 - Einschließlich Standort Variablen im jeweiligen Jahr gemessen (derzeit nur "Querschnitt" Perspektive)
 - Einschließlich zusätzlicher Variablen für die Messung grundlegenden Nachfragefaktoren (z.B.: demografische Entwicklung) und Versorgung (z.B.: verfügbare Bodenfläche)
 - Daten: Neue Datenquellen erschließen, die nicht von den Schwächen (Angebotsaufschlag) betroffen sind.
- Weitere Anwendungen
 - Disaggregierte Indizes, mit Expertengetriebene und / oder modellbasierte Marktabgrenzung
 - Mit Multi-Level-Modellen Abweichungen vom Grundimmobilienpreise schätzen

Danke für Ihre Aufmerksamkeit!

Wolfgang A. Brunauer

Statistical Model Development
wolfgang.brunauer@realevalue.at

Ronald S. Weberndorfer

Statistical Analysis & Geoinformatics
ronald.weberndorfer@realevalue.at

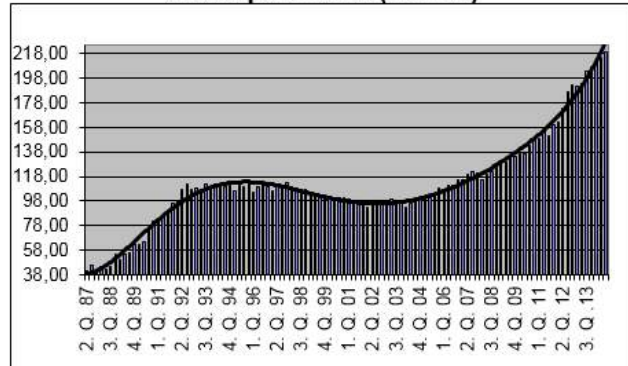
Anhang

Indizes in Österreich – Zeit Dummy Index - Modell

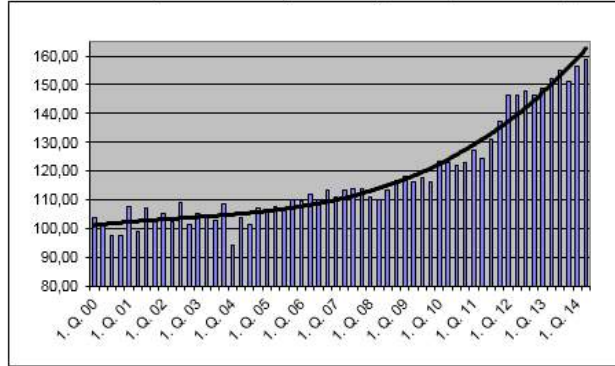
- Multiples lineares Regressionsmodell, in dem der (*log*) Preis durch Objekteigenschaften und Bezirks Dummies erklärt wird.
- Die Jahresmodelle enthalten Quartals-Dummies
- Die Modelle sind überlappend aufgebaut, wobei die Zeiteffekte mit dem letzten Quartal des vorherigen Modells verknüpft sind.
- Die Schätzung erfolgt mit sechs verschiedenen Modellen, welche nach Objekttyp und Region aufgeteilt sind:
 - "Vienna" und "Österreich, außer Wien"
 - EFH, ETW (neu oder gebraucht)
- Der Gesamteffekt wird durch ein Gewichtungsschema festgelegt (siehe unten)
- Probleme:
 - "Retrospektive" Variation der Zeiteffekte durch bestehende Modellstruktur
 - Qualitätsänderungen werden nicht explizit berücksichtigt
 - Keine Berücksichtigung von nicht beobachteten Regionen

Indizes in Österreich – Zeit Dummy Index - Ergebnisse

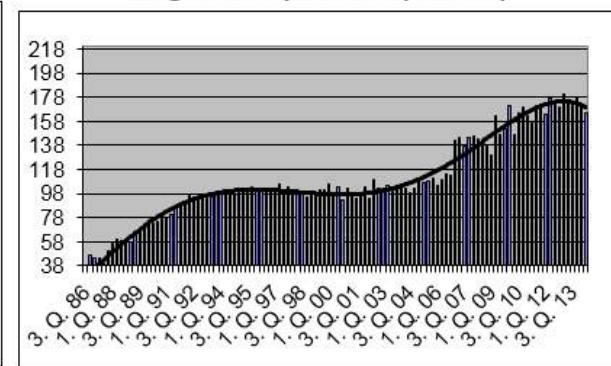
Used apartments (Vienna)



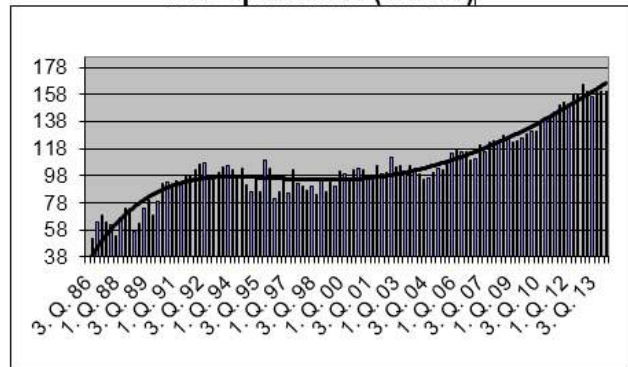
Used apartments (Austria, except Vienna)



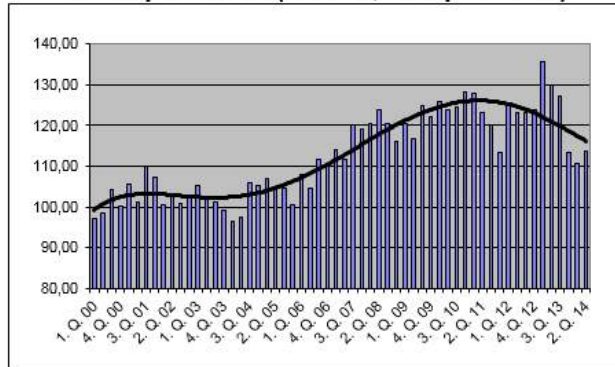
Single-family houses (Vienna)



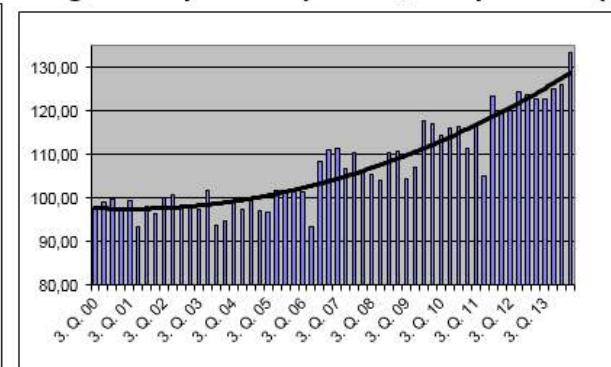
New apartments (Vienna)



New apartments (Austria, except Vienna)



Single-family houses (Austria, except Vienna)



<http://www.srf.tuwien.ac.at/feil/immobilienbewertung/Index214.pdf>

Indizes in Österreich – Zeit Dummy Index - Gewichtungsschema

Die Gewichte werden durch die Anzahl der Transaktionen aus dem Grundbuch (2008-2013) festgelegt.

1. Wohnungen: Gewichtung der Indizes für neue und gebrauchte ETW; für Wien und Österreich, außer Wien

	Wohnung (neu)	Wohnung (gebraucht)
Vienna	15%	85%
Austria, außer Wien	13%	87%

2. Regionale Indizes: Gewichtung ETW und EFH

	Wohnung	Haus
Wien	93%	7%
Österreich, außer Wien	70%	30%

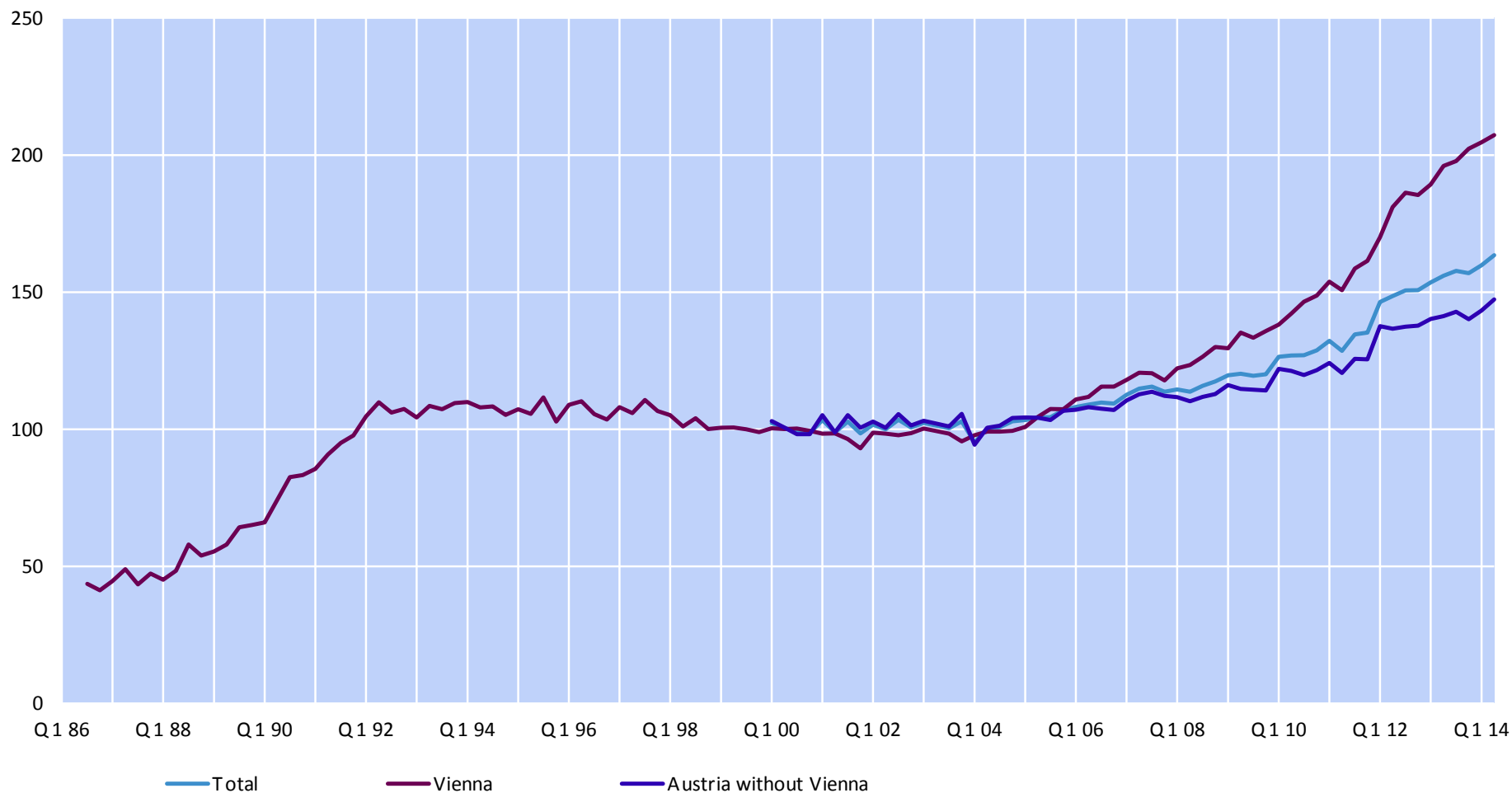
3. Index für Österreich: Gewichtung Wien und Österreich, mit Ausnahme Wien

Wien	Österreich, außer Wien
27%	73%

Indizes in Österreich – Zeit Dummy Index - Ergebnisse

House price developments

2000=100



Source: OeNB, Prof. Wolfgang Feilmayr, Department of Spatial Planning, Vienna University of Technology.