

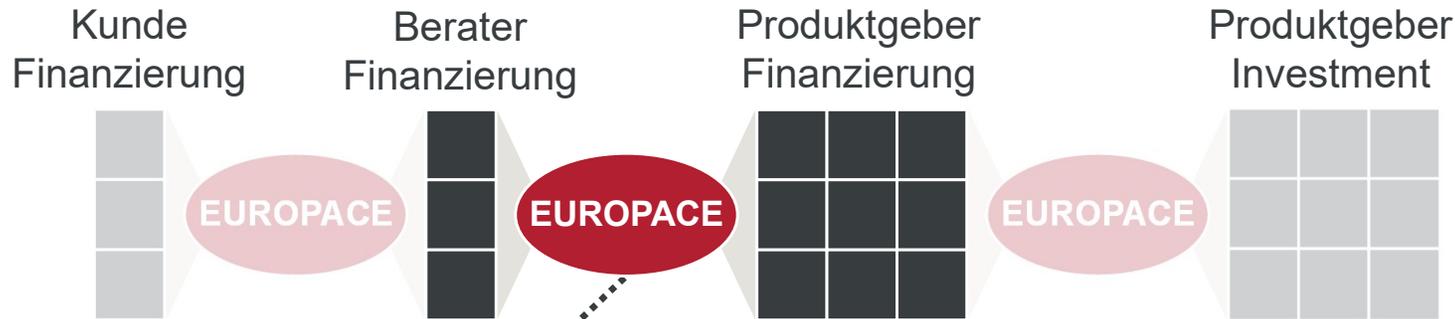
# Ein Immobilienpreisindex auf Basis von Kreditdaten von Hypoport Ergebnisse einer Methodenstudie

Workshop Immobilienpreise Bonn 29.01.2009

Hans-Joachim Dübel, Finpolconsult; Sören Iden Hypoport B.V.

## Datenquelle: Hypoport – ein B2B-Intermediär in der Hypothekenfinanzierung

Europace als Datenlieferant für den HPX



### HPX-Genesis

- 2003** Hypoport startet B2B-Plattform zur Kreditvergabe
- 2006** HPX-mean wird entwickelt, Basisperiode August 2005
- 2007/8** Machbarkeitsstudie zu hedonischem Immobilienpreisindex Finpolconsult im Auftrag von BBR
- 2008/9** Entwicklung der HPX-hedonic Indizes

### Information zur EUROPACE-Plattform:

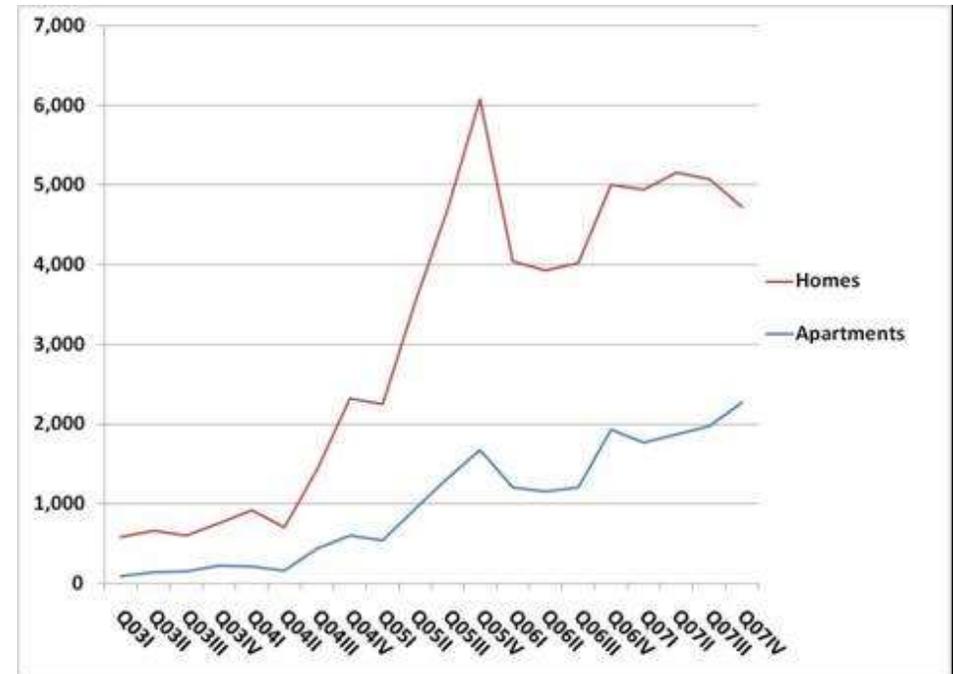
- Vermittlung von 5.000 – 8.000 Immobilienfinanzierungen pro Monat
- Transaktionsvolumen ~3,4 Mrd. € in Q2/2008
- Liefert realisierte Transaktionsdaten, keine Angebotspreise

## Datensatz: Beobachtungen Europace-Datensatz

### Grobstruktur des Datensatzes I. Quartal 2003 – IV. Quartal 2007

<b>Beobachtungen gesamt</b>	197023
nach Finanzierungszweck	
An-/ Umbau	3484
Kauf	84844
Neubau	42266
Umschuldung	60460
zur Besicherung	5969
nach Objektart	
davon Eigentumswohnung	45764
davon Ein/Zweifamilienhaus	144093
davon Mehrfamilienhaus	7166
nach Nutzungsart	
Eigennutzung	156187
Teilvermietung	20384
Vermietung	20452
Eigentumswohnungen/Nutzung	45764
davon Eigennutzung	32498
davon Teilvermietung	565
davon Vermietung	12701
davon Eigennutzung und Kauf	22762
davon Eigennutzung und Neubau	2184
davon Eigennutzung und Kauf nicht modernisiert	17783
davon Eigennutzung und Neubau nicht modernisiert	2183
Ein/Zweifamilienhaus/Nutzung	144093
davon Eigennutzung	123400
davon Teilvermietung	16169
davon Vermietung	4524
davon Eigennutzung und Kauf	44163
davon Eigennutzung und Neubau	36038
davon Eigennutzung und Kauf nicht modernisiert	25514
davon Eigennutzung und Neubau nicht modernisiert	36033

### Zeitprofil der Kreditfall-Beobachtungen je Quartal aus der Europace-Plattform I. Quartal 2003 – IV. Quartal 2007



## Variablen: Hypoport und Matching mit Variablen des Bundesamtes für Bau- und Raumwesen

### Hypoport / Europace

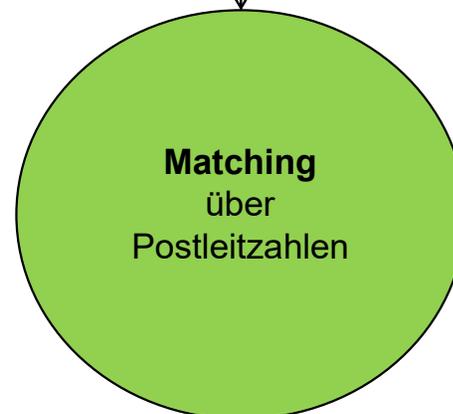
V.a. Hauseigenschaften:

- Objektart
- Nutzungsart
- Baualter
- Wohnfläche
- Grundstücksfläche
- Anzahl Wohneinheiten
- Finanzierungszweck
- Kaufpreis
- Modernisierungskosten
- Herstellkosten
- Grundstückspreis

### BBR

V.a. Lageeigenschaften:

- Bevölkerungszahl der Gemeinde
- Fahrtzeiten zum nächsten Oberzentrum
- Entfernungsklassen (Zentrum, Pendelbereiche)
- Regionseinteilung (Ost/West/Nord/Süd)
- Variable wirtschaftliche Dynamik
- Variable Beschäftigungspotential (Arbeitsmarkproxy)



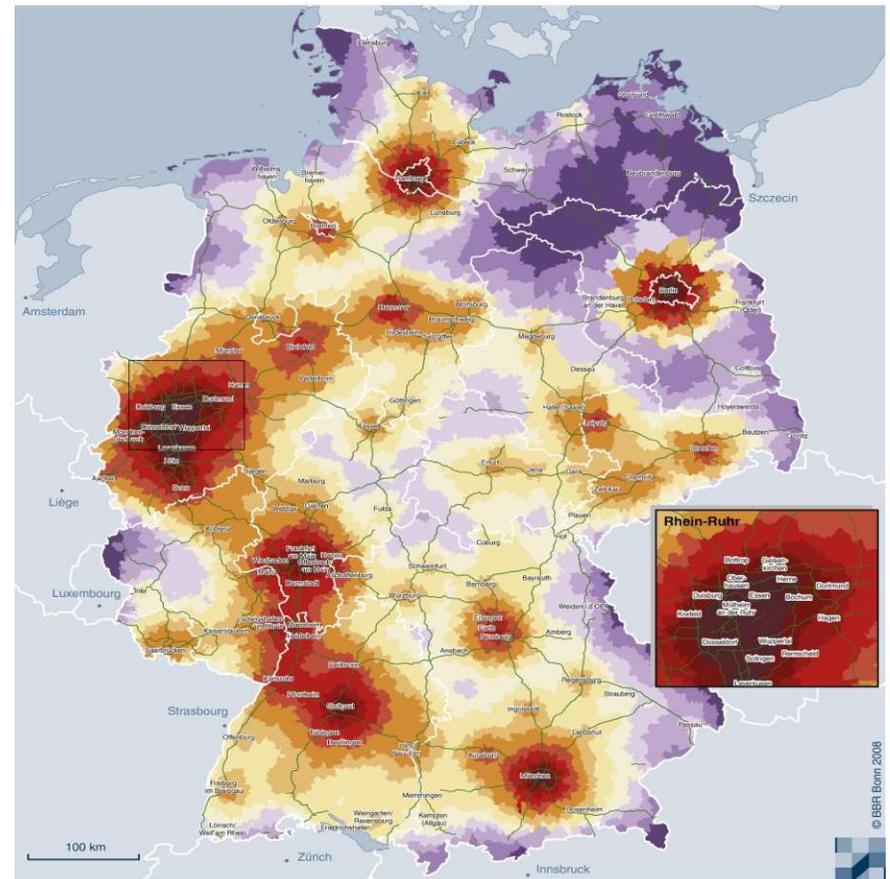
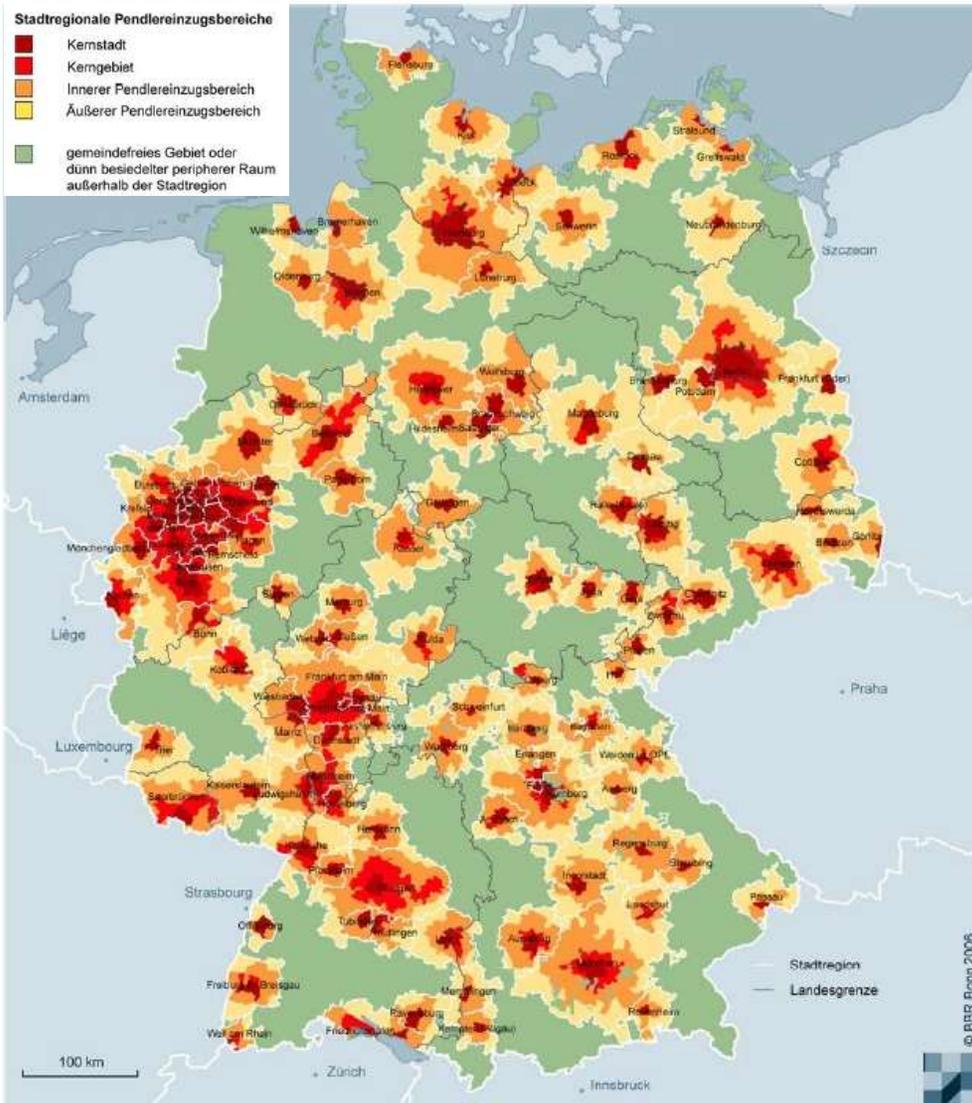
### Was fehlt ?

Genaue Objekteigenschaften wie

- Unterkellerung, Zahl Räume, Etage, Fahrstuhl, Bäder
- Bauzustand, Wohnlage (kaum Variation)

**Variablen: BBR-Vorgaben**

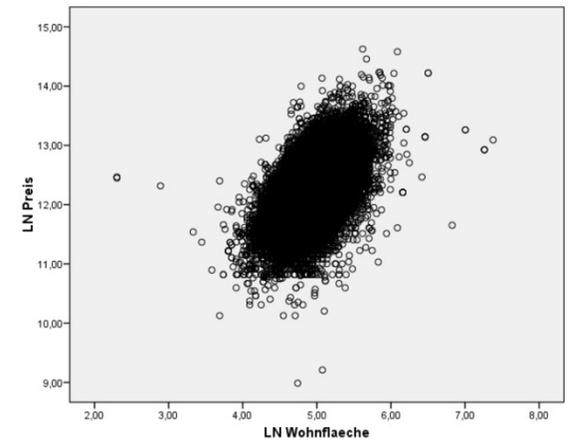
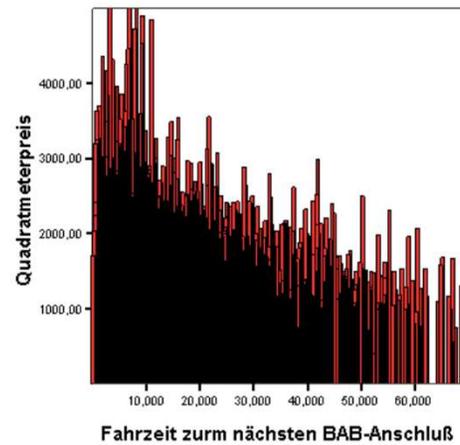
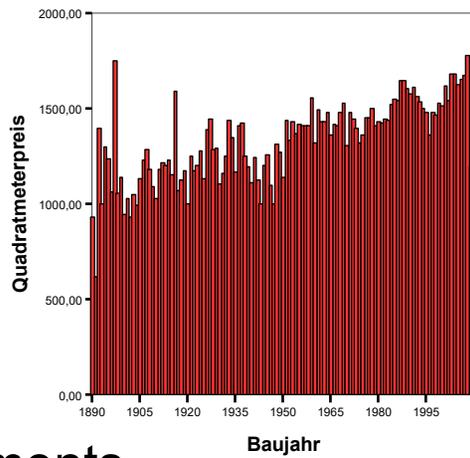
Pendlereinzugsbereiche, Beschäftigtenpotential



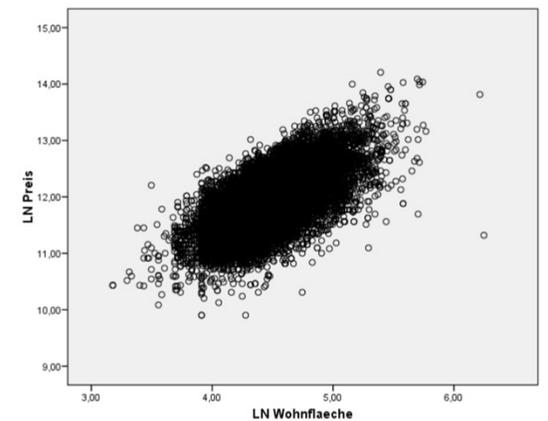
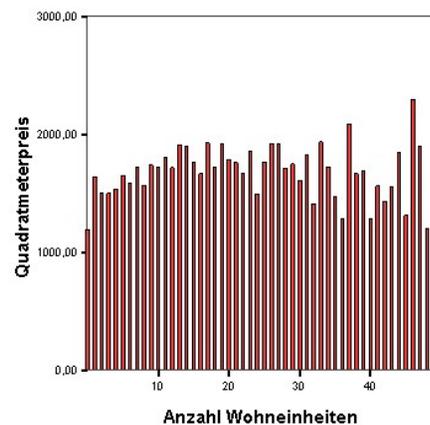
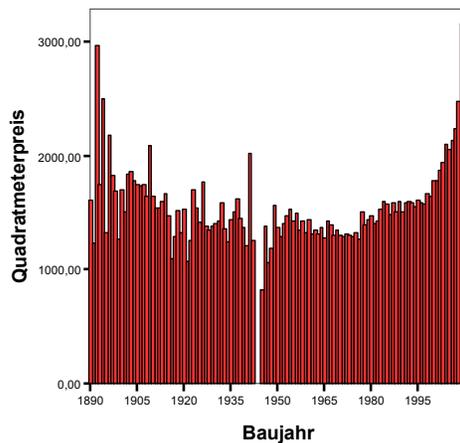
**Spezifikationsfragen: Variablen**

Stetig vs. Klassen, Potenzen / Logs

**Häuser**



**Apartments**



## Spezifikationsfragen: Modell und Variationen

### Hedonisches Regressionsmodell mit Zeitvariablenmethode

$$\ln P_t^h = \sum_{c=1}^C \beta_c X_{c,t}^h + \alpha_t T_t + v_t^h$$

- $P_t^h$  Preis einer Wohnung bzw. eines Hauses ( $h$ ) zum Zeitpunkt ( $t$ );
- $\beta_c X_{c,t}^h$   $\beta_c$  sind die impliziten Preise,  $c=1,2,\dots,C$  die Charakteristika (BETA-MATRIX)
- $\alpha_t T_t$  Dummyvariable für Zeitschätzung
- $v_t^h$  Fehlerterm

Zu testende Fragen:

- Ausreichende Fallzahlen? Variation der Abgrenzung der Beobachtungen?
- Signifikante Einflussfaktoren / Kollinearität?
- Hinreichende Signifikanz der Zeitdummyschätzer? Stabilität im Zeitablauf?
- Konstante Strukturfaktoren (Beta-Matrix)? Dynamisierung vs. hedonische Imputation und 'Nachziehen' der Basisperiode?
- Regionalregressionen vs. Bundesregression mit Bezug auf a) – d) vorzuziehen?

→ 3 Ansätze: Bundesregression, Gesamtperiode (1) und benachbart (2),  
Regionalregressionen, Gesamtperiode (3)

## Ansatz 1: Bundesweite Regression, statische/konstante Betamatrix

Homes: Teilvermietungen führen zu Verschlechterungen, funktionales Baualter zu Verbesserungen

		Selbstnutzung		Selbstnutzung und Teilvermietung					Selbstnutzung	
		Standard	Dummy	Standard	Dummy	.. mit Dummy	.. mit	.. mit	Baualter	Wie (8)
		(1)	Modernisierung (2)	(3)	Teilvermietung (4)	Modernisierung (5)	Bevoelkerung (6)	Entfernung>30 (7)	funktional (8)	ohne TV/Mod (9)
<b>Beobachtungen</b>		61478	80111	89618	89618	89618	89618	89618	89580	61451
<b>Teststatistiken</b>										
	R <sup>2</sup>	0.509	0.510	0.489	0.500	0.501	0.511	0.518	0.522	0.526
	Ajustiertes R <sup>2</sup>	0.509	0.510	0.489	0.500	0.500	0.511	0.518	0.521	0.526
<b>Variablen</b>	<b>Typ</b>									
KONSTANTE	stetig	8.357	8.448	8.903	8.537	8.538	8.159	8.212	8.229	8.111
FLAECHEN										
Wohnflaeche	stetig, LN	0.762	0.741	0.644	0.722	0.721	0.714	0.708	0.701	0.743
Grundstuecksflaeche	stetig, LN	0.047	0.049	0.052	0.051	0.051	0.061	0.067	0.069	0.064
BAUJAHR										
vor 1919	1/0	0.762	-0.421	-0.449	-0.429	-0.421	-0.416	-0.414	X	X
1919 - 1949	1/0	0.047	-0.310	-0.353	-0.335	-0.327	-0.334	-0.332	X	X
1950 - 1979	1/0	0.762	-0.230	-0.267	-0.251	-0.244	-0.249	-0.249	X	X
1980 - 1999	1/0	0.047	-0.115	-0.130	-0.124	-0.120	-0.121	-0.119	X	X
2000 und juenger	1/0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	X	X
Baualter in Jahren	1/0	X	X	X	X	X	X	X	-0.007140	-0.007295
Baualter i.J.quadriert	1/0	X	X	X	X	X	X	X	0.0000286	0.0000297
INVESTITION										
Teilvermietung	1/0	X	X	X	-0.170	-0.170	-0.167	-0.163	-0.160	X
Modernisierung	1/0	X	-0.011	X	X	-0.014	-0.014	-0.014	-0.010	X
VERDICHTUNG										
Bevoelkerung	stetig, LN	X	X	X	X	X	0.037	0.032	0.033	0.026
Entf. Oberzentrum >30min	1/0	X	X	X	X	X	X	-0.092	-0.092	-0.079
Metropole	1/0	0.222	0.257	0.270	0.266	0.266	0.113	0.106	0.111	0.093
Grossstadt	1/0	0.033	0.050	0.057	0.055	0.055	-0.035	-0.048	-0.046	-0.051
Laendlich	1/0	-0.167	-0.186	-0.188	-0.186	-0.185	-0.147	-0.134	-0.134	-0.127
Umland	1/0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
REGION										
Ostdeutschland	1/0	-0.216	-0.235	-0.238	-0.237	-0.238	-0.232	-0.223	-0.224	-0.206
Norddeutschland	1/0	-0.089	-0.095	-0.096	-0.097	-0.096	-0.086	-0.086	-0.089	-0.083
Sueddeutschland	1/0	0.165	0.162	0.162	0.163	0.162	0.175	0.173	0.172	0.173
Westdeutschland	1/0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ANTRAGSDATUM										

## Ansatz 1: Bundesweite Regression, statische/konstante Betamatrix

Apartments: bessere Fits ex-ante, Modernisierungsvariable verschlechtert Fit

		Selbstnutzung		Selbstnutzung einschl. Modernisierungen			Selbstnutzung	
		Standard	Modernisierungen	und Anzahl	und	Baualter	Wie (5)	
		(1)	(2)	Wohneinheiten	Bevoelkerung	funktional	ohne Modernis.	
		(1)	(2)	(3)	(4)	(5)	(6)	
<b>Beobachtungen</b>		19966	24946	24946	24946	24900	19923	
<b>Teststatistiken</b>								
	R <sup>2</sup>	0.632	0.627	0.629	0.635	0.640	0.647	
	Ajustiertes R <sup>2</sup>	0.631	0.627	0.628	0.634	0.639	0.646	
<b>Variablen</b>	<b>Typ</b>							
KONSTANTE	stetig	7.847	7.804	7.737	7.296	7.307	7.369	
FLAECHEN	Wohnflaeche stetig, LN	0.921	0.932	0.942	0.941	0.937	0.925	
BAUJAHR								
	vor 1919	1/0	-0.220	-0.219	-0.219	-0.225	X	X
	1919 - 1949	1/0	-0.362	-0.347	-0.344	-0.345	X	X
	1950 - 1979	1/0	-0.429	-0.415	-0.413	-0.413	X	X
	1980 - 1999	1/0	-0.261	-0.255	-0.255	-0.251	X	X
	2000 und juenger	1/0	0.000	0.000	0.000	0.000	X	X
	Baualter in Jahren	1/0	X	X	X	X	-0.020740	-0.021672
	Baualter i.J.quadriert	1/0	X	X	X	X	0.0002720	0.0002840
	Baualter i.J. hoch drei	1/0	X	X	X	X	-0.00000091	-0.00000093
INVESTITION								
	Modernisierung	1/0	X	-0.046	-0.045	-0.045	-0.038	X
VERDICHTUNG								
	Bevoelkerung stetig, LN	X	X	X	0.045	0.047	0.046	
	1-4 Wohneinheiten	1/0	X	0.000	0.000	0.000	0.000	
	5-20 Wohneinheiten	1/0	X	0.038	0.028	0.025	0.032	
	ueber 20 Wohneinheiten	1/0	X	0.001	-0.016	-0.022	-0.010	
	Metropole	1/0	0.258	0.266	0.261	0.106	0.103	
	Grossstadt	1/0	0.008	0.008	0.005	-0.083	-0.077	
	Laendlich	1/0	-0.124	-0.132	-0.130	-0.089	-0.083	
	Umland	1/0	0.000	0.000	0.000	0.000	0.000	
REGION								
	Ostdeutschland	1/0	-0.142	-0.156	-0.155	-0.180	-0.191	0.103
	Norddeutschland	1/0	-0.023	-0.034	-0.032	-0.035	-0.032	-0.077
	Sueddeutschland	1/0	0.176	0.170	0.169	0.177	0.181	-0.083
	Westdeutschland	1/0	0.000	0.000	0.000	0.000	0.000	
ANTRAGSDATUM								

## Ansatz 1: .. Wie oben, mit veränderten Regionalvariablen, Regionalabgrenzungen

Homes: neue Stadttypen und Beschäftigenpotentialvariable mit verbesserten Fits

		Bundesebene			Regionen			Bundesebene		
		(9) aus Tabelle 2	BBR Stadttypen	Wirtschaftl. Dynamik	Nord	Ost	Sued	West	Beschaeftigten- Potential	
		(1)	(2)	(3)	(5)	(6)	(7)	(8)	(9)	
<b>Beobachtungen</b>		61,451	61,451	45,873	14,733	15,750	11,015	19,953	60998	
<b>Teststatistiken</b>										
	R <sup>2</sup>	0.526	0.552	0.537	0.516	0.516	0.499	0.498	0.581	
	Ajustiertes R <sup>2</sup>	0.526	0.551	0.537	0.515	0.515	0.498	0.497	0.580	
<b>Variablen</b>		<b>Typ</b>								
KONSTANTE		stetig	8.111	8.275	8.063	8.334	7.922	8.498	8.145	8.050
FLAECHEN										
	Wohnflaeche	stetig, LN	0.743	0.735	0.759	0.700	0.753	0.718	0.740	0.726
	Grundstuecksflaeche	stetig, LN	0.064	0.077	0.091	0.086	0.063	0.074	0.099	0.091
BAUJAHR										
	Baualter in Jahren	1/0	-0.007295	-0.007521	-0.006418	-0.007803	-0.008489	-0.008711	-0.007355	-0.007773
	Baualter i.J.quadriert	1/0	0.00002971	0.00003209	0.00002762	0.00003887	0.00004946	0.00003126	0.00003127	0.00003379
VERDICHTUNG										
	Bevoelkerung	stetig, LN	0.026	0.020	0.024	0.015	0.028	0.035	0.018	0.009
	Metropole	1/0	0.093	X	X	X	X	X	X	X
	Grossstadt	1/0	-0.051	X	X	X	X	X	X	X
	Umland	1/0	0.000	X	X	X	X	X	X	X
	Laendlich	1/0	-0.127	X	X	X	X	X	X	X
	Kernstadt Agglomeration	1/0	X	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Kernstadt gross	1/0	X	-0.160	-0.133	-0.174	-0.150	-0.272	-0.132	-0.043
	Kernstadt klein	1/0	X	-0.243	-0.216	-0.250	-0.241	-0.343	-0.156	-0.051
	Pendel in	1/0	X	-0.173	-0.159	-0.173	-0.134	-0.256	-0.165	-0.053
	Pendel aus	1/0	X	-0.288	-0.289	-0.315	-0.216	-0.398	-0.253	-0.125
	Laendlich	1/0	X	-0.345	1)	-0.429	-0.205	-0.469	-0.329	-0.170
REGION										
	Ostdeutschland	1/0	-0.206	-0.250	-0.214	X	X	X	X	-0.184
	Norddeutschland	1/0	-0.083	-0.111	-0.134	X	X	X	X	-0.034
	Sueddeutschland	1/0	0.173	0.174	0.155	X	X	X	X	0.219
	Westdeutschland	1/0	0.000	0.000	0.000	X	X	X	X	0.000
	Entf. Oberzentrum >30min	1/0	-0.079	-0.015	-0.031	0.015	-0.027	0.036	-0.075	-0.010
	Beschaeftigtenpotential	stetig	X	X	X	X	X	X	X	0.0000075743
	Wachsend	1/0	X	X	0.000	X	X	X	X	X
	Stagnierend	1/0	X	X	-0.053	X	X	X	X	X
	Schrumpfend	1/0	X	X	-0.141	X	X	X	X	X
	Divergierend BevWa	1/0	X	X	-0.082	X	X	X	X	X
	Divergierend BeschWa	1/0	X	X	-0.084	X	X	X	X	X
ANTRAGSDATUM										

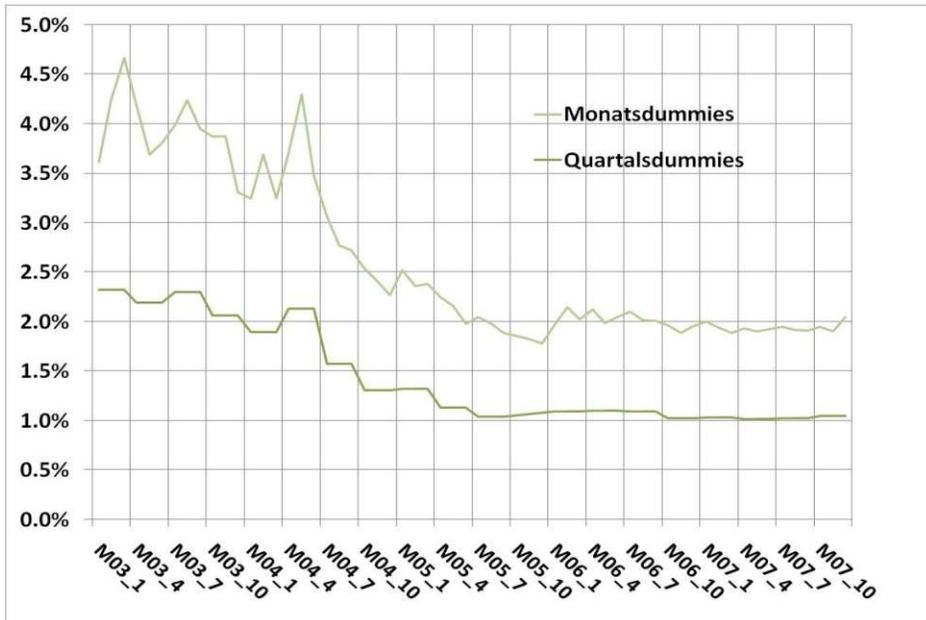
## Ansatz 1: .. Wie oben, mit veränderten Regionalvariablen, Regionalabgrenzungen

Apartments: wie Homes, zusätzlich wirtschaftliche Dynamikvariable signifikante Verbesserung

		Bundesebene			Regionen			Bundesebene		
		(6) aus	BBR Stadttypen	Wirtschaftl.	Nord	Ost	Sued	West	Beschaeftigten-	
		Tabelle 3	(2)	Dynamik	(5)	(6)	(7)	(8)	Potential	
		(1)	(2)	(3)	(5)	(6)	(7)	(8)	(9)	
<b>Beobachtungen</b>		19,923	19,923	18,071	2,900	2,698	6,440	7,210	19,850	
<b>Teststatistiken</b>										
	R <sup>2</sup>	0.647	0.665	0.691	0.726	0.652	0.688	0.683	0.699	
	Ajustiertes R <sup>2</sup>	0.646	0.664	0.690	0.723	0.647	0.686	0.682	0.699	
<b>Variablen</b>	<b>Typ</b>									
KONSTANTE	stetig	7.369	7.530	7.342	6.746	7.146	7.863	7.458	7.527	
FLAECHE	Wohnflaeche	0.925	0.926	0.958	1.031	0.945	0.896	0.955	0.926	
BAUJAHR										
	Baualter in Jahren	1/0	-0.021672	-0.021579	-0.020476	-0.022843	-0.018407	-0.021391	-0.020194	-0.020797
	Baualter i.J.quadriert	1/0	0.000284	0.000284	0.000268	0.000334	0.000224	0.000284	0.000265	0.000273
	Baualter i.J. hoch drei	1/0	-0.00000933	-0.00000933	-0.00000871	-0.00001269	-0.00000663	-0.00000970	-0.00000866	-0.00000899
VERDICHTUNG										
	Bevoelkerung	stetig, LN	0.046	0.043	0.049	0.066	0.044	0.048	0.039	0.018
	1-4 Wohneinheiten	1/0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5-20 Wohneinheiten	1/0	0.032	0.021	0.016	0.050	0.036	0.014	0.009	0.006
	ueber 20 Wohneinheiten	1/0	-0.010	-0.026	-0.040	-0.010	0.018	-0.039	-0.065	-0.059
	Metropole	1/0	0.103	X	X	X	X	X	X	X
	Grossstadt	1/0	-0.077	X	X	X	X	X	X	X
	Laendlich	1/0	-0.083	X	X	X	X	X	X	X
	Umland	1/0	0.000	X	X	X	X	X	X	X
	Kernstadt Agglomeration	1/0	X	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Kernstadt gross	1/0	X	-0.187	-0.123	-0.102	-0.068	-0.263	0.009	0.015
	Kernstadt klein	1/0	X	-0.260	-0.208	-0.140	-0.301	-0.289	-0.065	-0.008
	Pendel in	1/0	X	-0.149	-0.115	-0.033	-0.079	-0.178	-0.138	-0.021
	Pendel aus	1/0	X	-0.247	1)	1)	-0.146	-0.321	1)	-0.118
	Laendlich	1/0	X	-0.276	1)	1)	-0.019	-0.376	1)	-0.146
REGION										
	Ostdeutschland	1/0	-0.179	-0.183	-0.098	X	X	X	X	0.033
	Norddeutschland	1/0	-0.028	-0.041	-0.072	X	X	X	X	0.049
	Sueddeutschland	1/0	0.182	0.195	0.160	X	X	X	X	0.192
	Westdeutschland	1/0	0.000	0.000	0.000	X	X	X	X	0.000
	Beschaeftigtenpotential	stetig	X	X	X	X	X	X	X	0.0000071148
	Wachsend	1/0	X	X	0.000	0.000	X	X	0.000	0.000
	Stagnierend	1/0	X	X	-0.129	-0.172	X	X	-0.083	-0.126
	Schrumpfend	1/0	X	X	-0.203	-0.118	X	X	-0.206	-0.237
	Divergierend BevWa	1/0	X	X	-0.169	-0.181	X	X	-0.221	-0.225
	Divergierend BeschWa	1/0	X	X	-0.092	-0.194	X	X	-0.079	-0.143
ANTRAGSDATUM										

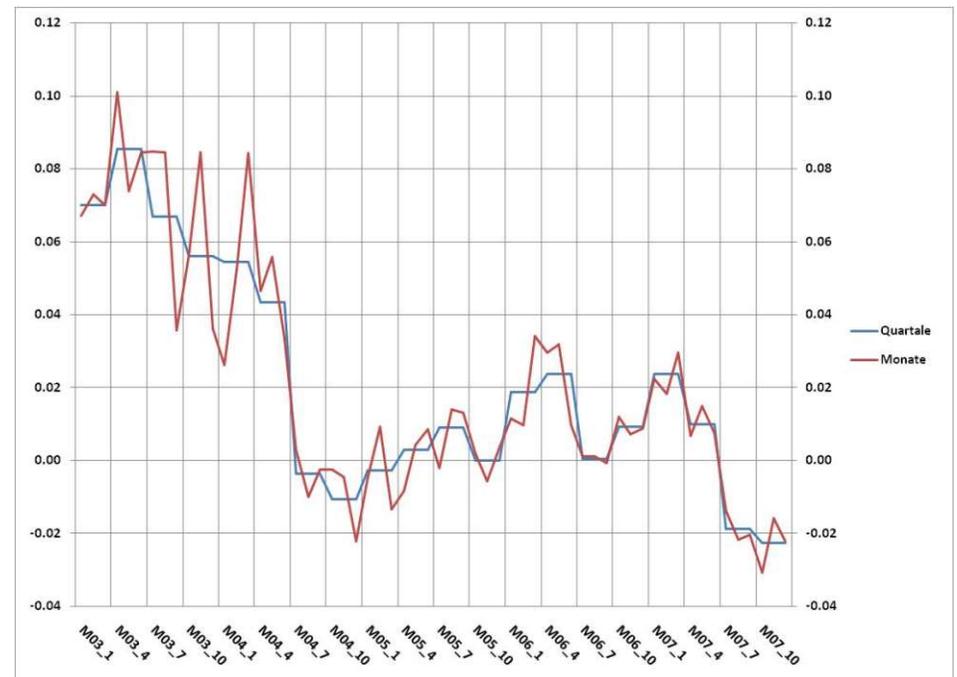
## Zeitvariablen

Monats- vs Quartalsdummies, zeitliche Abgrenzung notwendig → Datensatz ab 2005 verwertbar



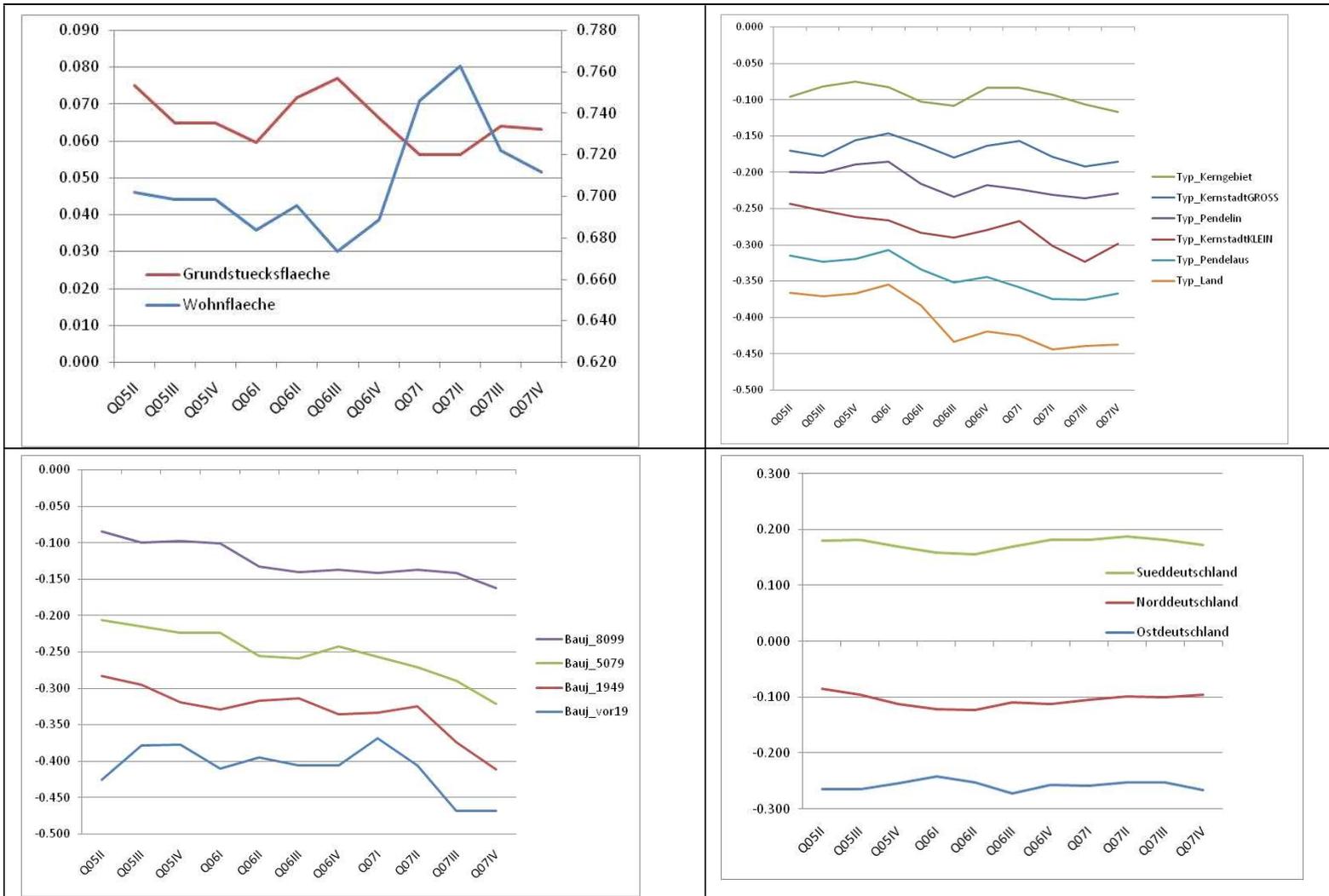
Schätzer

Standardfehler



## Ansatz 2: Bundesweit, benachbarte Regressionen zum Test auf Veränderungen der Beta-Matrix

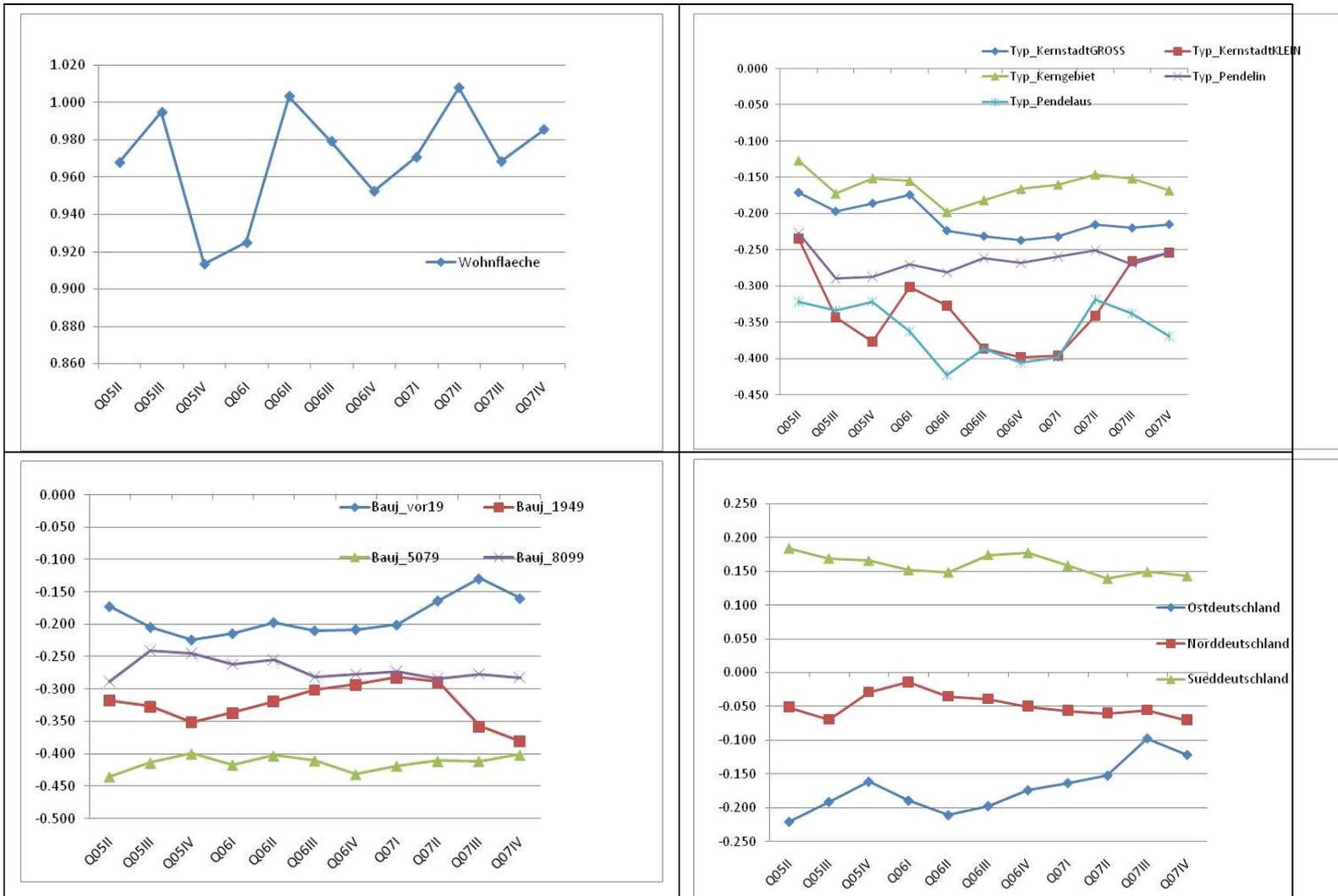
Homes: Probleme bei alten Homes (Modernisierungsgrad, Lagen?), Einfluss der Pendlerpauschale?



Anmerkung: Null ist Westdeutschland, Baujahr 2000 und jünger, Metropole, Selbstnutzung, keine Modernisierungsinvestition, Fahrzeit zum Oberzentrum unter 30 Minuten. Ohne Berücksichtigung der Variable Beschäftigungspotential.

## Ansatz 2: Bundesweit, benachbarte Regressionen zum Test auf Veränderungen der Beta-Matrix

Apartments: Probleme bei Kleinstädten, Aufholwachstum Ost, Zwischenkriegszeit (grosse Deals?)

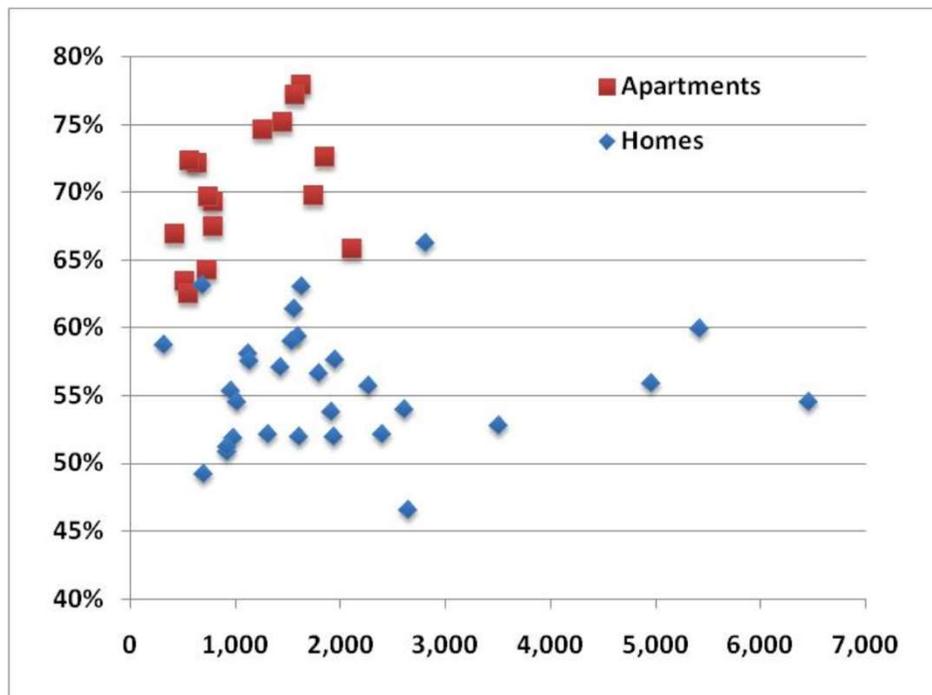


Anmerkung: Null ist Westdeutschland, Baujahr 2000 und jünger, Metropole, Selbstnutzung, keine Modernisierungsinvestition, Fahrzeit zum Oberzentrum unter 30 Minuten. Ohne Berücksichtigung der Variable Beschäftigungspotential.

## Ansatz 3: Berechnung von Regionalindizes

Work in Progress..

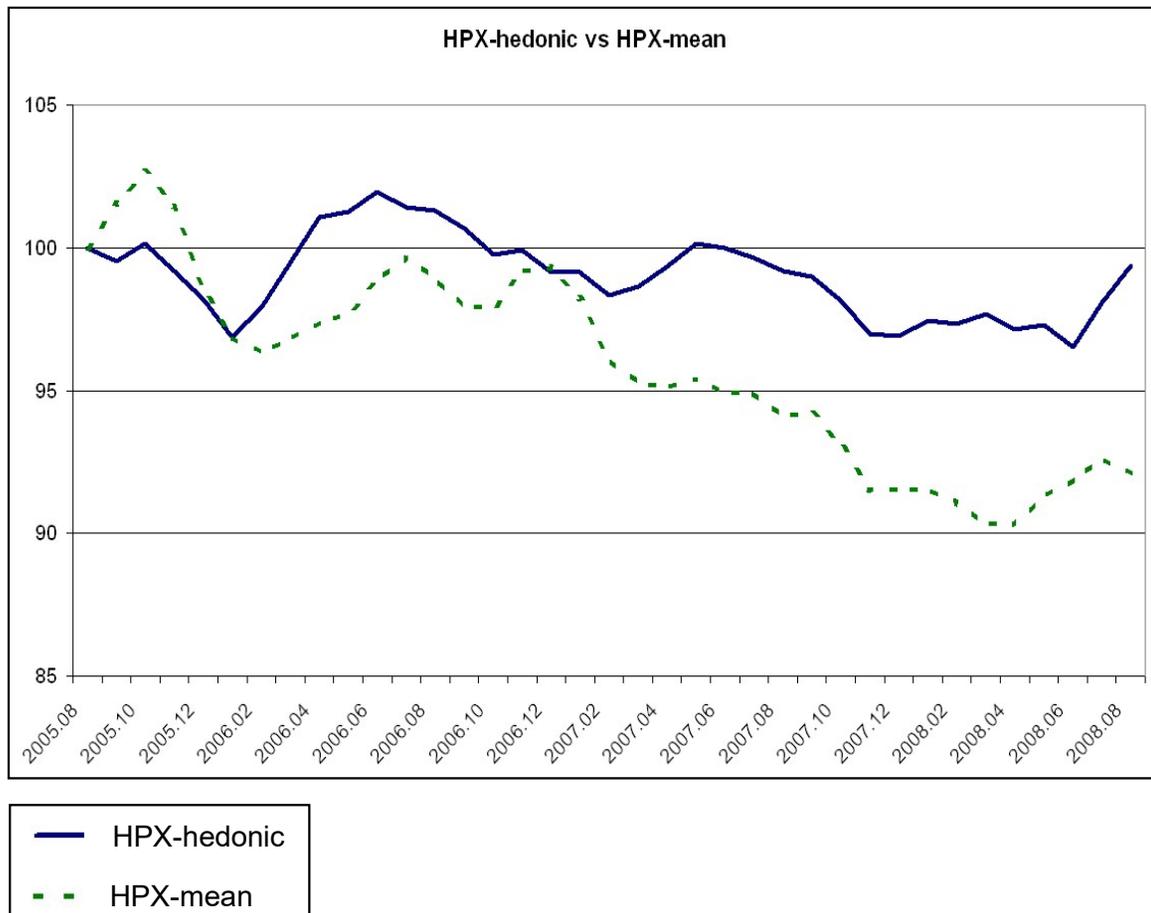
### Fit-Qualität (R-Quadrat) und Teilstichprobengröße bei 30 Regionalregressionen Homes und 16 Regionalregressionen Apartments



- Die Regionen basieren auf Einteilungen des BBR
  - Die Indizes werden wegen der geringeren Stichprobengrößen quartalsweise erstellt
  - Erstanalyse für weite Abgrenzung: 30 Homes und 16 Apartments
  - Momentan in 7 Regionen für Apartments und in 15 Regionen für Häuser (z.B. München, Berlin, Hamburg, Stuttgart, Rhein-Main Gebiet etc.)
- Weiteres Research (verfeinerte Regionsunterteilung) und Stabilisierung der Datenmengen notwendig
- Hochrechnungsfaktoren FEHLEN fuer eine Aggregation (BBR?)

## Ergebnis: gegenwärtige Berechnung des HPX-hedonic

Unterschiede in der Indexentwicklung zu nicht-hedonischem Index



### Zeitvariablenmethode mit

- Bundesregression
- überrollendem Zeitfenster von 8 Quartalen
- Regionsdummyvariablen statt Gewichtungen

### Mögliche Ursachen der unterschiedlichen Preisentwicklung zu altem HPX

- Schwankende regionale Marktanteile (z.B. Sueden) werden durch den Regionsdummy nicht abgefangen
- HPX-mean kann Lageeffekte, z.B. mehr Objekte in Pendelbereichen o.ä. nicht kontrollieren
- HPX-mean inkl. Modernisierungen und z.T. Teilvermietungen

## Welche Erweiterungen wären notwendig und wünschenswert?

Vergleich und Ausblick

Notwendig:

- Gewichtete Schätzung bzw. Hochrechnung, z.B. mit Wohnungsbeständen, bessere Daten zu Wohnungsbeständen und Neubauten, um regionale Einflüsse genauer abzufangen
  - ⇒ Datenbasis bislang nicht detailliert genug, Updates
- Bereinigung saisonaler Einflüsse
  - ⇒ Datenhistorie zu kurz, Verfahren für Bereinigung identifizieren

Wünschenswert:

- Verbesserung der Datenbasis aus EUROPACE (Informationen über Keller, Anzahl Badezimmer, Etagennummer, Fahrstuhl etc.)
  - ⇒ nur möglich wenn Kunden dadurch ein Preisvorteil entsteht
- Verbesserte Dynamisierung der Beta-Matrix (hedonische Imputation mit dynamischen Schattenpreisen)
  - ⇒ Hoher Datenbedarf, in der Literatur verwendete AR-Formulierung unbefriedigend, Dt. Politikrisiko (Strukturbrüche) extrem hoch

Zwischen wünschenswert und notwendig:

- Bottom-up Index nach INSEE-Modell mit ausreichend entwickelten Regionalregressionen und Hochrechnungsfaktoren

**VIelen DANK FÜR IHRE AUFMERKSAMKEIT.**