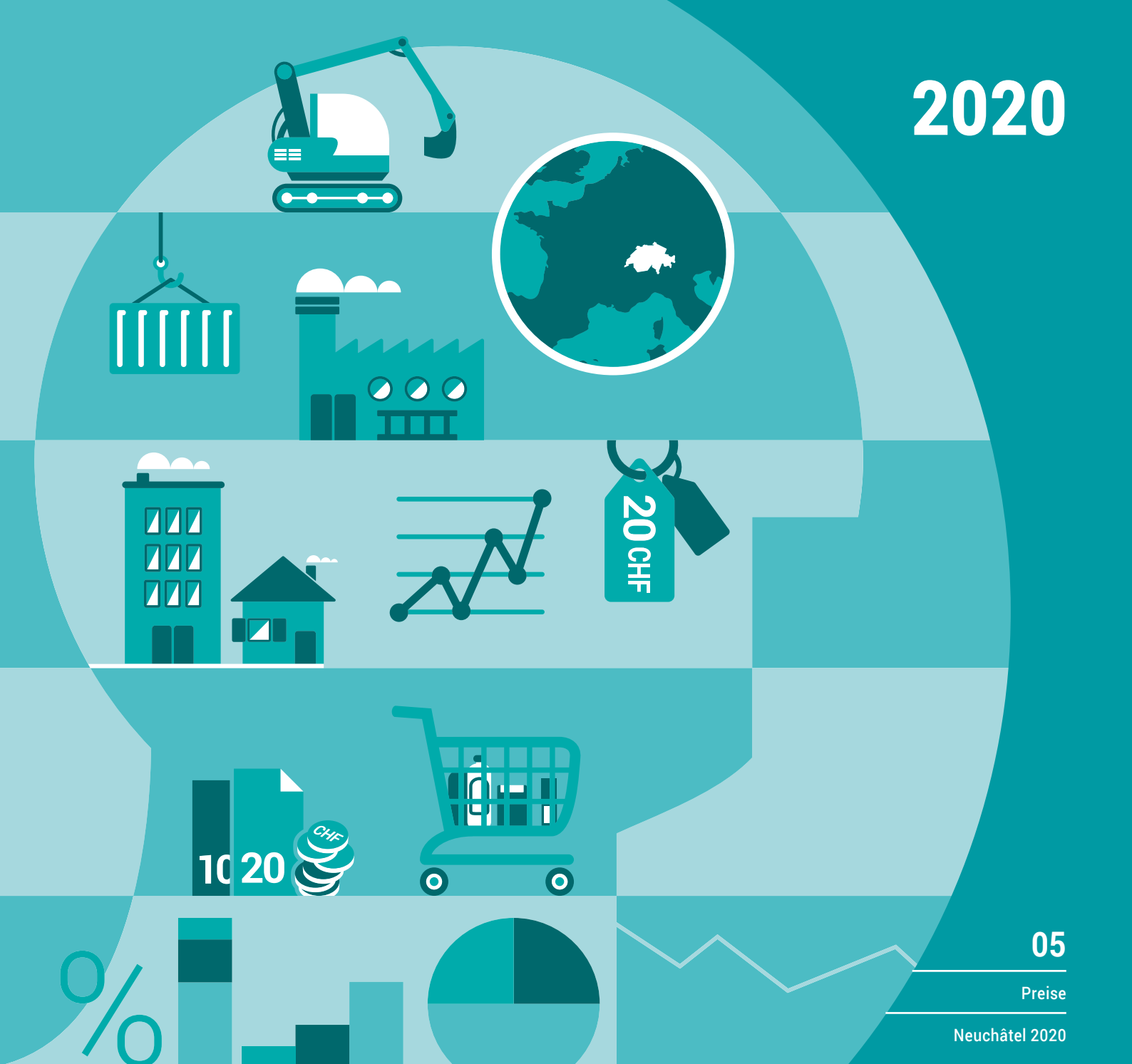


2020



05

Preise

Neuchâtel 2020

Schweizerischer Wohnimmobilienpreisindex

Qualitätsbereinigungsverfahren



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Statistik BFS

Themenbereich «Preise»

Aktuelle themenverwandte Publikationen

Fast alle vom BFS publizierten Dokumente werden auf dem Portal www.statistik.ch gratis in elektronischer Form zur Verfügung gestellt. Gedruckte Publikationen können bestellt werden unter der Telefonnummer 058 463 60 60 oder per E-Mail an order@bfs.admin.ch.

Landesindex der Konsumentenpreise (Dezember 2015 = 100), Methodische Grundlagen, Neuchâtel 2016, 68 Seiten, BFS Nummer: 853-1500

Harmonisierter Verbraucherpreisindex (HVPI), Methodenübersicht und Gewichtung 2018, Neuchâtel 2018, 28 Seiten, BFS Nummer: 930-1800-05

Produzenten- und Importpreisindex Dezember 2015 = 100, Grundlagen, Neuchâtel 2016, 76 Seiten, BFS Nummer: 666-1501

Produzentenpreisindizes für Dienstleistungen (SPPI), Methodische Grundlagen, Neuchâtel 2020, 112 Seiten, BFS Nummer: 2023-2000

Schweizerische Baupreisstatistik. Oktober 2015 = 100, Methodische Grundlagen, Neuchâtel 2016, 60 Seiten, BFS Nummer: 622-1501

2020: Révision de la statistique des prix de la construction, Neuchâtel 2019, 4 Seiten, BFS Nummer: do-d-05.05-NL-24

Vergleichsprogramm von EUROSTAT und der OECD, Kaufkraftparitäten – BIP pro Kopf - Preisniveau, Neuchâtel 2015, 6 Seiten, BFS Nummer: 914-1500

Themenbereich «Preise» im Internet

www.statistik.ch → Statistiken finden → 05 – Preise

Schweizerischer Wohnimmobilienpreisindex

Qualitätsbereinigungsverfahren

Redaktion	Manuel Brand, BFS
Inhalt	Team IMPI, BFS
Herausgeber	Bundesamt für Statistik (BFS)

Neuchâtel 2020

Herausgeber: Bundesamt für Statistik (BFS)
Auskunft: IMPI@bfs.admin.ch, Tel. 058 463 60 69
Redaktion: Manuel Brand, BFS
Inhalt: Team IMPI, BFS
Reihe: Statistik der Schweiz
Themenbereich: 05 Preise
Originaltext: Deutsch
Layout: Sektion DIAM, Prepress/Print
Grafiken: Sektion PREIS
Karten: Sektion PREIS
Online: www.statistik.ch
Print: www.statistik.ch
Bundesamt für Statistik, CH-2010 Neuchâtel,
order@bfs.admin.ch, Tel. 058 463 60 60
Druck in der Schweiz
Copyright: BFS, Neuchâtel 2020
Wiedergabe unter Angabe der Quelle
für nichtkommerzielle Nutzung gestattet
BFS-Nummer: 2071-2002
ISBN: 978-3-303-05764-3

Inhaltsverzeichnis

1	Ausgangslage	5		
<hr/>			3.5.3	Slope 11
			3.5.4	Exposure 11
			3.5.5	LakeView 12
			3.5.6	MountainView 12
			3.5.7	DistanceToLakes 12
			3.5.8	DistanceToRivers 12
			3.5.9	DistanceToHighVoltagePowerLines 12
2	Methodische Grundlagen	6		
<hr/>			3.6	MatchingType 12
2.1	Stratifizierung	6		
2.2	Hedonisches Modell	7		
3	Datengrundlage	8		
<hr/>			4	Vorgehen bei der Modellierung 13
3.1	Erhebung	8	4.1	Deskriptive Analyse 13
3.2	Transaktionsvariablen	9	4.2	Variablentransformation und Modellierung 13
3.2.1	TransactionDate	9	4.3	Erste Analyse der Modellvarianten 14
3.2.2	TransactionPrice	9	4.4	Vertiefte Analyse der besten Modelle 15
3.3	Struktur- und Nutzungsvariablen	9	4.5	Behandlung der Ausreisser 16
3.3.1	ObjectType	9	4.6	Auswahl des definitiven Modells 16
3.3.2	SingleFamilyHouseType	9	4.7	Modellrevisionen 17
3.3.3	CondominiumType	9	4.8	Gutachten 17
3.3.4	PrimaryOrSecondaryHome	9		
3.3.5	OwnerOccupiedOrRented	9	5	Ergebnisse und Modellevaluation 18
3.3.6	YearOfConstruction	9	<hr/>	
3.3.7	LandArea	9	5.1	Deskriptive Analyse 18
3.3.8	VolumeOfBuilding	9	5.2	Hedonisches Modell für Einfamilienhäuser 19
3.3.9	StandardOfVolume	9	5.3	Hedonisches Modell für Eigentumswohnungen 22
3.3.10	NetLivingArea	10	5.4	Allgemeine Modellanalyse 25
3.3.11	NumberOfRooms	10	5.5	Analyse der Residuen 26
3.3.12	NumberOfBathrooms	10	5.6	Analyse der Schätzgenauigkeit 30
3.3.13	NumberOfParkings	10	5.7	Fazit 31
3.3.14	ConstructionQuality	10		
3.3.15	PropertyCondition	10		
3.4	Makrolagevariablen	10		
3.4.1	Canton	10		
3.4.2	MajorStatisticalRegion	10		
3.4.3	CommunityType	10		
3.4.4	SecondAppartementQuota	10		
3.4.5	TaxBurden	11		
3.4.6	TravelTimeToCenters	11		
3.5	Mikrolagevariablen	11		
3.5.1	PublicTransportQuality	11		
3.5.2	NoiseExposure	11		

6	Anhang	32
<hr/>		
6.1	Variablenliste der Ausgangsdaten	32
6.2	Variablenliste Hedonisches Modell Einfamilienhäuser	35
6.3	Variablenliste Hedonisches Modell Eigentumswohnungen	37
6.4	Geoinformationen	39
6.5	Schätzgenauigkeit in den Straten	58
6.5.1	Einfamilienhäuser aufgeteilt nach Gemeindetypen	58
6.5.2	Eigentumswohnungen aufgeteilt nach Gemeindetypen	59
7	Literaturverzeichnis	60
<hr/>		

1 Ausgangslage

Das Bundesamt für Statistik (BFS) liefert Informationen über die Preise und deren Entwicklung in zahlreichen Wirtschaftszweigen, sei es im Konsum, in der Produktion, im Import oder im Baugewerbe. Ein Wirtschaftssektor wurde von der amtlichen Statistik bis anhin allerdings noch nicht vollständig erfasst. Während die Mietpreisentwicklung seit geraumer Zeit durch den offiziellen Mietpreisindex abgedeckt ist, wurde die Preisentwicklung von Immobilientransaktionen bislang nicht berücksichtigt. Dieser Markt ist jedoch äusserst wichtig, sowohl was die getätigten Investitionen als auch die gewährten Hypotheken betrifft. Auf politischen Impuls hin wurde das BFS beauftragt, einen Immobilienpreisindex bzw. einen Preisindex für Wohneigentum zu erstellen. Ab Ende 2020 werden nun erste Quartalsergebnisse veröffentlicht, die Teilindizes für Einfamilienhäuser und Eigentumswohnungen umfassen.

Bei der Erstellung eines Preisindex gilt es, neben den Preisen auch die Qualität der Güter zu betrachten. Tatsächlich ist es so, dass ein Teil des beobachteten Preisunterschieds zwischen den verschiedenen Perioden nicht auf eine reine Preisveränderung, sondern auf die ungleiche Qualität der Güter zurückzuführen ist. Immobilien sind ein äusserst heterogenes Gut, bei dem sich die einzelnen Objekte aufgrund zahlreicher Eigenschaften voneinander unterscheiden. Aus diesem Grund ist die Gefahr einer Qualitätsverzerrung bzw., dass nicht Gleiches mit Gleichem verglichen wird, bei einem Immobilienpreisindex im Vergleich zu Preisindizes für Konsumgüter und Industrieprodukte stark akzentuiert.

Damit es trotzdem möglich ist, die Preise von verschiedenen Liegenschaften miteinander zu vergleichen, werden sogenannte Qualitätsbereinigungsverfahren angewendet. Diese haben zum Ziel, die Preisdifferenzen, die rein durch die abweichende Qualität der einzelnen Objekte verursacht werden, zu neutralisieren und die wahre Preisentwicklung zu extrahieren. Es existieren verschiedene Qualitätsbereinigungsverfahren, die das BFS im Rahmen der konzeptionellen Arbeiten¹ evaluiert hat. Beim Wohnimmobilienpreisindex (IMPI) kommen eine Stratifizierung sowie ein hedonisches Modell des Typs Hedonic Repricing zur Anwendung. Dieses Vorgehen ist auf internationaler Ebene verbreitet, hat sich beim Mietpreisindex des BFS bereits bewährt und wird auch durch die externe Machbarkeitsstudie zur statistischen Erfassung der Immobilienpreise² gestützt.

Im Rahmen des Projekts zur Erstellung eines offiziellen Wohnimmobilienpreisindex (IMPI) galt es, ein effektives Qualitätsbereinigungsverfahren aufzusetzen. Dies beinhaltet neben Wahl der Stratifizierungskriterien insbesondere die Erstellung von hedonischen Modellen für die beiden Objekttypen Einfamilienhäuser und Eigentumswohnungen. Als Grundlage für die Modellierung dienten die erhobenen Transaktionsdaten der Jahre 2017 bis 2019. Es handelt sich um insgesamt 83 324 Transaktionen (35 724 Einfamilienhäuser, 47 600 Eigentumswohnungen). Neben der Implementierung der hedonischen Modelle galt es auch zu definieren, wie und in welchen Zyklen die hedonischen Modelle künftig erneuert werden. Das gesamte Qualitätsbereinigungsverfahren wurde im Nachgang durch einen Experten begutachtet³.

Der vorliegende Bericht stellt die Funktionsweise des Qualitätsbereinigungsverfahrens für den IMPI vor und erläutert die Herleitung und Resultate der hedonischen Modelle.

¹ Bundesamt für Statistik (2016). Detailkonzept Projekt Immobilienpreisindex (auf Anfrage erhältlich)

² ARGE EPFL, Econability and HEG (2012). Machbarkeitsstudie für die statistische Erfassung der Immobilienpreise

³ Silver, M. (2020). Report to the Swiss Federal Statistical Office (FSO) on two hedonic models: the residential property (house) price index (RPPI) and the rent price index (RPI)

2 Methodische Grundlagen

2.1 Stratifizierung

Bei der Methode der Stratifizierung werden die Transaktionen in einzelne Untergruppen, sogenannte Schichten oder Straten aufgeteilt. Für sämtliche Unterkategorien werden Subindizes (bzw. Elementarindizes) berechnet, die anschliessend gewichtet zu einem Gesamtindex aggregiert werden. Anhand der Schichtung wird die Heterogenität verringert und eine Qualitätsbereinigung entlang der gewählten Stratifizierungskriterien durchgeführt. Mittels Stratifizierung können zudem Subindizes für bestimmte Segmente berechnet werden. Im Fall des Wohnimmobilienpreisindex werden die Transaktionen anhand der beiden Variablen Objekttyp und Gemeindetyp strukturiert. Bei Letzterer dienen als Basis die neun Gemeindetypen der offiziellen BFS-Typologie⁴, die dann zu folgenden fünf Kategorien weiter aggregiert werden: Städtische Gemeinde einer grossen Agglomeration, Städtische Gemeinde einer mittelgrossen Agglomeration, Städtische Gemeinde einer kleinen oder ausserhalb einer Agglomeration, Intermediäre Gemeinde, Ländliche Gemeinde. Zusammen mit den beiden Kategorien für den Objekttyp (Einfamilienhäuser und Eigentumswohnungen) ergibt sich eine Stratifizierungsmatrix von zehn Zellen.

Verschiedene Analysen haben gezeigt, dass auf der Grundlage dieser Aggregation relativ homogene Stratifizierungszellen entstehen. Zudem ist sichergestellt, dass für sämtliche Stratifizierungszellen genügend Transaktionen für die Indexberechnung vorhanden sind. Zwei Kriterien, die z.B. bei einer Stratifizierung entlang des Objekttyps und den Kantonen nicht erfüllt wären. Anhand der oben dargestellten Schichtung können Subindizes berechnet und eine Qualitätsbereinigung entlang der beiden einflussreichen Variablen Objekttyp und Gemeindetyp durchgeführt werden. Weil allerdings bei einer einfachen Ex-Post-Stratifizierung nicht alle preisbeeinflussenden Variablen berücksichtigt werden können, ist es mit diesem Verfahren nicht möglich, sämtliche Qualitätsunterschiede zu bereinigen. Aus diesem Grund wird die Stratifizierung zusätzlich mit je einem separaten hedonischen Modell für die Einfamilienhäuser und die Eigentumswohnungen kombiniert.

**Stratifizierungsschema
beim Wohnimmobilienpreisindex** **T1**

	Einfamilienhäuser	Eigentums- wohnungen
Städtische Gemeinde einer grossen Agglomeration	x	x
Städtische Gemeinde einer mittelgrossen Agglomeration	x	x
Städtische Gemeinde einer kleinen Agglomeration oder ausserhalb einer Agglomeration	x	x
Intermediäre Gemeinde	x	x
Ländliche Gemeinde	x	x

© BFS 2020

⁴ Bundesamt für Statistik, Gemeindetypologie und Stadt/Land-Typologie 2012

2.2 Hedonisches Modell

Hedonische Modelle basieren auf der Definition von Gütern als Bündel ihrer verschiedenen Merkmale oder Eigenschaften. Bei Immobilien sind damit sowohl Informationen zur physischen Struktur, zur Nutzung als auch zur Lage der Objekte gemeint (vgl. Kapitel 3). Anhand dieser Objekteigenschaften lässt sich die Qualität der Liegenschaften erfassen. Umgekehrt können die Preise der Immobilien durch ihre Eigenschaften geschätzt werden, ähnlich wie bei einem Warenkorb, dessen Preis durch seinen Inhalt bzw. die Einzelpreise der darin enthaltenen Produkte bestimmt wird. Der einzige Unterschied besteht darin, dass die Preise der verschiedenen Objekteigenschaften von Immobilien nicht einzeln beobachtet werden können. Die marginalen bzw. impliziten Preise lassen sich allerdings anhand von Regressionen eruieren. Die daraus resultierende hedonische Gleichung soll den Preis jeder Immobilie bestmöglich reproduzieren. Mittels der impliziten Preise kann anschliessend die Qualität der gehandelten Objekte bewertet und eine Qualitätsbereinigung durchgeführt werden. Die hedonische Gleichung weist folgende Form auf:

$$p_{it} = \beta x_{it} + \mu_{it}$$

- p_{it} Transaktionspreis von Immobilie i in Periode t
- x_{it} Vektor der erklärenden Variablen (Struktur, Nutzung, Lage) für die Immobilie i in Periode t
- β Vektor der Koeffizienten der erklärenden Variablen (implizite Preise)
- μ_{it} Fehlerterm für die Transaktion der Immobilie i in Periode t

Es gibt verschiedene Ansätze, um anhand von hedonischen Modellen Preisindizes zu erstellen. Grob gesehen können sie in die Time Dummy-, die Characteristics Prices-, die Hedonic Imputation- und die Hedonic Repricing-Methode unterteilt werden. Verschiedene Analysen während der Konzeptphase⁵ haben gezeigt, dass die Methode des Hedonic Repricing⁶ für das Vorhaben des BFS die passendste ist. Bei dieser Methode werden die Preisveränderungen innerhalb der einzelnen Straten zusätzlich durch die Entfernung von Qualitätsunterschieden bereinigt bzw. purifiziert. Hierzu wird für jede Zelle der Stratifizierung ein Index für die Bruttopreisveränderungen sowie ein Index für die Qualitätsveränderungen berechnet. Der Quotient dieser beiden Indizes entspricht anschliessend der qualitätsbereinigten Preisveränderung. Der Index der Qualitätsveränderungen wird anhand einer hedonischen Gleichung ermittelt, in die die Eigenschaften sämtlicher Objekte aus der entsprechenden Zelle, die in der Periode verkauft wurden,

eingesetzt werden. Die Formel zur Berechnung eines Index I für eine Stratifizierungszelle l in der Periode t sieht folgendermassen aus:

$$I_l = \frac{(\prod_{i \in N^t} p_i^t)^{\frac{1}{N^t}}}{(\prod_{i \in N^0} p_i^0)^{\frac{1}{N^0}}} \div \frac{(\prod_{i \in N^t} \hat{p}_{i|z_i^t}^b)^{\frac{1}{N^t}}}{(\prod_{i \in N^0} \hat{p}_{i|z_i^0}^b)^{\frac{1}{N^0}}}$$

- p_i^0 Transaktionspreis von Immobilie i in Periode 0
- p_i^t Transaktionspreis von Immobilie i in Periode t
- \hat{p}_i^b Vektor der Koeffizienten der erklärenden Variablen (implizite Preise) geschätzt für Periode b
- z_i^0 Vektor der erklärenden Variablen (Struktur, Nutzung, Lage) für die Immobilie i in Periode 0
- z_i^t Vektor der erklärenden Variablen (Struktur, Nutzung, Lage) für die Immobilie i in Periode t

Der erste Quotient der Formel entspricht dem Bruttopreisindex. Dieser setzt sich aus dem geometrischen Mittel der Transaktionspreise p der aktuellen Periode t im Zähler sowie dem geometrischen Mittel der Transaktionspreise in der Basisperiode 0 im Nenner zusammen. Der zweite Quotient der Gleichung stellt den Index der Qualitätsveränderungen dar. Er besteht aus dem geometrischen Mittel der geschätzten Preise der Periode t geteilt durch die geschätzten Preise der Basisperiode 0 . Die verschiedenen Subindizes der einzelnen Straten l können anschliessend anhand einer Gewichtung w zu übergeordneten Indizes aggregiert werden. Dabei werden die Zellen anhand ihres Anteils am Transaktionsvolumen aus dem Vorjahr gewichtet.

$$I_N = \sum_l w_l I_l$$

Aufgrund der Tatsache, dass die hedonische Gleichung beim Ansatz des Hedonic Repricing einzig zur Gewichtung der Objektmerkmale und zur Ableitung der Qualitätsbereinigungsfaktoren verwendet wird, muss sie im Unterschied zu anderen hedonischen Verfahren, die die Gleichung zur Schätzung der qualitätsbereinigten Preisveränderungen verwenden, nicht in jeder Periode neu berechnet werden. Dies bedeutet, dass eine breitere Stichprobe zur Schätzung des hedonischen Modells beigezogen werden kann. Zudem erlaubt das stabile Modell eine Berechnung der Indizes allein mit den periodenspezifischen Preisen und Qualitätsmerkmalen.

⁵ Bundesamt für Statistik (2016). Detailkonzept Projekt Immobilienpreisindex (auf Anfrage erhältlich)

⁶ Im Gutachten zu den hedonischen Modellen wird das gewählte Verfahren als Hedonic Imputation dargestellt. Dieser Ansatz ist unter gewissen Voraussetzungen äquivalent zum Ansatz des Hedonic Repricing. Da die Darstellungsweise der Formel mit je einer separaten Komponente zur Berechnung des Bruttopreisindex und einer Komponente zur Berechnung des Qualitätsindex vom Hedonic Repricing herrührt, hält das BFS weiter an dieser Bezeichnung fest. Einen detaillierteren Vergleich der beiden Methoden kann dem Gutachten von Mick Silver entnommen werden.

3 Datengrundlage

In diesem Kapitel wird auf die für die Herleitung der beiden hedonischen Modelle verwendeten Daten und deren Beschaffung eingegangen. Detailliertere Informationen zur Kategorisierung der einzelnen Variablen können dem Anhang entnommen werden.

3.1 Erhebung

Für die Erstellung und Produktion des neuen Wohnimmobilienpreisindex werden verschiedene Daten benötigt. Zusätzlich zu den Informationen über die einzelnen Transaktionen (Transaktionsdatum und Transaktionspreis) bedarf es ergänzender Daten für die Qualitätsanpassung. Dazu gehören Angaben zur physischen Struktur (Strukturvariablen), zur Nutzung (Nutzungsvariablen) und zum Standort (Lagevariablen) der Immobilien. Eine effektive Qualitätsanpassung und ein methodisch valider Immobilienpreisindex sind nur möglich, wenn die Qualität der erhobenen Immobilien realistisch wiedergegeben wird. Im Rahmen der Konzeptionsphase untersuchte das BFS, welche Institutionen Immobilientransaktionen erfassen und welche davon als Datenquellen für den neuen Index dienen könnten.

Diese Analyse hat gezeigt, dass für die Erstellung eines offiziellen Wohnimmobilienpreisindex verschiedene Datenquellen benötigt werden. Die Kerninformationen über Transaktionsdaten, Preise und physische Eigenschaften der Immobilien werden vierteljährlich bei den grössten Hypothekarinstituten (26 Banken) erhoben. Dies ist die einzige Quelle, die die erforderlichen Transaktionsdaten, einschließlich der Preise, vollständig und zeitnah erfasst. Hypothekarinstitute führen auch zusätzliche Informationen über die Qualität der von ihnen finanzierten Immobilien in ihren Akten. Da die operativen und strategischen Geschäftsentscheidungen der Hypothekeninstitute auf diesen Daten aufbauen, sind sie von hoher Qualität.

Die von den Hypothekarinstituten erhobenen Transaktions-, Struktur- und Nutzungsvariablen werden durch verschiedene Informationen über die Standortqualität der Immobilien ergänzt. Diese Daten stammen aus Geoinformationsdatensätzen, die von der Bundesverwaltung frei zur Verfügung gestellt werden. Die Anreicherung der Bankdaten mit den geolokalisierten Datensätzen basiert auf der Objektadresse als Identifikationsmerkmal. Aufgrund von Bedenken auf Seiten der Datenlieferanten wegen einer möglichen Verletzung des Datenschutzes und des Bankkündengeheimnisses, erhebt das BFS die Adressinformationen jedoch nicht mit. In Zusammenarbeit mit der Schweizerischen Bankiervereinigung hat das BFS ein Open Source Tool entwickelt, das es den meldepflichtigen Hypothekarinstituten ermöglicht, Daten ohne die Adresse der Objekte zu liefern. Dieses Instrument, einschliesslich einer Datenbankdatei mit den Adressen aller relevanten Liegenschaften in der Schweiz zusammen mit den entsprechenden Standortvariablen, stellt das BFS den Datenlieferanten zur Verfügung. Die Datenbankdatei wird vierteljährlich durch das BFS aktualisiert. Die Adressen sind dem eidgenössischen Gebäude- und Wohnungsregister (GWR) entnommen. Die Standortvariablen stammen aus öffentlich zugänglichen Geodatensätzen (vgl. Kapitel 3.4. und 3.5). Aus Datenschutzgründen stellt das BFS sicher, dass keine Kombination von Standortvariablen in der Datenbankdatei weniger als dreimal vorkommt. Die Standortvariablen wurden so kategorisiert, dass sie aber dennoch genügend Aussagekraft im hedonischen Modell entfalten. Um die Datenlieferung vorzubereiten, importieren die Banken ihre Transaktionsdaten inklusive Objektadresse in das Tool. Das Tool validiert die Eingabedaten und reichert sie über die Objektadresse mit Informationen aus der Datenbankdatei an. Die resultierende Ausgabedatei enthält schliesslich keine Objektadressen mehr und ist als anonymisierter Datensatz für den Versand an das BFS bereit.⁷

⁷ Für weitere Informationen zum Open-Source-Erhebungstool sowie zur Datenbankdatei konsultieren sie bitte <https://github.com/bfs-preis/impi/wiki/> Einführung. Zudem können dort sowohl der Code des Tools als auch die Datei eingesehen werden.

3.2 Transaktionsvariablen

3.2.1 TransactionDate

Als Transaktionsdatum wird das Auszahlungsdatum der ersten Kredittranche verwendet. Die Variable dient der eindeutigen Zuweisung der Transaktionen zu einem Quartal.

3.2.2 TransactionPrice

Als Transaktionspreis gilt der finale Kaufpreis von Freihandtransaktionen in Schweizer Franken. Schenkungen, Vorzugspreise, Zwangsversteigerungen, Teilkäufe und Vorerbschaften werden ausgeschlossen. Transaktionskosten wie Grunderwerbssteuern, Notar- und Grundbuchgebühren sowie Maklerprovisionen werden dabei nicht berücksichtigt.

3.3 Struktur- und Nutzungsvariablen

Strukturvariablen umfassen Informationen, die die physischen Eigenschaften einer Immobilie beschreiben. Nutzungsvariablen beinhalten Informationen zur künftigen Nutzung der gehandelten Objekte. Das BFS erhebt diese Variablen zusammen mit dem Transaktionsdatum und dem Transaktionspreis bei den Hypothekarinstituten.

3.3.1 ObjectType

In der Kategorie Wohneigentum unterscheidet das BFS zwischen Einfamilienhäusern und Eigentumswohnungen. Die Variable wird als Stratifizierungskriterium verwendet, wobei für beide Objekttypen Subindizes berechnet werden. Zudem werden separate hedonische Modelle für die Einfamilienhäuser und Eigentumswohnungen berechnet.

3.3.2 SingleFamilyHouseType

Die Banken unterscheiden bei Einfamilienhäusern weitere Unterkategorien. Dazu gehören z.B. die Typologien freistehendes Einfamilienhaus, Reihemittelhaus, Reiheneckhaus oder Doppelhaus. Der größte gemeinsame Nenner ist die Unterscheidung zwischen freistehenden und zusammengebauten Einfamilienhäusern. Die für die Modellierung verwendete Variable umfasst daher diese beiden Kategorien.

3.3.3 CondominiumType

Die Banken unterscheiden bei den Eigentumswohnungen weitere Unterkategorien. Dazu gehören z.B. die Typologien Erdgeschosswohnung, Gartenwohnung, Etagenwohnung, Attikawohnung oder Dachgeschosswohnung. Der größte gemeinsame Nenner ist die

Unterscheidung zwischen Wohnungen im obersten Stockwerk und sämtlichen anderen Wohnungen. Die für die Modellierung verwendete Variable umfasst daher diese beiden Kategorien.

3.3.4 PrimaryOrSecondaryHome

Diese Variable enthält die Information, ob das Objekt als Erst- oder Zweitwohnung verwendet wird. Mit der Verabschiedung der Zweitwohnungsinitiative im Jahr 2012 stimmte die Schweiz dafür, den Bau von Zweitwohnungen einzuschränken. In Gemeinden mit einem Zweitwohnungsanteil von mehr als 20% (siehe Variable Zweitwohnungsquote) werden keine neuen Zweitwohnungen mehr genehmigt. Es stellt sich daher die Frage, ob sich die Preise für Zweitwohnungen in den von der 20-Prozent-Quote betroffenen Gemeinden anders entwickeln als die Preise für Erstwohnungen.

3.3.5 OwnerOccupiedOrRented

Diese Variable enthält die Information, ob ein Objekt zur Selbstnutzung oder zur Vermietung vorgesehen ist.

3.3.6 YearOfConstruction

Diese Variable enthält Informationen, wann das entsprechende Objekt errichtet wurde. Insgesamt werden sieben Bauperioden unterschieden.

3.3.7 LandArea

Einfamilienhäuser verfügen per Definition über einen Grundstücksanteil, der nur dem konkreten Objekt zugeordnet werden kann. Für Eigentumswohnungen gibt es jedoch keine individuell zuordenbare Landfläche. Daher wird die Parzellenfläche nur für Einfamilienhäuser berücksichtigt. Sie wird in Quadratmetern gemessen.

3.3.8 VolumeOfBuilding

Die Größe von Einfamilienhäusern wird als Gebäudevolumen in Kubikmetern erfasst.

3.3.9 StandardOfVolume

In der Schweiz gibt es drei verschiedene Normen zur Messung des Gebäudevolumens (SIA 116, SIA 416 und Gebäudeversicherungsnorm). Da das Gebäudevolumen von ein und demselben Objekt je nach gewählter Norm stark variieren kann, ist es sinnvoll, die Variable als Zusatz zum Gebäudevolumen zu verwenden.

3.3.10 NetLivingArea

Die Grösse von Eigentumswohnungen wird als Nettowohnfläche in Quadratmetern angegeben.

3.3.11 NumberOfRooms

Diese Variable enthält die Anzahl der Zimmer sowohl für Einfamilienhäuser als auch für Eigentumswohnungen. Die Art und Weise, wie die Anzahl der Räume erfasst wird, unterscheidet sich jedoch von Datenlieferant zu Datenlieferant und von Region zu Region. In einigen Fällen werden Küchen als halbe Räume gezählt, in anderen ist dies nicht der Fall. Es ist daher angebracht, auf die nächste ganze Zahl abzurunden.

3.3.12 NumberOfBathrooms

Diese Variable enthält die Anzahl der Badezimmer bzw. Nasszellen (inkl. separater Toiletten) sowohl für Einfamilienhäuser als auch für Eigentumswohnungen.

3.3.13 NumberOfParkings

Diese Variable enthält die Anzahl der Garagen, Parkplätze und Carports sowohl für Einfamilienhäuser als auch für Eigentumswohnungen.

3.3.14 ConstructionQuality

Die Banken erfassen den Ausbaustandard zum Zeitpunkt der Transaktion. Dazu verwenden sie ein Bewertungssystem. Da sich die Notenskalen von Hypothekarinstitut zu Hypothekarinstitut unterscheiden, hat das BFS die Bewertungen von 1 bis 100 standardisiert.

3.3.15 PropertyCondition

Die Banken erfassen den Gebäude- bzw. Wohnungszustand zum Zeitpunkt der Transaktion. Dazu verwenden sie ein Bewertungssystem. Da sich die Notenskalen von Institut zu Institut unterscheiden, hat das BFS die Bewertungen von 1 bis 100 standardisiert.

3.4 Makrolagevariablen

Lagevariablen geben Auskunft über die Qualität des Objektsstandorts. Wir unterscheiden dabei zwischen Makro- und Mikro-lagevariablen. Bei Makrolagevariablen handelt es sich um Lagevariablen, die Gültigkeit für Gemeinden oder für übergeordnete Gebietskörperschaften haben.

3.4.1 Canton

Es handelt sich hierbei um eine kategorielle Variable, die Auskunft darüber gibt, in welchem Kanton ein Objekt steht.

3.4.2 MajorStatisticalRegion

Die sieben statistischen Grossregionen sind vom BFS definierte Bezugsräume in der Schweiz, die auf einer hierarchischen Ebene über den Kantonen stehen. Es handelt sich um eine Aggregation basierend auf den Kantonen. In der Nomenklatur der Europäischen Union entspricht diese Einteilung der NUTS 2 Ebene⁸.

3.4.3 CommunityType

Die Gemeindetypen sind eine Klassifikation der Schweizer Gemeinden⁹, die vom BFS erstellt und veröffentlicht wird. Ihre Differenzierung geschieht anhand der Kriterien Dichte, Grösse und Erreichbarkeit. Das Kriterium der Gemeindetypen wird für den Schweizer Wohnimmobilienpreisindex neben dem Objekttyp als zweites Stratifizierungskriterium verwendet. Für ihre Zellen werden ebenfalls Subindizes berechnet. Zu diesem Zweck werden die ursprünglich neun Gemeindetypen zu fünf Kategorien weiter aggregiert (vgl. Kapitel 2.1).

3.4.4 SecondAppartementQuota

Hierbei handelt es sich um eine Dummy-Variable, die Aufschluss darüber gibt, ob der Anteil der Zweitwohnungen in der Gemeinde 20% übersteigt. Die Rohdaten bezüglich der Zweitwohnungsquote pro Gemeinde wird vom Bundesamt für Raumentwicklung (ARE) aufbereitet¹⁰. In Kombination mit der Variable PrimaryOrSecondaryHome dient die Information zur Identifizierung der Zweitwohnungen in Gemeinden mit einem Anteil an Zweitwohnungen von mehr als 20%.

⁸ eurostat, Nomenclature of territorial units for statistics (NUTS)

⁹ Bundesamt für Statistik, Gemeindetypologie und Stadt/Land-Typologie 2012

¹⁰ Bundesamt für Raumentwicklung, Zweitwohnungen

3.4.5 TaxBurden

Die Eidgenössische Steuerverwaltung (ESTV)¹¹ veröffentlicht statistische Informationen über die durchschnittliche Steuerbelastung pro Gemeinde für verschiedene Haushaltstypen und Einkommensklassen. Die vorliegende Variable enthält Informationen über die Steuerbelastung in Prozent des jährlich zu versteuernden Einkommens für eine verheiratete Person mit zwei Kindern und einem jährlich zu versteuernden Bruttoarbeitseinkommen von 100 000 Franken. Für den IMPI werden drei Kategorien unterschieden (niedrige, mittlere und hohe Steuerbelastung).

3.4.6 TravelTimeToCenters

Das Bundesamt für Raumentwicklung (ARE)¹² gibt pro Gemeinde Auskunft über die durchschnittliche Fahrzeit in Minuten mit den öffentlichen Verkehrsmitteln und dem motorisierten Individualverkehr zur nächstgelegenen Kernstadt oder zum nächstgelegenen Agglomerationszentrum. Als Kernstädte gelten Genf, Lausanne, Bern, Basel, Zürich und Lugano. Die Kategorie der Agglomerationszentren umfasst rund 50 Städte, die für das Umland von einer gewissen Bedeutung sind. Die beim Wohnimmobilienpreisindex verwendete Variable basiert auf dem standardisierten Durchschnitt der vier Reisezeitkategorien Öffentlicher Verkehr zu Kernstädten, Öffentlicher Verkehr zu Agglomerationszentren, Motorisierter Individualverkehr zu Kernstädten, Motorisierter Individualverkehr zu Agglomerationszentren. Die standardisierte Variable wird anschliessend in drei Kategorien unterteilt.

3.5 Mikrolagevariablen

Mikrolagevariablen umfassen alle Lagevariablen mit einer räumlichen Auflösung feiner als auf der kommunalen Ebene. Um diese Daten einem bestimmten Standort zuordnen zu können, werden sie als Geoinformationen erfasst.

3.5.1 PublicTransportQuality

Die ÖV-Güteklassen sind ein Indikator des ARE¹³ für die Beurteilung der Erschliessung mit dem öffentlichen Verkehr. Zu diesem Zweck werden alle Haltestellen des öffentlichen Verkehrs in der Schweiz kategorisiert. Dies geschieht anhand der Kriterien Art des Verkehrsmittels (Bahnknotenpunkte, Bahnlinien, Trams, Busse, Postautos, Bereitschaftsbusse, Schiffe und Seilbahnen) und Kursintervall. Um die Erreichbarkeit der Haltestellen für Fussgänger zu berücksichtigen, wird in einem zweiten Schritt die Distanz zur Haltestelle als weiterer Parameter hinzugefügt. Dieser wird in konzentrischen Kreisen als lineare Entfernung zur Haltestelle gemessen und dient in Kombination mit der Haltestellenkategorie zur Bestimmung der Qualitätsklassen des öffentlichen Verkehrs.

¹¹ Eidgenössische Steuerverwaltung, Steuerbelastung in den Gemeinden

¹² Bundesamt für Raumentwicklung, Reisezeiten

¹³ Bundesamt für Raumentwicklung, Reisezeiten

Dabei ist zu beachten, dass sich diese konzentrischen Kreise über maximal einen Kilometer erstrecken. Alle Punkte, die weiter entfernt liegen, werden der Kategorie «keine Erreichbarkeit» zugeordnet. Insgesamt werden fünf Kategorien unterschieden.

3.5.2 NoiseExposure

Das Bundesamt für Umwelt (BAFU)¹⁴ und das Bundesamt für Zivilluftfahrt (BAZL)¹⁵ haben eine umfassende Geodatenbank zur Lärmbelastung in der Schweiz erstellt. Berücksichtigt werden die Lärmquellen Strassen-, Eisenbahn- und Flugverkehr. Für die drei Lärmarten werden jeweils Daten für den Tag (6 bis 22 Uhr) und für die Nacht (22 bis 6 Uhr) veröffentlicht. Für den Wohnimmobilienpreisindex verwendet das BFS aggregierte Lärmbelastungsdaten. Zu diesem Zweck werden der maximale Tag- und der maximale Nachtwert der drei Lärmarten in die Day-Night-Average-Sound-Level-Formel¹⁶ eingefügt. Diese Formel trägt der Tatsache Rechnung, dass Lärm in der Nacht grundsätzlich als störender empfunden wird, und bestraft Nachtlärm mit einem Zuschlag von 10 Dezibel. Die resultierende Lärmbelastung wird anschliessend in drei Kategorien eingeteilt.

3.5.3 Slope

Diese Variable enthält Informationen über die Steilheit des Geländes am jeweiligen Objektstandort. Die Hangneigung (in Grad) wird aus dem digitalen Höhenmodell swissALTI3D¹⁷ des Bundesamtes für Landestopografie (swisstopo) abgeleitet. Für den Wohnimmobilienpreisindex wird die Hangneigung anschliessend in drei Kategorien eingeteilt. Die Hangneigung kann sowohl positiven als auch negativen Einfluss auf den Immobilienpreis haben. Einerseits werden durch die Steilheit des Geländes die Baukosten in die Höhe getrieben. Andererseits werden durch Hangneigung und Exposition (vgl. Kapitel 5.5.4) besonders gute Hanglagen definiert.

3.5.4 Exposure

Die Exposition bzw. die Ausrichtung eines Hanges wird ebenso wie die Hangneigung aus dem digitalen Höhenmodell swissALTI3D von swisstopo abgeleitet. Die Expositionsrichtung wird in Grad Abweichung von der Nordrichtung gemessen und bewegt sich deshalb innerhalb des Spektrums von 0 bis 360 Grad. Für den Wohnimmobilienpreisindex werden schliesslich zwei Kategorien unterschieden (keine Exposition bzw. Exposition nach Norden, Exposition nach Süden). Südhänge weisen dabei eine gute Sonneneinstrahlung auf.

¹⁴ Bundesamt für Umwelt, Lärmbelastung

¹⁵ Bundesamt für Zivilluftfahrt, Lärmbelastungskataster

¹⁶ $Lr_{DN} = 10 \log \left[\frac{15}{24} \times 10^{(Lr_D/10)} + \frac{9}{24} \times 10^{((Lr_N/10)/10)} \right]$

¹⁷ Bundesamt für Landestopografie (swisstopo), Digitales Höhenmodell swissALTI3D

3.5.5 LakeView

Sicht auf einen See ist ein gewichtiges Verkaufsargument bei Immobilien. Aus dem digitalen Höhenmodell swissALTI3D von swisstopo können diejenigen Bereiche bestimmt werden, die von einem bestimmten Ort aus sichtbar sind. Diese werden mit den Seeflächen verschnitten, um anschliessend die sichtbare Seefläche in Hektaren berechnen zu können. In der finalen Variable werden drei Kategorien unterschieden.

3.5.6 MountainView

Wie bei der Seesicht wird auch die Variable Bergsicht aus dem digitalen Höhenmodell swissALTI3D von swisstopo abgeleitet. Zusätzlich wird das digitale Höhenmodell EU-DEM¹⁸ der Europäischen Umweltagentur verwendet, um die angrenzenden Regionen der Nachbarländer Frankreich, Italien, Österreich, Deutschland und Liechtenstein abzudecken. Die sichtbaren Bereiche der digitalen Höhenmodelle werden mit einem Geoinformationslayer verschnitten, der die dominantesten¹⁹ Gipfel der Schweiz und der angrenzenden Regionen der vier Nachbarländer enthält. Schlussendlich werden für die Variable Bergsicht drei Kategorien unterschieden.

3.5.7 DistanceToLakes

Diese Variable gibt Auskunft darüber, ob sich ein Einfamilienhaus oder eine Eigentumswohnung in weniger als 100 Metern Entfernung zum nächsten See befindet oder nicht. Die Geoinformationen zu den Seen werden aus dem Topografische Landschaftsmodell swissTLM3D von swisstopo bezogen.

3.5.8 DistanceToRivers

Diese Variable gibt Auskunft darüber, ob sich ein Einfamilienhaus oder eine Eigentumswohnung unter 100 Metern Entfernung zum nächsten Fluss befindet oder nicht. Die Geoinformationen zu den Flüssen werden aus dem Topografische Landschaftsmodell swissTLM3D von swisstopo bezogen.

3.5.9 DistanceToHighVoltagePowerLines

Diese Variable gibt Auskunft darüber, ob sich ein Einfamilienhaus oder eine Eigentumswohnung in weniger als 100 Metern Entfernung zur nächsten Hochspannungsleitung befindet oder nicht. Die Geoinformationen zu den Hochspannungsleitungen werden aus dem Topografische Landschaftsmodell swissTLM3D von swisstopo bezogen.

3.6 MatchingType

Diese Variable wird erstellt, wenn die Datenlieferanten ihre Daten in das vom BFS zur Verfügung gestellte Tool importieren. Sie enthält Informationen über die Qualität des Adressabgleichs zwischen den Inputdaten des Datenlieferanten und den Adressen des Gebäude- und Wohnungsregisters. Es werden vier Kategorien unterschieden. Die Kategorie Point-Matching enthält Objekte, deren Adresse das Tool exakt zuordnen konnte. Wenn kein Point-Matching möglich ist, versucht das Tool, die Lagevariablen des Objekts zu imputieren, das der Mitte der Strasse (Center-Street-Matching) oder der Mitte der Gemeinde (Center-Community-Matching) am nächsten liegt. Falls auch diese Versuche fehlschlagen, fällt das Objekt in die Kategorie No-Matching. In letztem Fall können keine Lagevariablen angereichert werden. Bei der Erstellung des hedonischen Modells werden nur Objekte berücksichtigt, die ein Point-Matching oder ein Center-Street-Matching aufweisen. Im Gegensatz zur Modellierung werden bei der Indexberechnung auch Center-Community-Matchings berücksichtigt.

¹⁸ European Environment Agency, Digital surface model EU-DEM

¹⁹ Die Dominanz eines Berggipfels entspricht der horizontalen Distanz bis zum nächsten höheren Berggipfel.

4 Vorgehen bei der Modellierung

In nachfolgenden Kapitel wird die Vorgehensweise bei der Erstellung der beiden hedonischen Modelle erläutert. Was hier als linearer Prozess dargestellt wird, war in Tat und Wahrheit ein iterativer Prozess, bei dem die verschiedenen Etappen nach neuen Erkenntnissen jeweils wiederholt wurden.

4.1 Deskriptive Analyse

In einem allerersten Schritt werden die erhobenen Daten, die als Basis für die Erstellung der hedonischen Modelle dienen, analysiert und auf allfällige Lücken, Fehler und Auffälligkeiten überprüft. Grundsätzlich können beim Modellieren nur Transaktionen verwendet werden, die keine Missings oder eindeutige Fehler bei den verwendeten Variablen aufweisen. Je nach Variablenumfang eines Modells werden deshalb in dieser Testphase auch unterschiedlich viele Transaktionen berücksichtigt. Neben Missings und Fehlern bei Variablen kann die Nichtverwendbarkeit einer Transaktion auch auf sogenannte No-Matchings zurückzuführen sein. Bei diesen Fällen konnte das Erhebungstool keinen erfolgreichen Adressabgleich vornehmen. In der Folge weist das Objekt keine Lagevariablen auf. Aus diesem Grund könnten die betroffenen Transaktionen nur in Modellen ohne Lagevariablen verwendet werden. Neben den No-Matchings werden bei der Modellierung auch Transaktionen mit einem Center-Community-Matching ausgeschlossen. Dies, weil die imputierten Lagevariablen, die vom Objekt stammen, das dem geographischen Mittelpunkt der Gemeinde am nächsten liegt, erheblich von den wirklichen Lagevariablen abweichen können.

Neben der Überprüfung der Verwendbarkeit der einzelnen Transaktionen ist es auch wichtig, sich ein gesamthafes Bild bezüglich der Qualität der Daten zu machen. Dabei werden beispielsweise die Verteilungen der einzelnen Variablen angeschaut. Zudem werden die Korrelationen zwischen der abhängigen Variable (Preis) und den unabhängigen Variablen (Struktur-, Nutzungs-, Lagevariablen) bzw. zwischen den einzelnen unabhängigen Variablen studiert. Die aus diesen Analysen gewonnenen Erkenntnisse fließen dann in die Modellierung und in die Interpretation der Resultate ein.

4.2 Variablentransformation und Modellierung

Nach der deskriptiven Analyse der Daten werden die vorhandenen Variablen auf verschiedene Art und Weise transformiert und aggregiert. So werden die linearen Variablen (VolumeOfBuilding, LandArea, NetLivingArea) beispielsweise logarithmiert oder quadriert. Bei kategoriellen Variablen bieten sich alternative Aggregationen der Kategorien an. Anschliessend werden sämtliche als sinnvoll erachteten Variablenkombinationen modelliert. Um sich ein Bild über den Einfluss einzelner Variablen machen zu können, werden zusätzlich die Verfahren der Vorwärtsselektion¹⁹ und der Rückwärtselimination²⁰ beigezogen. Neben den einzelnen Variablen werden auch Interaktionen²¹ zwischen verschiedenen Variablen getestet.

¹⁹ Bei der Vorwärtsselektion werden die unabhängigen Variablen eine nach der anderen in das Modell aufgenommen. Dabei wird jeweils diejenige Variable ausgewählt, die die grösste partielle Korrelation mit der abhängigen Variable aufweist. Dieses Prozedere wird wiederholt, bis ein Abbruchkriterium erfüllt ist oder alle Variablen in das Modell aufgenommen wurden.

²⁰ Bei der Rückwärtselimination wird mit allen verfügbaren Variablen gestartet. Anschliessend werden die unabhängigen Variablen eine nach der anderen aus dem Modell entfernt. Dabei wird jeweils diejenige Variable aus dem Modell entfernt, die die kleinste partielle Korrelation mit der abhängigen Variable aufweist. Dieses Prozedere wird wiederholt, bis ein Abbruchkriterium erfüllt ist oder keine Variable mehr im Modell enthalten ist.

²¹ Von Interaktionen wird gesprochen, wenn der Effekt einer unabhängigen Variable auf die abhängige Variable vom Wert einer anderen unabhängigen Variable abhängt. In einer Regressionsgleichung wird ein Interaktionseffekt als das Produkt von zwei oder mehr unabhängigen Variablen dargestellt:

$$\hat{y} = b_0 + b_1X_1 + b_2X_2$$

4.3 Erste Analyse der Modellvarianten

Aus den verschiedenen Variablenkombinationen resultieren für beide Objekttypen rasch einige hundert Modell-Varianten. Selbstverständlich wäre es nicht zielführend, diese Modelle alle von Hand zu analysieren und zu beurteilen. Aus diesem Grund werden die Modelle zuerst anhand verschiedener Kennzahlen beurteilt. Für sämtliche der unten aufgeführten und erläuterten Kennzahlen zur Modellgüte werden die jeweils ungefähr 10 besten Modelle ausgewählt.

R-Squared (R^2)

Das R^2 ist ein statistisches Maß dafür, wie nahe die beobachteten Daten an der geschätzten Regressionslinie liegen. Bei der Berechnung des R^2 wird die anhand des Modells erklärte Varianz durch die beobachtete Varianz dividiert.

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$

Das R-Quadrat liegt immer zwischen 0 und 100%, wobei bei 100% das Modell die gesamte beobachtete Varianz erklärt. Das R^2 hat die Eigenschaft, dass es mit jeder Variable steigt, die zum Modell hinzugefügt wird. Würde man die verschiedenen Modelle nur aufgrund dieser Kennzahl beurteilen, wäre es demnach optimal, das Modell mit den meisten erklärenden Variablen zu wählen. Dies birgt die Gefahr der Überanpassung (Overfitting) des Modells.

Adjusted R-Squared (Adj. R^2)

Ein Gütemass, das neben der Modellanpassung auch das Problem der Überanpassung berücksichtigt, ist das Adjusted R-Squared. Es besteht aus dem R^2 sowie einem Strafterm, der mit jeder zum Modell hinzugefügten Variable weiter ansteigt. Aus diesem Grund nimmt das Adj. R^2 einen geringeren Wert als das R^2 an. Beim Hinzufügen einer neuen Variable steigt das Adj. R^2 nur dann an, wenn der zusätzliche Erklärungsgehalt den Anstieg des Strafterms ausgleichen kann. Im Strafterm werden die Anzahl Transaktionen (n) sowie die Anzahl Variablen (k) berücksichtigt.

$$Adj. R^2 = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - k - 1} \right]$$

Akaike Information Criterion (AIC)

Das Akaike Information Criterion vergleicht die verschiedenen Modellkandidaten anhand der Werte der log-Likelihood-Funktion (\hat{L}). Diese steigen, je grösser der Erklärungsgehalt des Modells ist. Daneben wird auch ein Strafterm berücksichtigt, der aus der Anzahl der im Modell enthaltenen Parameter (k) besteht. Das Modell mit dem kleinsten AIC wird bevorzugt.

$$AIC = 2k - 2\ln(\hat{L})$$

Bayesian Information Criterion (BIC)

Das Bayesian Information Criterion basiert ebenfalls auf der log-Likelihood-Funktion (\hat{L}) und unterscheidet sich nur beim Strafterm gegenüber dem Akaike Information Criterion. Das BIC ist neben der im Modell enthaltenen Parameter auch von der Anzahl Transaktionen (n) abhängig. Bereits ab einer Stichprobe von acht Observationen bestraft das BIC stärker als das AIC.

$$BIC = \ln(n)k - 2\ln(\hat{L})$$

Mean Absolute Error (MAE)

Der Mean Absolute Error (MAE) einer Regression entspricht dem arithmetischen Mittel der absoluten Differenz zwischen den Vorhersagewerten (\hat{y}) und den Beobachtungswerten (y) des berücksichtigten Samples.

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j|$$

4.4 Vertiefte Analyse der besten Modelle

Die aufgrund der im Kapitel 4.3 erwähnten Kennzahlen ausgewählten Modelle werden in einem nächsten Schritt genauer betrachtet. Dabei werden die Modelle nun einzeln und von Hand analysiert. Zuerst werden dabei die Koeffizienten der einzelnen Modellvariablen genauer studiert und plausibilisiert. Es stellt sich dabei die Frage, ob die Koeffizienten den theoretisch zu erwartenden Einflüssen entsprechen (z.B. hat Lärm im Modell wirklich einen negativen Einfluss auf den Preis?). Daneben wird auch die Signifikanz der einzelnen Koeffizienten angeschaut. Zudem wird anhand des Variance Inflation Factors (VIF)²² überprüft, ob die verwendeten unabhängigen Variablen Multikollinearität²³ aufweisen.

Lineare Regressionen unterliegen verschiedenen Annahmen, die erfüllt sein sollten, damit die Methode ordnungsgemäss funktioniert. Dazu gehören neben der Abwesenheit von Multikollinearität, dem linearen Zusammenhang zwischen den metrischen unabhängigen Variablen und der abhängigen Variable auch die Normalverteilung und Homoskedastizität der Residuen²⁴. Bei einer Verletzung einer der beiden letzten Annahmen sind die Schätzer zwar weiterhin konsistent, allerdings setzt beispielsweise der t-Test eine Normalverteilung und Homoskedastizität der Residuen voraus. Aus diesem Grund ist es wichtig, die beiden Hypothesen bei den ausgewählten Modellen zu überprüfen.

Die Normalverteilung der Residuen wird anhand eines QQ-Plots der Residuen beurteilt. Liegen die Residuen im Plot auf einer diagonalen Geraden, sind sie normalverteilt. Die Normalverteilungshypothese kann auch anhand eines Histogramms der Residuen mit darübergelegter Normalverteilungskurve überprüft werden. Neben den graphischen Check's wird der Jarque-Bera-Test²⁵ beigezogen.

Die Homoskedastizität der Residuen wird ebenfalls grafisch und mit Tests überprüft. Hierzu wird ein Scatterplot der Residuen versus die geschätzten Preisen verwendet. Wenn die Abbildung eine zufällige Streuung der Observationen ohne Clusterbildung und systematische Muster aufweist, kann Homoskedastizität angenommen werden. Da die Interpretation von Diagrammen immer auch eine subjektive Komponente aufweist, werden zusätzlich der Breusch-Pagan-Test²⁶ und eine Version des White Tests²⁷ zur Überprüfung der Homoskedastizitätsannahme beigezogen. Daneben werden auch sämtliche unabhängigen Variablen gegen die Residuen geplottet. Es ist nicht überraschend, wenn zu diesem Zeitpunkt der Modellierung nicht normalverteilte und heteroskedastische Residuen beobachtet werden können. Eine Massnahme zur Reduktion der Heteroskedastizität ist die Logarithmierung der abhängigen Variablen. Zudem haben Ausreisser einen grossen Einfluss auf die beiden zu Grunde liegenden Annahmen.

²² Anhand des Variance-Inflation-Factor-Tests (VIF) kann die Multikollinearität zwischen verschiedenen Regressionsvariablen gemessen werden. Mathematisch gesehen ist der VIF für eine Regressionsmodellvariable gleich dem Verhältnis der gesamten Modellvarianz zur Varianz eines Modells, das nur diese eine unabhängige Variable enthält.

²³ Von Multikollinearität wird dann gesprochen, wenn es in einem Modell eine starke Korrelation zwischen zwei oder mehreren unabhängigen Variablen gibt. Bei starker Multikollinearität, wird die Schätzung der Regressionskoeffizienten unsicher, da die betroffenen Variablen zumindest teilweise denselben Teil der Varianz erklären.

²⁴ Die Residuen entsprechen dem Teil des Transaktionspreises, der durch das hedonische Modell nicht erklärt wird bzw. der Differenz zwischen dem beobachteten und dem geschätzten Transaktionspreis.

²⁵ Der Jarque-Bera-Test prüft anhand der Schiefe und Kurtosis ob eine Normalverteilung vorliegt. Vgl. Bera, A. K. and C. M. Jarque (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. In *Economic Letters*. Nr.3, p. 255–259.

²⁶ Breusch, T.S. and A. R. Pagan (1980). A simple test for heteroscedasticity and random coefficient variation. In: *Journal of the Econometric Society*, *Econometrica*, p. 817–838

²⁷ Wooldridge, J. M. (2013), *Introductory Econometrics, a modern approach*, 5th edition, South-Western, chap. 8.3, p. 280

4.5 Behandlung der Ausreisser

Bis zu diesem Zeitpunkt wurden nur Transaktionen mit Missings aus dem Datensatz entfernt. Daneben sind aber auch auffällige Transaktionen genauer anzuschauen. In einem nächsten Schritt werden die Daten deshalb auf Ausreisser überprüft und wenn nötig bereinigt. Wir unterscheiden dabei zwischen uni- und multivariaten Ausreissern. Bei univariaten Ausreissern handelt es sich um einen auffällig hohen bzw. niedrigen Wert in einer Variable (z.B. 200 Zimmer). Solche Fälle sind relativ einfach herauszufiltern. Komplizierter wird es hingegen bei der Detektion von multivariaten Ausreissern. Damit sind Transaktionen gemeint, die aufgrund der Kombination verschiedener Variablen auffallen. Dabei können die einzelnen Ausprägungen durchaus Sinn machen (z.B. 6 Zimmer; 30m² Nettowohnfläche).

In einem ersten Schritt kümmern wir uns um die multivariaten Ausreisser. Das Verfahren, das dabei angewendet wird, heisst Cook's Distance²⁸. Die Outlier detection wird basierend auf dem hedonischen Modell durchgeführt. Cook's Distance gibt wieder, wie stark sich die Residuen verändern, falls eine bestimmte Observation aus der Schätzung des Regressionsmodells entfernt wird. Dadurch können Observationen ausfindig gemacht werden, die einen grossen Einfluss auf das Schätzmodell haben. Die Formel zur Berechnung der Cook's Distance für eine Transaktion lautet folgendermassen:

$$D_i = \frac{\{\beta - \beta_i\}^T X^T X \{\beta - \beta_i\}}{ps^2}$$

D_i	Cook's Distance für Transaktion i
X	Matrix ($n \times p$) mit den Werten der unabhängigen Variablen
T	Transposition einer Matrix bzw. eines Vektors
β	Vektor ($p \times 1$) der Least Squares Schätzer
β_i	Vektor ($p \times 1$) der Least Squares Schätzer ohne Berücksichtigung der Transaktion i im hedonischen Modell
p	Anzahl der unabhängigen Variablen plus 1
s^2	Geschätzte Varianz unter Verwendung des ganzen Datensatzes

Das Resultat der Analyse ist ein Cook's Distance Wert für sämtliche Transaktionen des Datensatzes. Grundsätzlich gilt: Mit hohen Cook's Distance Werten steigt die Wahrscheinlichkeit, dass es sich um einen Outlier handelt. Beim Wohnimmobilienpreisindex werden sämtliche Transaktionen, die eine Cook's Distance grösser als «4/Anzahl Transaktionen» aufweisen, aus dem Sample gelöscht. Diese «cut-off»-Grenze ist weit verbreitet (z.B. offizieller Immobilienpreisindex Irland²⁹) und wird unter anderem auch von Bollen und Jackman³⁰ propagiert.

Im Anschluss an die Eliminierung der multivariaten Ausreisser werden in einem zweiten Schritt gewisse Regeln zur Auffindung der allenfalls übrig gebliebenen univariaten Ausreisser angewendet. Dabei geht es vor allem darum, extreme Werte bei der abhängigen, wie auch bei den wichtigsten unabhängigen Variablen (VolumeOfBuilding, LandArea und NetLivingArea) zu eliminieren, die allenfalls die hedonischen Modelle verzerren könnten.

4.6 Auswahl des definitiven Modells

Nach der Ausreisserbehandlung werden die ausgewählten Modelltypen mit dem reduzierten Datensatz neu kalibriert. Wegen den bearbeiteten Stichproben können sich auch die Kennzahlen der einzelnen Modelle verändern. Durch die Entnahme von Ausreissern sollte sich die Schätzqualität verbessern. Es kann aber auch zu Veränderungen bei den Koeffizienten kommen. Deshalb müssen die neu kalibrierten Modelle nochmals analog zum Vorgehen in Kapitel 4.4. analysiert werden. Dazu gehört auch eine Überprüfung der der linearen Regression zu Grunde liegenden Hypothesen. Nach der Ausreisserbehandlung sollten sich die zu Grunde liegenden Annahmen für die lineare Regression verbessert haben. Aufgrund der vorliegenden Resultate wird anschliessend ein definitives Modell ausgewählt.

²⁸ Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics* Vol. 19 (1): p. 15–18

²⁹ O'Hanlon, N. (2011). Constructing a National House Price Index for Ireland. *Journal of the Statistical and Social Inquiry Society of Ireland*, Vol XL, Central Statistics Office and Centre for Policy Studies

³⁰ Bollen, K. A. and R. W. Jackman (1990). Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases. *Modern methods of data analysis*, p. 257–291

4.7 Modellrevisionen

Auch wenn die ökonometrischen Modelle beim Ansatz des Hedonic Repricing eine gewisse Zeit lang stabil belassen werden können, gilt es zu berücksichtigen, dass sich die impliziten Preise der Qualitätseigenschaften mittel- bis langfristig verändern können. Um diesem Aspekt Rechnung zu tragen, wird das Hedonic Repricing-Modell in regelmässigen Abständen unter Berücksichtigung der neu zur Verfügung stehenden Daten kalibriert. Zusätzlich wird das BFS, parallel zur Anwendung der Methode des Hedonic Repricing einen zweiten Index mit der Rolling-Time-Dummy-Methode berechnen. Die Rolling-Time-Dummy-Methode ist eine Abwandlung bzw. Erweiterung der klassischen Time-Dummy-Methode, bei der neben den Struktur-, Nutzungs- und Lagevariablen auch periodenspezifische Dummy-Variablen in das hedonische Modell integriert werden. Das hedonische Modell wird dabei in jeder Periode neu berechnet. Als Basis dienen jeweils die Transaktionen aus der aktuellen und den drei vorangegangenen Perioden. Die Preisentwicklung kann direkt aus den periodenspezifischen Dummies abgeleitet werden. Der Rolling-Time-Dummy-Index wird nicht veröffentlicht und dient einzig als interner Benchmark sowie zur Überwachung der Entwicklung der impliziten Preise.

4.8 Gutachten

Im Nachgang zur Erstellung der Ausgangsmodelle mit den Daten der Jahre 2017 bis 2019 wurden die Resultate einem Gutachten unterzogen. Das Gutachten wurden von Mick Silver, einem emeritierten Professor für Economic Statistics an der Cardiff University und Senior Economist beim Internationalen Währungsfonds (IWF) durchgeführt. Aus Datenschutzgründen hatte der Experte keinen Zugang zu den Daten. Als Grundlage für das Gutachten wurden aggregierte Auswertungen sowie ein detaillierter Beschrieb der Vorgehensweise bereitgestellt. Ausserdem standen die Vertreter des BFS für alle Fragen zur Verfügung.

5 Ergebnisse und Modellevaluation

Im nachfolgenden Kapitel werden die Ergebnisse der hedonischen Modelle für Einfamilienhäuser und Eigentumswohnungen dargestellt. Dabei handelt es sich um die Ausgangsmodelle, die auf der Grundlage der erhobenen Transaktionsdaten der Jahre 2017 bis 2019 erstellt wurden. Berücksichtigt wurden insgesamt 83 324 Transaktionen (35 724 Einfamilienhäuser; 47 600 Eigentumswohnungen). Diese Modelle werden zur Berechnung der ersten Indexpunkte von 2019 bis 2020 verwendet.

Für die beiden Objekttypen Einfamilienhäuser und Eigentumswohnungen wurden insgesamt je etwas über 500 verschiedene Modelle berechnet. Für sämtliche der in Kapitel 4.3 aufgelisteten Kennzahlen wurden jeweils die zehn besten Modelle pro Objekttyp ausgewählt. Diese Modelle wurden anschliessend genauer analysiert. Die aufgeführten Resultate beziehen sich auf die definitiven hedonischen Modelle für die beiden Objekttypen Einfamilienhäuser und Eigentumswohnungen.

5.1 Deskriptive Analyse

Im Rahmen der deskriptiven Analyse stellte sich heraus, dass die Variablen `NumberOfParkings`, `CondominiumType` sowie `OwnerOccupiedOrRented` eine zu geringe Abdeckung aufweisen, um sie in der Modellierung zu verwenden. Zum heutigen Zeitpunkt sind leider nicht alle Hypothekarinstitute in der Lage, (flächendeckende) Angaben zu diesen drei Variablen zu liefern. Bei der Anzahl Parkplätze kommt erschwerend hinzu, dass die Kriterien, was als Parkplatz gezählt wird, nicht überall die gleichen sind. So werden beispielsweise Aussenparkplätze auf dem Grundstück nicht immer in die Berechnung einbezogen. Dasselbe gilt für Garagen, die im Gebäudevolumen inbegriffen sind. Um den Einfluss dieser drei Variablen dennoch testen zu können, wurden auch hedonische Modelle mit Subsamples gemacht, die nur die Transaktionen umfassen, die No-Missings enthalten. Insbesondere die Anzahl Parkplätze zeigten dabei einen signifikanten Einfluss auf den Preis. Falls sich die Anzahl Missings künftig reduzieren sollte, wird eine Aufnahme der Variablen in das hedonische Modell neu geprüft.

Abgesehen von den drei erwähnten Variablen weisen die Daten der Jahre 2017 bis 2019 eine gute Qualität auf. Wobei festgehalten werden muss, dass für die Jahre 2017 und 2018 einzelne Datenlieferanten noch nicht in der Lage waren, flächendeckende und vollständig Datensätze zu liefern. Seit anfangs 2019 ist aber eine lückenlose Datenerhebung bei sämtlichen Lieferanten gewährleistet. Zudem ist die Datenqualität sowie der Anteil der durch das Erhebungstool richtig gemachten Transaktionen über die drei Jahre kontinuierlich angestiegen. Bei den Einfamilienhäusern flossen insgesamt 30 673 Transaktionen in das finale hedonische Modell ein. Dies entspricht 85,86% der erhobenen Daten. Bei den Eigentumswohnungen ist der Anteil der verwendeten Transaktionen mit 83,55% (39 771) etwas geringer. Dies ist auf das etwas schlechtere Matching bei den Eigentumswohnungen zurückzuführen.

	Einfamilienhaus	Eigentumswohnung
Ausgangsstichprobe:	35 724 (100,0%)	47 600 (100,0%)
Center-Community-Matchings:	-2 252 (-6,30%)	-4 386 (-9,21%)
No-Matchings:	-344 (-0,96%)	-529 (-1,11%)
Missings/Fehler:	-545 (-1,53%)	-427 (-0,90%)
Ausreisser:	-1 910 (-5,35%)	-2 487 (-5,22%)
Endstichprobe:	30 673 (85,86%)	39 771 (83,55%)

© BFS 2020

5.2 Hedonisches Modell für Einfamilienhäuser

Das definitive hedonische Modell für die Einfamilienhäuser umfasst insgesamt 78 Variablen. Von den flächendeckend zur Verfügung stehenden Variablen wurden ausser der Variable `SingleFamilyHouseType`, die sich als nicht signifikant erwies, sämtliche anderen Variablen verwendet. Eine mögliche Interpretation besteht darin, dass der Erklärungsgehalt der Unterscheidung zwischen freistehenden und zusammengebauten Einfamilienhäusern bereits in der Variable `LandArea` enthalten ist. Freistehende Einfamilienhäuser dürften im Schnitt eine grössere Grundstücksfläche aufweisen als zusammengebaute Einfamilienhäuser.

Die beiden linearen Variablen `VolumeOfBuilding` und `LandArea` fliessen logarithmiert in das Modell ein. Genauso wie die abhängige Variable (Preis). Die Anzahl Zimmer werden kategorisiert (3 oder weniger Zimmer, 4 Zimmer, 5 Zimmer, 6 Zimmer, 7 Zimmer, 8 und mehr Zimmer). Dasselbe gilt für die Anzahl Badezimmer bzw. Nasszellen (1 Nasszelle, 2 Nasszellen, 3 Nasszellen, 4 Nasszellen, 5 oder mehr Nasszellen). Die beiden Variablen `SecondAppartementQuota` und `PrimaryOrSecondaryHome` werden zu drei Kategorien kategorisiert (Erstwohnungen, Zweitwohnungen in Gemeinden mit mehr als 20% Zweitwohnungen, Zweitwohnungen in Gemeinden mit weniger als 20% Zweitwohnungen). Die restlichen Variablen fliessen so in das Modell ein, wie sie erhoben werden. Andere Transformationen haben sich nicht als zielführend herausgestellt. Dasselbe gilt für die Verwendung von Interaktionsvariablen.

Zu guter Letzt wird noch eine Time-Dummy-Variable für das Erhebungsjahr in das Modell aufgenommen. Deren Einfluss auf den Erklärungsgehalt des Modells ist gering. Allerdings hilft diese Variable die Koeffizienten der restlichen Variablen zu bereinigen. Die Time-Dummy-Variable wird bei der Indexberechnung nicht mitberücksichtigt. Dasselbe gilt für die Variable `PrimaryOrSecondaryHome_Quota`. Auch wenn die Regulierung des Zweitwohnungsmarktes Einfluss auf die Immobilienpreise hat, sollen bei der Qualitätsbereinigung nur jene Merkmale berücksichtigt werden, die die Qualität des Objekts direkt beeinflussen. Bei der Modellerstellung, wo die Effizienz und Konsistenz der Schätzung wichtig ist, werden hingegen möglichst sämtliche preisbestimmenden Variablen verwendet. Eine ähnliche Diskussion ergibt sich bei der Variable `TaxBurden`. Im Unterschied zur Diskussion bei den Zweitwohnungen kann die Steuerbelastung allerdings auch als Proxy für andere lagespezifischen Merkmale herhalten. Nach vertiefter Analyse und Diskussion mit dem Gutachter wurde entschieden, die Variable `TaxBurden` sowohl im Modell als auch bei der Indexberechnung zu berücksichtigen.

Variable	Estimate	Std.Error	t-value	Pr(> t)
Intercept	9,0315657	0,04179125	216,111	<0,0000000000000002 ***
Ln_VolumeOfBuilding	0,47398024	0,00564127	84,02	<0,0000000000000002 ***
StandardOfVolume_1	0,07421288	0,00394849	18,795	<0,0000000000000002 ***
StandardOfVolume_2	0,08329522	0,00422233	19,727	<0,0000000000000002 ***
Ln_LandArea	0,14654441	0,00262138	55,904	<0,0000000000000002 ***
PrimaryOrSecondaryHome_Quota_2	0,22913116	0,0107528	21,309	<0,0000000000000002 ***
PrimaryOrSecondaryHome_Quota_3	-0,11580631	0,01160152	-9,982	<0,0000000000000002 ***
YearOfConstruction_2	0,13364008	0,00649185	20,586	<0,0000000000000002 ***
YearOfConstruction_3	0,15375814	0,00570923	26,932	<0,0000000000000002 ***
YearOfConstruction_4	0,23412193	0,00558198	41,942	<0,0000000000000002 ***
YearOfConstruction_5	0,3092231	0,00602537	51,32	<0,0000000000000002 ***
YearOfConstruction_6	0,37843752	0,00686483	55,127	<0,0000000000000002 ***
YearOfConstruction_7	0,3383864	0,00729143	46,409	<0,0000000000000002 ***
NumberOfRooms_4	0,07382005	0,00789634	9,349	<0,0000000000000002 ***
NumberOfRooms_5	0,11566851	0,00789912	14,643	<0,0000000000000002 ***
NumberOfRooms_6	0,14000158	0,00831512	16,837	<0,0000000000000002 ***
NumberOfRooms_7	0,16201799	0,00921115	17,589	<0,0000000000000002 ***
NumberOfRooms_8	0,1701903	0,01012952	16,801	<0,0000000000000002 ***
NumberOfBathrooms_2	0,05775446	0,00308096	18,746	<0,0000000000000002 ***
NumberOfBathrooms_3	0,10995569	0,0048095	22,862	<0,0000000000000002 ***

Variable	Estimate	Std.Error	t-value	Pr(> t)
NumberOfBathrooms_4	0,17138035	0,0105044	16,315	<0,0000000000000002 ***
NumberOfBathrooms_5	0,61796326	0,0275311	22,446	<0,0000000000000002 ***
ConstructionQuality	0,00325168	0,0000699	46,522	<0,0000000000000002 ***
PropertyCondition	0,00082646	0,00006111	13,524	<0,0000000000000002 ***
Canton_1	0,51358608	0,01920283	26,745	<0,0000000000000002 ***
Canton_2	0,27419527	0,01540323	17,801	<0,0000000000000002 ***
Canton_3	0,50668523	0,01669259	30,354	<0,0000000000000002 ***
Canton_4	0,36709181	0,04601584	7,978	0,00000000000000154 ***
Canton_5	0,49904647	0,02223519	22,444	<0,0000000000000002 ***
Canton_6	0,69902152	0,03537493	19,76	<0,0000000000000002 ***
Canton_7	0,48610032	0,03917479	12,408	<0,0000000000000002 ***
Canton_8	0,35433787	0,0267348	13,254	<0,0000000000000002 ***
Canton_9	0,87675578	0,02488961	35,226	<0,0000000000000002 ***
Canton_10	0,24816734	0,01733942	14,312	<0,0000000000000002 ***
Canton_11	0,27104698	0,01590651	17,04	<0,0000000000000002 ***
Canton_12	0,62712209	0,01926245	32,557	<0,0000000000000002 ***
Canton_13	0,4583321	0,01771147	25,878	<0,0000000000000002 ***
Canton_14	0,32454673	0,01958518	16,571	<0,0000000000000002 ***
Canton_15	0,39960485	0,02087386	19,144	<0,0000000000000002 ***
Canton_16	0,55255587	0,07829395	7,057	0,00000000000173191 ***
Canton_17	0,31430402	0,01773879	17,718	<0,0000000000000002 ***
Canton_18	0,41525643	0,01937179	21,436	<0,0000000000000002 ***
Canton_19	0,34442079	0,01730893	19,898	<0,0000000000000002 ***
Canton_20	0,34591512	0,01790865	19,316	<0,0000000000000002 ***
Canton_21	0,21384552	0,02080995	10,276	<0,0000000000000002 ***
Canton_22	0,49545262	0,01584552	31,268	<0,0000000000000002 ***
Canton_23	0,11524008	0,02036098	5,66	0,00000001528603916 ***
Canton_24	0,30459629	0,02039254	14,937	<0,0000000000000002 ***
Canton_25	0,5846959	0,02071608	28,224	<0,0000000000000002 ***
CommunityType_2	-0,18488423	0,00506899	-36,474	<0,0000000000000002 ***
CommunityType_3	-0,27443007	0,00578743	-47,418	<0,0000000000000002 ***
CommunityType_4	-0,16528598	0,00563255	-29,345	<0,0000000000000002 ***
CommunityType_5	-0,19815456	0,00493073	-40,188	<0,0000000000000002 ***
CommunityType_6	-0,25998289	0,006051	-42,965	<0,0000000000000002 ***
CommunityType_7	-0,31751818	0,00785078	-40,444	<0,0000000000000002 ***
CommunityType_8	-0,34812416	0,00616471	-56,47	<0,0000000000000002 ***
CommunityType_9	-0,41491464	0,00943306	-43,985	<0,0000000000000002 ***
TaxBurden_2	-0,13175254	0,00869791	-15,148	<0,0000000000000002 ***
TaxBurden_3	-0,19010493	0,01159549	-16,395	<0,0000000000000002 ***
TravelTimeToCenters_2	-0,09467984	0,00353978	-26,747	<0,0000000000000002 ***
TravelTimeToCenters_3	-0,15340326	0,00489575	-31,334	<0,0000000000000002 ***

Variable	Estimate	Std.Error	t-value	Pr(> t)
PublicTransportQuality_2	-0,08153461	0,00792745	-10,285	<0,0000000000000002 ***
PublicTransportQuality_3	-0,13211196	0,00761033	-17,36	<0,0000000000000002 ***
PublicTransportQuality_4	-0,16864596	0,00767559	-21,972	<0,0000000000000002 ***
PublicTransportQuality_5	-0,19563253	0,00803267	-24,355	<0,0000000000000002 ***
NoiseExposure_2	-0,02432109	0,00303219	-8,021	0,0000000000000109 ***
NoiseExposure_3	-0,05743664	0,00340711	-16,858	<0,0000000000000002 ***
Slope_2	0,0213533	0,00300983	7,095	0,00000000000132614 ***
Slope_3	0,02340586	0,00338482	6,915	0,00000000000477233 ***
Exposure_2	0,00807342	0,00274147	2,945	0,00323 **
LakeView_2	0,0294696	0,0031221	9,439	<0,0000000000000002 ***
LakeView_3	0,14125793	0,00365238	38,676	<0,0000000000000002 ***
MountainView_2	0,03589583	0,00339559	10,571	<0,0000000000000002 ***
MountainView_3	0,03912677	0,00372321	10,509	<0,0000000000000002 ***
DistanceToLakes_2	-0,08781904	0,01024465	-8,572	<0,0000000000000002 ***
DistanceToRivers_2	0,03529945	0,00733159	4,815	0,00000148119852766 ***
DistanceToHighVoltagePowerLines_2	0,02620769	0,01101495	2,379	0,01735 *
Year_2018	0,02798861	0,00301712	9,277	<0,0000000000000002 ***
Year_2019	0,06654407	0,00305893	21,754	<0,0000000000000002 ***
Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1				
Residual standard error: 0,2155 on 30594 degrees of freedom				
Multiple R-squared: 0,813				
Adjusted R-squared: 0,8125				
F-statistic: 1705 on 78 and 30594 DF, p-value: < 0,00000000000000022				
Mean Absolute Error: 0,1673825				

5.3 Hedonisches Modell für Eigentumswohnungen

Das hedonische Modell für die Eigentumswohnungen umfasst insgesamt 74 Variablen. Von den flächendeckend zur Verfügung stehenden Variablen wurden ausser der Variable DistanceToRivers sämtliche anderen Variablen verwendet. Diese Variable zeigte sich in sämtlichen Modellvarianten als wenig signifikant. Eine mögliche Interpretation besteht darin, dass Eigentumswohnungen weniger stark von allfälligen Hochwasserschäden betroffen sind als Einfamilienhäuser, bei denen die Nähe zu Fließgewässern im Modell einen Preisabschlag zur Folge hat.

Die abhängige Variable (Preis) sowie die lineare Variable NetLivingArea fließen in logarithmierter Form in das Modell ein. Die Anzahl Zimmer werden kategorisiert (1 bis 3 Zimmer, 4 Zimmer, 5 Zimmer, 6 und mehr Zimmer). Dasselbe gilt für die Anzahl Nasszellen (1 Nasszelle, 2 Nasszellen, 3 Nasszellen, 4 Nasszellen oder mehr). Die beiden Variablen SecondAppartementQuota und PrimaryOrSecondaryHome werden zu vier Kategorien aggregiert

(Erstwohnungen in Gemeinden mit mehr als 20% Zweitwohnungen, Erstwohnungen in Gemeinden mit weniger als 20% Zweitwohnungen, Zweitwohnungen in Gemeinden mit mehr als 20% Zweitwohnungen, Zweitwohnungen in Gemeinden mit weniger als 20% Zweitwohnungen). Die restlichen Variablen fließen so in das Modell ein, wie sie erhoben werden. Andere Transformationen haben sich nicht als zielführend herausgestellt. Dasselbe gilt für die Verwendung von Interaktionsvariablen.

Wie bei dem Modell für Einfamilienhäuser werden die Time-Dummy-Variable für das Erhebungsjahr sowie die Variable PrimaryOrSecondaryHome_Quota nur für die Modellherstellung verwendet, nicht aber bei der Indexberechnung.

Variable	Estimate	Std.Error	t-value	Pr(> t)
Intercept	9,20087438	0,02841157	323,842	<0,0000000000000002 ***
Ln_NetLivingArea	0,86263109	0,00622994	138,465	<0,0000000000000002 ***
PrimaryorSecondaryHome_Quota_2	-0,076381	0,00546883	-13,967	<0,0000000000000002 ***
PrimaryorSecondaryHome_Quota_3	0,31668503	0,00680817	46,515	<0,0000000000000002 ***
PrimaryorSecondaryHome_Quota_4	-0,1073321	0,00791314	-13,564	<0,0000000000000002 ***
YearOfConstruction_2	0,00824964	0,01217776	0,677	0,498134
YearOfConstruction_3	-0,0684441	0,00834722	-8,2	0,00000000000000248 ***
YearOfConstruction_4	-0,0590597	0,00761163	-7,759	0,000000000000008755 ***
YearOfConstruction_5	0,02761908	0,00777509	3,552	0,000382 ***
YearOfConstruction_6	0,13275662	0,00799518	16,605	<0,0000000000000002 ***
YearOfConstruction_7	0,15184701	0,00830411	18,286	<0,0000000000000002 ***
NumberOfRooms_2	0,08948845	0,00851304	10,512	<0,0000000000000002 ***
NumberOfRooms_3	0,0901545	0,00916526	9,837	<0,0000000000000002 ***
NumberOfRooms_4	0,09172111	0,01008831	9,092	<0,0000000000000002 ***
NumberOfRooms_5	0,11025005	0,01112246	9,912	<0,0000000000000002 ***
NumberOfRooms_6	0,11722841	0,01305282	8,981	<0,0000000000000002 ***
NumberOfBathrooms_2	0,0644601	0,00285163	22,605	<0,0000000000000002 ***
NumberOfBathrooms_3	0,1711669	0,00729767	23,455	<0,0000000000000002 ***
NumberOfBathrooms_4	0,47160824	0,03520233	13,397	<0,0000000000000002 ***
ConstructionQuality	0,00235584	0,00005518	42,697	<0,0000000000000002 ***
PropertyCondition	0,00099975	0,00006044	16,54	<0,0000000000000002 ***
Canton_1	0,2788336	0,0062737	44,445	<0,0000000000000002 ***
Canton_2	0,21927124	0,0107079	20,478	<0,0000000000000002 ***
Canton_3	0,42065575	0,00979163	42,961	<0,0000000000000002 ***
Canton_4	0,55840587	0,04624944	12,074	<0,0000000000000002 ***
Canton_5	0,23957816	0,00877487	27,303	<0,0000000000000002 ***

Variable	Estimate	Std.Error	t-value	Pr(> t)
Canton_6	0,51938675	0,0198951	26,106	<0,0000000000000002 ***
Canton_7	0,39931335	0,01658133	24,082	<0,0000000000000002 ***
Canton_8	0,19955684	0,0289925	6,883	0,000000000005945232 ***
Canton_9	0,54830761	0,00984245	55,708	<0,0000000000000002 ***
Canton_10	0,17240194	0,01002999	17,189	<0,0000000000000002 ***
Canton_11	0,14017876	0,01231949	11,379	<0,0000000000000002 ***
Canton_12	0,43692846	0,01320424	33,09	<0,0000000000000002 ***
Canton_13	0,28601087	0,01011827	28,267	<0,0000000000000002 ***
Canton_14	0,19081751	0,01636749	11,658	<0,0000000000000002 ***
Canton_15	0,32311504	0,01797504	17,976	<0,0000000000000002 ***
Canton_16	0,38392974	0,11706597	3,28	0,001040 **
Canton_17	0,18617694	0,00844055	22,057	<0,0000000000000002 ***
Canton_18	0,41153021	0,00877191	46,915	<0,0000000000000002 ***
Canton_19	0,20346723	0,00843763	24,114	<0,0000000000000002 ***
Canton_20	0,17912603	0,00936491	19,127	<0,0000000000000002 ***
Canton_21	0,09028801	0,00729469	12,377	<0,0000000000000002 ***
Canton_22	0,46944891	0,0111812	41,986	<0,0000000000000002 ***
Canton_24	0,16970302	0,01575227	10,773	<0,0000000000000002 ***
Canton_25	0,39325776	0,00814853	48,261	<0,0000000000000002 ***
Canton_26	0,02758039	0,02252821	1,224	0,220861
CommunityType_2	-0,1733526	0,00404403	-42,866	<0,0000000000000002 ***
CommunityType_3	-0,2361069	0,00456633	-51,706	<0,0000000000000002 ***
CommunityType_4	-0,1590807	0,00459521	-34,619	<0,0000000000000002 ***
CommunityType_5	-0,2066412	0,00432611	-47,766	<0,0000000000000002 ***
CommunityType_6	-0,2619576	0,00636455	-41,159	<0,0000000000000002 ***
CommunityType_7	-0,219518	0,0063784	-34,416	<0,0000000000000002 ***
CommunityType_8	-0,309327	0,00578645	-53,457	<0,0000000000000002 ***
CommunityType_9	-0,3595389	0,00815003	-44,115	<0,0000000000000002 ***
TaxBurden_2	-0,1428704	0,00633073	-22,568	<0,0000000000000002 ***
TaxBurden_3	-0,2589482	0,00896097	-28,897	<0,0000000000000002 ***
TravelTimeToCenters_2	-0,0745974	0,00285017	-26,173	<0,0000000000000002 ***
TravelTimeToCenters_3	-0,1356675	0,00439335	-30,88	<0,0000000000000002 ***
PublicTransportQuality_2	-0,0507863	0,00403437	-12,588	<0,0000000000000002 ***
PublicTransportQuality_3	-0,0791149	0,00399183	-19,819	<0,0000000000000002 ***
PublicTransportQuality_4	-0,0947679	0,00421151	-22,502	<0,0000000000000002 ***
PublicTransportQuality_5	-0,1210709	0,00508201	-23,823	<0,0000000000000002 ***
NoiseExposure_2	-0,0266268	0,00246713	-10,793	<0,0000000000000002 ***
NoiseExposure_3	-0,0437301	0,0025929	-16,865	<0,0000000000000002 ***
Slope_2	0,05316386	0,00255377	20,818	<0,0000000000000002 ***
Slope_3	0,09854547	0,00311621	31,623	<0,0000000000000002 ***
Exposure_2	0,00610517	0,00224429	2,72	0,006525 **

Variable	Estimate	Std.Error	t-value	Pr(> t)
LakeView_2	0,04104882	0,00272956	15,039	<0,0000000000000002 ***
LakeView_3	0,15505632	0,00296262	52,338	<0,0000000000000002 ***
MountainView_2	0,0330689	0,00291352	11,35	<0,0000000000000002 ***
MountainView_3	0,03826563	0,00332001	11,526	<0,0000000000000002 ***
DistanceToLakes_2	-0,0680977	0,00687699	-9,902	<0,0000000000000002 ***
DistanceToHighVoltagePowerLines_2	0,01737588	0,01070091	1,624	0,104431
Year_2018	0,0288674	0,00254526	11,342	<0,0000000000000002 ***
Year_2019	0,05375467	0,00253236	21,227	<0,0000000000000002 ***

Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1
 Residual standard error: 0,2155 on 30594 degrees of freedom
 Multiple R-squared: 0,813
 Adjusted R-squared: 0,8125
 F-statistic: 1705 on 78 and 30594 DF, p-value: < 0,0000000000000002
 Mean Absolute Error: 0,1673825

5.4 Allgemeine Modellanalyse

Bereits in Rahmen der Modellauswahl wurden die Modelle ein erstes Mal plausibilisiert. Dabei wird jeweils überprüft, ob die Koeffizienten der Variablen signifikant sind und den zu erwartenden ökonomischen Effekten entsprechen. Sowohl bei den Struktur-, Nutzungs- als auch bei den Mikrolagevariablen entsprechen Koeffizienten in beiden Modelle ausnahmslos den theoretisch zu erwartenden Effekten. Auch bei den Makrolagevariablen ist dies grundsätzlich der Fall. Bei den beiden Variablen Canton und CommunityType, die Gebietskörperschaften widerspiegeln, ist die Beurteilung allerdings ein bisschen schwieriger. Dies lässt sich am Beispiel der Kantone festmachen. In beiden Modellen wurde jeweils der Kanton mit dem tiefsten Preisniveau als Referenzkategorie verwendet. Dies hat zur Folge, dass die Koeffizienten sämtlicher anderen Kantone ein positives Vorzeichen erhalten. Bei den Einfamilienhäusern bildet der Kanton Jura die Referenzkategorie, bei den Eigentumswohnungen ist es der Kanton Wallis. Dass diese beiden Kantone ein tieferes Preisniveau aufweisen als beispielsweise die Kantone Zürich oder Genf, liegt auf der Hand. Da aber keine genaue Reihenfolge der theoretisch zu erwartenden kantonalen Preisniveaus existiert, lassen sich die Koeffizienten für die verschiedenen Kantone nur grob plausibilisieren. Soweit sich aber eine Aussage machen lässt, ergeben sich auch für die Kantone und die Gemeindetypen vernünftige Resultate. Daneben gilt es auch festzuhalten, dass die Signifikanz der einzelnen Koeffizienten gegeben ist. Dies gilt auch für alle anderen Modellvariablen. Als einzige berücksichtigte Variable ist die Distanz zu Hochspannungsleitungen im Modell für Eigentumswohnungen nur schwach signifikant.

Mit Hilfe des Variance-Inflation-Factor-Tests wurde in beiden Modellen eine leichte Multikollinearität für die Variablen Canton und TaxBurden gefunden. Da Steuersätze zwar von Kanton zu Kanton stark variieren, innerhalb der Kantone aber relativ homogen sind, überrascht eine gewisse Korrelation zwischen den Kantonen und der Steuerbelastung wenig. Dies ist allerdings nicht weiter problematisch, da beide Variablen durchaus unterschiedliche Teile der beobachteten Varianz erklären.

Die zwei Modelle weisen mit einem R^2 von 0,813 (Einfamilienhäuser) und 0,8574 (Eigentumswohnungen) einen relativ hohen Erklärungsgehalt auf. Dass das Modell für Einfamilienhäuser ein tieferes R^2 hat als dasjenige für Eigentumswohnungen, ist allenfalls dadurch zu erklären, dass Eigentumswohnungen einen etwas homogenen Objekttyp darstellen. Das Adjusted R^2 ist mit 0,8125 und 0,8571 nur geringfügig tiefer als das R^2 . Der Standardfehler der Regression beträgt 0,2155 für Einfamilienhäuser bzw. 0,2023 für Eigentumswohnungen und der Mean Absolute Error liegt bei 0,1673825 respektive 0,159704.

In beiden hedonischen Modellen wird sowohl die abhängige Variable als auch vereinzelte unabhängige Variablen (VolumeOfBuilding, LandArea und NetLivingArea) logarithmiert. In diesem Zusammenhang spricht man von einem Log-Log-Model. Die Koeffizienten der logarithmierten unabhängigen Variablen können als prozentuale Veränderung des Transaktionspreises bei einer prozentualen Änderung des bestimmten Merkmals interpretiert werden. Der Koeffizient der Variable $\ln_VolumeOfBuilding$ beträgt beispielsweise 0,47398024. Dies bedeutet, dass eine Erhöhung des Gebäudevolumens um ein Prozent eine Erhöhung des Transaktionspreises um 0,47398024% zur Folge hat. Beim Koeffizienten einer nicht logarithmierten unabhängigen Dummyvariablen β beträgt die Veränderung des Transaktionspreises im Vergleich zur Referenzkategorie $100(e^\beta - 1)$ Prozent.

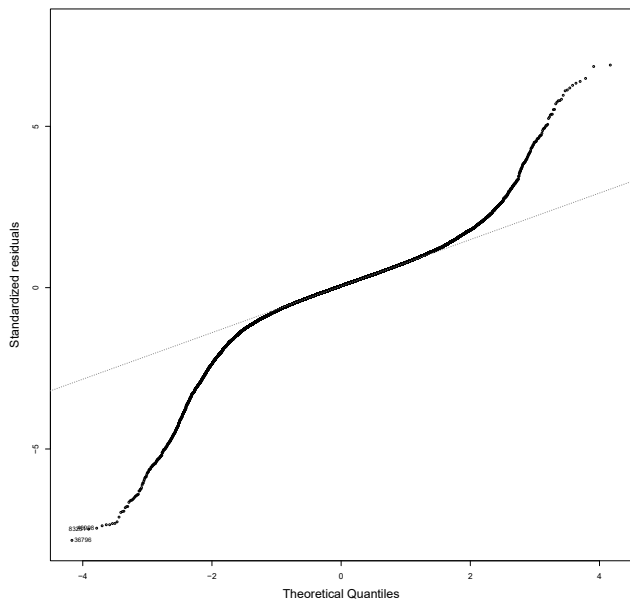
Es gilt an dieser Stelle noch darauf hinzuweisen, dass die beiden hedonischen Modelle für die Bestimmung der aggregierten Qualität der Objekte der verschiedenen Straten erstellt wurden. Auch wenn dabei die Qualität von jedem Objekt in einem ersten Schritt separat bestimmt und erst dann aggregiert wird, liegt der Fokus nicht auf der Marktwertbestimmung von einzelnen Einfamilienhäusern oder Eigentumswohnungen. Aus diesem Grund lehnt das BFS sämtliche dahingehenden Verwendungen der publizierten Modelle ab.

5.5 Analyse der Residuen

Aus den untenstehenden QQ-Plots wird deutlich, dass die Normalverteilung der Residuen im hedonischen Modell für Einfamilienhäuser vor der Ausreisserbehandlung nicht gegeben ist. Die aufgereichte Linie mit den Observationen weicht klar von der diagonalen Gerade ab. Nach der Entfernung der Ausreisser sieht die Verteilung hingegen deutlich besser aus.

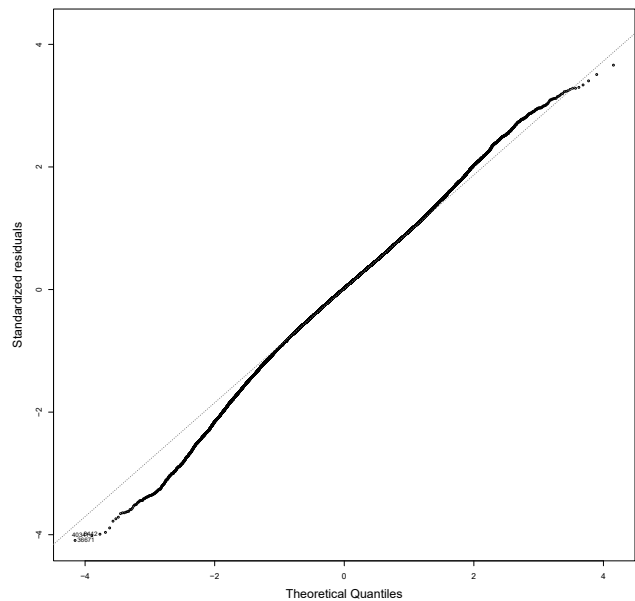
Beim Modell für die Eigentumswohnungen ist die Verbesserung der Verteilung noch offensichtlicher. Die Linie gleicht einer diagonalen Geraden und damit der Normalverteilung an. Einzig an den beiden Enden ist noch eine Abweichung ersichtlich.

Modell EFH: QQ-Plot vor Ausreisserbehandlung G1



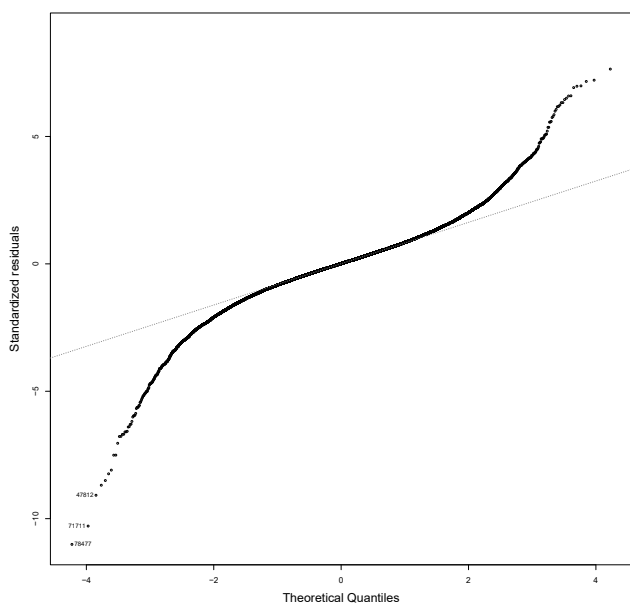
© BFS 2020

Modell EFH: QQ-Plot nach Ausreisserbehandlung G2



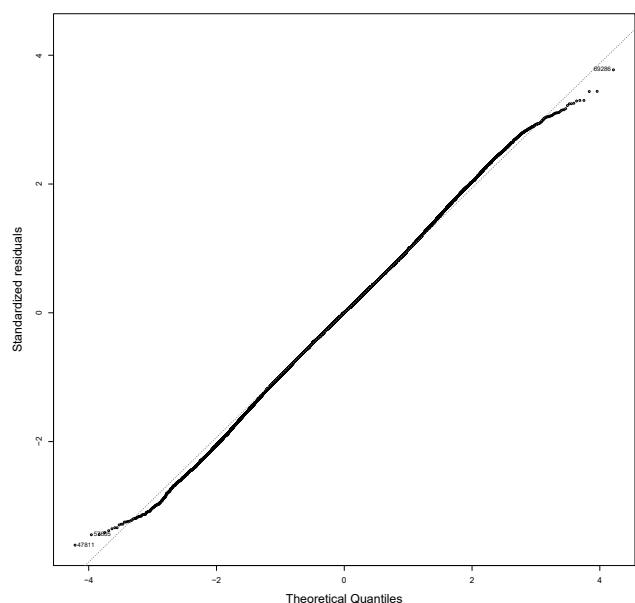
© BFS 2020

Modell EWG: QQ-Plot vor Ausreisserbehandlung G3



© BFS 2020

Modell EWG: QQ-Plot nach Ausreisserbehandlung G4



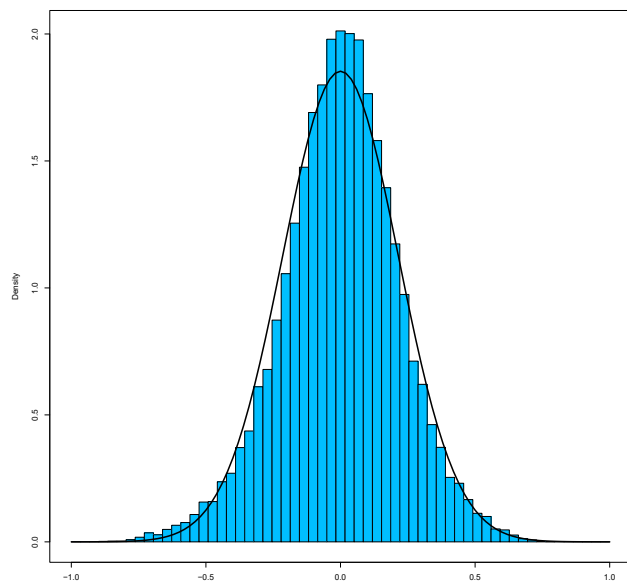
© BFS 2020

Die Annäherung beider Samples an die Normalverteilung wird durch die Verbesserung der Jarque-Berra-Test-Resultate bestätigt. Allerdings muss die Nullhypothese der Normalverteilung weiterhin verworfen werden. In beiden Samples ist eine gewisse Überschusskurtosis vorhanden. Bei den Einfamilienhäusern kommt eine leichte Schiefe hinzu. Auch wenn die beiden Samples

trotz Ausreisserbehandlung und Logarithmierung der abhängigen Variable immer noch nicht perfekt normalverteilt sind, genügen sie den Ansprüchen an eine lineare Regression. Es ist bei Immobilientransaktionsdaten üblich, leichte Abweichungen von der Normalverteilung vorzufinden.

EFH: Histogramm Residuenverteilung

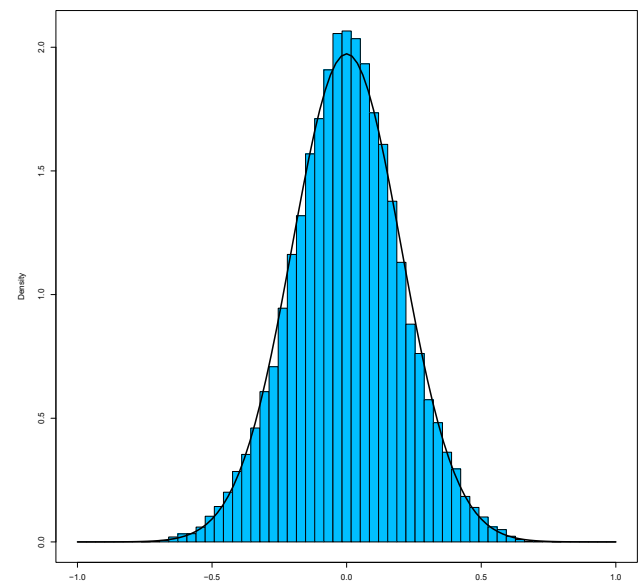
G5



© BFS 2020

EWG: Histogramm Residuenverteilung

G6



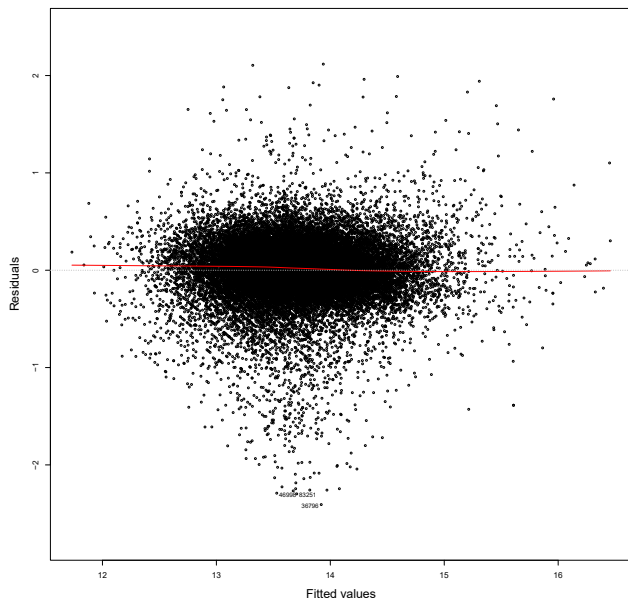
© BFS 2020

Aus den untenstehenden Diagrammen wird ersichtlich, dass die Residuen im ursprünglichen Sample der Einfamilienhäuser allenfalls eine leichte Heteroskedastizität aufweisen. Dies wird durch den Breusch-Pagan-Test bestätigt, bei dem die Nullhypothese der Homoskedastizität der Residuen mit einem p-Wert von 0,053 nur ganz knapp erfüllt ist. Nach der Ausreisserbehandlung ist die Annahme der homoskedastischen Verteilung der Residuen

dann hingegen klar gegeben (p-Wert = 0,754). Gemäss dem White-Test, der nicht von der Normalverteilung der Residuen abhängig ist, kann die Heteroskedastizität bereits im Sample vor der Ausreisserbehandlung deutlich verworfen werden.

**EFH: Residual vs. Fitted Plot
vor Ausreisserbehandlung**

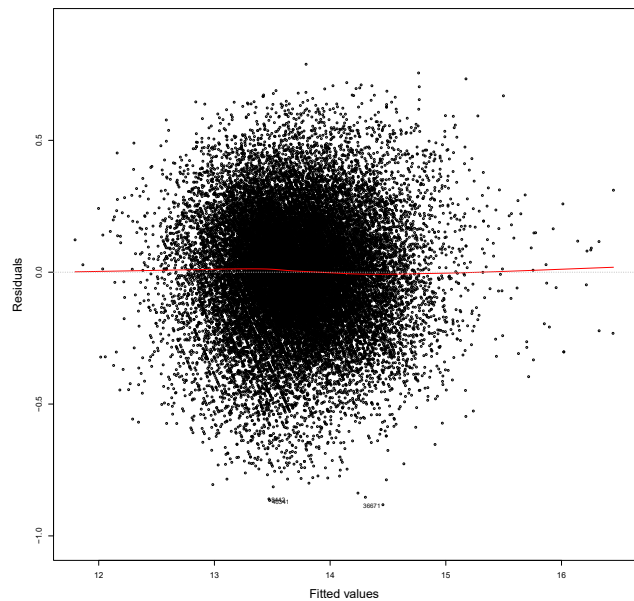
G7



© BFS 2020

**EFH: Residual vs. Fitted Plot
nach Ausreisserbehandlung**

G8



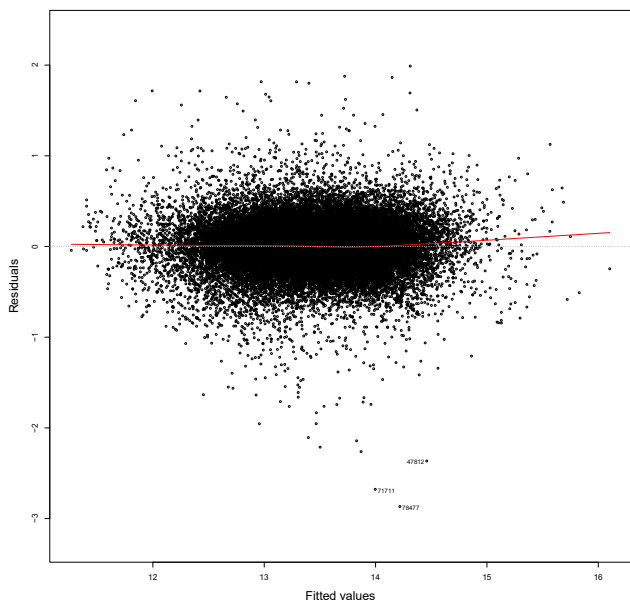
© BFS 2020

Auch bei dem Sample der Eigentumswohnungen sind die Residuen gemäss White-Test bereits vor der Ausreisserbehandlung homoskedastisch verteilt ($p\text{-Wert} = 0,1859$). Im Unterschied zu den Einfamilienhäusern kann laut dem Breusch-Pagan-Test die Heteroskedastizität der Residuen allerdings auch nach der Eliminierung der Ausreisser immer noch knapp nicht verworfen werden.

Auch wenn die Residuen beim Modell für Einfamilienhäuser etwas konstanter verteilt sind als beim Modell für Eigentumswohnungen, kann festgehalten werden, dass nach der Ausreisserbehandlung in beiden Modellen keine problematische Heteroskedastizität mehr erkennbar ist. Neben der Beseitigung der Ausreisser hat sicherlich auch die Logarithmierung der abhängigen Variable bei der Beseitigung von Heteroskedastizität geholfen.

**EWG: Residual vs. Fitted Plot
vor Ausreisserbehandlung**

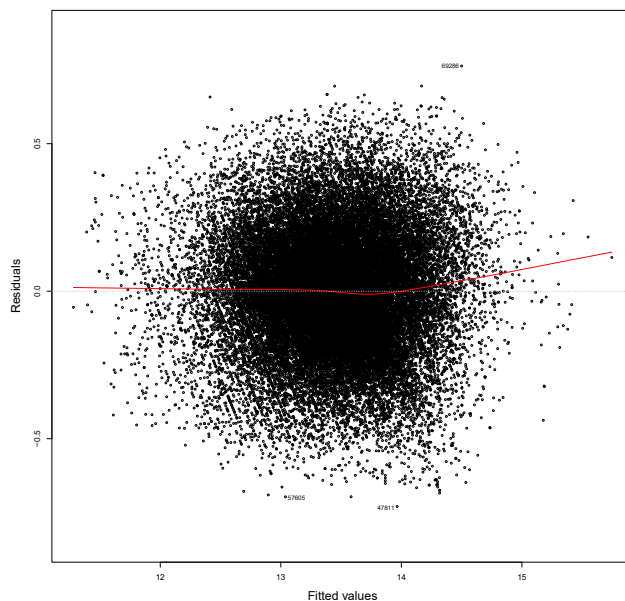
G9



© BFS 2020

**EWG: Residual vs. Fitted Plot
nach Ausreisserbehandlung**

G10



© BFS 2020

5.6 Analyse der Schätzgenauigkeit

In diesem Abschnitt wird die Schätzgenauigkeit der beiden hedonischen Modelle unter die Lupe genommen. Dabei interessiert uns die prozentuale Abweichung der geschätzten Preise von den effektiv beobachteten Transaktionspreisen. Insgesamt ist bei den Eigentumswohnungen eine leicht höhere Schätzgenauigkeit als bei den Einfamilienhäusern zu beobachten. Bei beiden Modellen weisen knapp vier von fünf Objekte eine Abweichung von unter 25 Prozent auf. Bei der genauen Betrachtung der Objekte mit einer grösseren Abweichung als 25% fällt auf, dass darin vermehrt Objekte mit einem sehr hohen Transaktionspreis enthalten sind, die von den hedonischen Modellen tendenziell unterschätzt werden. Zu einem gewissen Teil handelt es sich dabei um Objekte, die bei der Indexberechnung nicht berücksichtigt werden, da dort eine Obergrenze von 10 Millionen CHF angewendet wird.

Auch wenn die gute Schätzqualität pro Objekt erfreulich ist, gilt es festzuhalten, dass die hedonischen Modelle beim Wohnimmobilienpreisindex nicht in erster Linie dafür bestimmt sind, einzelne

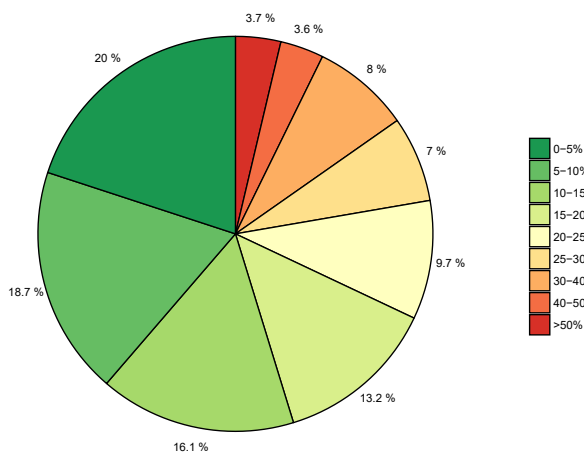
Objekte zu bewerten. Viel wichtiger ist es, dass die aggregierte bzw. mittlere Qualität der gehandelten Objekte adäquat geschätzt werden kann. Über das gesamte Transaktionsvolumen der Einfamilienhäuser weisen die geschätzten Werte 2,3% Abweichung auf. Bei den Eigentumswohnungen sind es 2,2%. In beiden Fällen wird das Transaktionsvolumen leicht unterschätzt. Dies hat wiederum zu einem guten Teil mit den Transaktionen zu tun, die einen sehr hohen Transaktionspreis aufweisen. Bei diesen Objekten scheinen beide Modelle zu tief zu bewerten, was sich dann relativ deutlich in dem Vergleich des tatsächlichen und des geschätzten Transaktionsvolumens niederschlägt.

Neben den Schätzgenauigkeiten der hedonischen Modelle für Einfamilienhäuser und für Eigentumswohnungen wurde auch die Schätzgenauigkeit der Modelle in den jeweiligen Zellen der Stratifizierung untersucht (vgl. Kapitel 6.4). Wenig überraschend ist die Abweichung der Schätzgenauigkeit bei den mittleren Gemeindetypologien am geringsten. Die grössten Abweichungen der Schätzgenauigkeit kann beim Gemeindetyp der städtischen Gemeinden einer grossen Agglomeration beobachtet werden.

EFH: Schätzgenauigkeit gesamtes Sample

%uale Abweichung der Transaktions- von den Schätzpreisen

G11

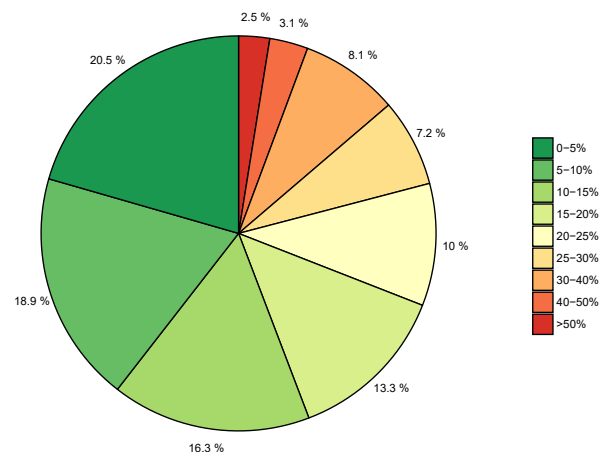


© BFS 2020

EWG: Schätzgenauigkeit gesamtes Sample

%uale Abweichung der Transaktions- von den Schätzpreisen

G12



© BFS 2020

5.7 Fazit

Mit der Kombination einer Stratifizierung mit einem hedonischen Modell des Typs Hedonic Repricing greift das BFS auf ein weit verbreitetes und erprobtes Qualitätsbereinigungsverfahren für die Berechnung eines Immobilienpreisindex zurück. Durch Stratifizierung entlang der zwei Objekttypen (Einfamilienhäuser und Eigentumswohnungen) und der fünf Gemeindetypen (Städtische Gemeinden einer grossen Agglomeration, Städtische Gemeinden einer mittelgrossen Agglomeration, Städtische Gemeinden einer kleinen oder ausserhalb einer Agglomeration, Intermediäre Gemeinden, Ländliche Gemeinden) werden relativ homogene Straten gebildet, die genügend Transaktionen zur Berechnung von Subindizes aufweisen. Die hedonischen Modelle für die beiden Objekttypen weisen einen hohen Erklärungsgehalt auf, erfüllen die zugrunde liegenden Hypothesen an eine lineare Regression und gewährleisten aufgrund der guten Schätzgenauigkeit für sämtliche Stratifizierungszellen die gewünschte Qualitätsbereinigung. Künftig werden die beiden Modelle regelmässig neu kalibriert. Dabei wird nach dem in Kapitel 4 geschilderten Schema vorgegangen. Die nächste Revision ist für den Beginn des Jahres 2021 terminiert. In diesem Zug soll auch getestet werden, ob das verwendete Sample für die Modellierung auf Daten aus zwei Jahren (2019 und 2020) reduziert werden kann. Aufgrund der Tatsache, dass es in den Daten der Jahre 2017 und 2018 noch gewisse Lücken hatte, diese aber seit Anfang 2019 nicht mehr auftreten, könnte eine Eingrenzung des Samples möglich sein.

6 Anhang

6.1 Variablenliste der Ausgangsdaten

Variable	Beschreibung	Nomenklatur
TransactionDate	Transaktionsdatum	
Price	Finaler Kaufpreis	
ObjectType	Objektart	1 = Einfamilienhaus 2 = Eigentumswohnung
SingleFamilyHouseType	Einfamilienhaus-Typ	1 = Freistehend 2 = Zusammengebaut
CondominiumType	Eigentumswohnungs-Typ	1 = Wohnung im obersten Stockwerk 2 = Wohnung in unteren Stockwerken
PrimaryOrSecondaryHome	Erst- bzw. Zweitwohnung	1 = Erstwohnung 2 = Zweitwohnung
OwnerOccupiedOrRented	selbstgenutzt bzw. vermietet	1 = Selbstgenutzt 2 = Vermietet
YearOfConstruction	Baujahr	1 = vor 1919 2 = 1919 bis 1945 3 = 1946 bis 1970 4 = 1971 bis 1990 5 = 1991 bis 2005 6 = 2006 bis 2015 7 = nach 2015
LandArea	Grundstückfläche (bei Einfamilienhäusern)	
VolumeOfBuilding	Kubatur (bei Einfamilienhäusern)	
StandardOfVolume	Erfassungsart Kubatur	1 = GVA 2 = SIA416 3 = SIA116
NetLivingArea	Nettowoohnfläche	
NumberOfRooms	Anzahl Zimmer	
NumberOfBathrooms	Anzahl Nasszellen	
NumberOfParkings	Anzahl Parkplätze	
ConstructionQuality	Ausbaustandard	1 (schlecht) bis 100 (sehr gut)
PropertyCondition	Gebäudezustand	1 (schlecht) bis 100 (sehr gut)

© BFS 2020

Variable	Beschreibung	Nomenklatur
Canton	Kanton	1 = Zürich 2 = Bern 3 = Luzern 4 = Uri 5 = Schwyz 6 = Obwalden 7 = Nidwalden 8 = Glarus 9 = Zug 10 = Fribourg 11 = Solothurn 12 = Basel-Stadt 13 = Basel-Landschaft 14 = Schaffhausen 15 = Appenzell Ausserrhoden 16 = Appenzell Innerrhoden 17 = St. Gallen 18 = Graubünden 19 = Aargau 20 = Thurgau 21 = Ticino 22 = Vaud 23 = Valais 24 = Neuchâtel 25 = Genève 26 = Jura
MajorStatisticalRegion	Grossregionen	1 = Région lémanique 2 = Espace Mittelland 3 = Nordwestschweiz 4 = Zürich 5 = Ostschweiz 6 = Zentralschweiz 7 = Ticino
CommunityType	Gemeindetyp	1 = Städtische Gemeinde einer grossen Agglomeration 2 = Städtische Gemeinde einer mittelgrossen Agglomeration 3 = Städtische Gemeinde einer kleinen oder ausserhalb einer Agglomeration 4 = Periurbane Gemeinde hoher Dichte 5 = Periurbane Gemeinde mittlerer Dichte 6 = Periurbane Gemeinde geringer Dichte 7 = Ländliche Zentrumsgemeinde 8 = Ländliche Gemeinde nahe grosser Agglomerationen 9 = Ländliche periphere Gemeinde
SecondAppartementQuota	Zweitwohnungsquote der Gemeinde (über 20%)	Ja/Nein
TaxBurden	Steuerbelastung pro Gemeinde in %	1 = 0 bis 5,4% 2 = über 5,4 bis 6,8% 3 = über 6,8%
TravelTimeToCenters	Reisezeit zu Zentren in Minuten	1 = 0 bis 11 min 2 = über 11 bis 18 min 3 = über 18 min
PublicTransportQuality	ÖV-Güteklasse	1 = Klasse A 2 = Klasse B 3 = Klasse C 4 = Klasse D 5 = Klasse E
NoiseExposure	Aggregierte Lärmbelastung (in Dezibel)	1 = 0 bis 45 db 2 = über 45 bis 52 db 3 = über 52 db
Slope	Hangneigung	1 = 0 bis 4 Grad 2 = über 4 bis 9 Grad 3 = über 9 Grad
Exposure	Exposition (Himmelsrichtungen)	1 = Nord, Nordwest, Ost, Nordost 2 = West, Südwest, Süd, Südost
LakeView	Sichtbare Seefläche in Hektaren	1 = 0ha 2 = bis 100ha 3 = über 100ha

Variable	Beschreibung	Nomenklatur
MountainView	Anzahl sichtbare Berggipfel	1 = 0 bis 4 Gipfel 2 = 5 bis 13 Gipfel 3 = über 13 Gipfel
DistanceToLakes	Distanz zum nächsten See in Meter	1 = 0 bis 100 m 2 = über 100 m
DistanceToRivers	Distanz zum nächsten Fluss in Meter	1 = 0 bis 100 m 2 = über 100 m
DistanceToHighVoltagePowerLines	Distanz zur nächsten Hochspannungsleitung in Meter	1 = 0 bis 100 m 2 = über 100 m
MatchingType	Qualität des Adressabgleichs	0 = objektgenaues Matching 1 = Matching Strassenmitte 2 = Matching PLZ 3 und 4 = kein Matching möglich

© BFS 2020

6.2 Variablenliste Hedonisches Modell Einfamilienhäuser

Variable	Beschreibung
Ln_VolumeOfBuilding	Natürlicher Logarithmus Gebäudevolumen
StandardOfVolume_1	Dummy Erfassung Gebäudevolumen gemäss GVA
StandardOfVolume_2	Dummy Erfassung Gebäudevolumen gemäss SIA 416
StandardOfVolume_3	Dummy Erfassung Gebäudevolumen gemäss SIA 116
Ln_LandArea	Natürlicher Logarithmus Grundstücksfläche
PrimaryOrSecondaryHome_Quota_1	Dummy Erstwohnungen
PrimaryOrSecondaryHome_Quota_2	Dummy Zweitwohnungen in Gemeinden mit mehr als 20% Zweitwohnungen
PrimaryOrSecondaryHome_Quota_3	Dummy Zweitwohnungen in Gemeinden mit weniger als 20% Zweitwohnungen
YearOfConstruction_1	Dummy Baujahr vor 1919
YearOfConstruction_2	Dummy Baujahr 1919 bis 1945
YearOfConstruction_3	Dummy Baujahr 1946 bis 1970
YearOfConstruction_4	Dummy Baujahr 1971 bis 1990
YearOfConstruction_5	Dummy Baujahr 1991 bis 2005
YearOfConstruction_6	Dummy Baujahr 2005 bis 2015
YearOfConstruction_7	Dummy Baujahr nach 2015
NumberOfRooms_3	Dummy 3 oder weniger Zimmer
NumberOfRooms_4	Dummy 4 Zimmer
NumberOfRooms_5	Dummy 5 Zimmer
NumberOfRooms_6	Dummy 6 Zimmer
NumberOfRooms_7	Dummy 7 Zimmer
NumberOfRooms_8	Dummy 8 oder mehr Zimmer
NumberOfBathrooms_1	Dummy 1 Badezimmer
NumberOfBathrooms_2	Dummy 2 Badezimmer
NumberOfBathrooms_3	Dummy 3 Badezimmer
NumberOfBathrooms_4	Dummy 4 Badezimmer
NumberOfBathrooms_5	Dummy 5 oder mehr Badezimmer
ConstructionQuality	Ausbaustandard
PropertyCondition	Gebäudezustand
Canton_1	Dummy Kanton Zürich
Canton_2	Dummy Kanton Bern
Canton_3	Dummy Kanton Luzern
Canton_4	Dummy Kanton Uri
Canton_5	Dummy Kanton Schwyz
Canton_6	Dummy Kanton Obwalden
Canton_7	Dummy Kanton Nidwalden
Canton_8	Dummy Kanton Glarus
Canton_9	Dummy Kanton Zug
Canton_10	Dummy Kanton Fribourg
Canton_11	Dummy Kanton Solothurn
Canton_12	Dummy Kanton Basel-Stadt
Canton_13	Dummy Kanton Basel-Landschaft
Canton_14	Dummy Kanton Schaffhausen
Canton_15	Dummy Kanton Appenzell Ausserrhoden
Canton_16	Dummy Kanton Appenzell Innerrhoden
Canton_17	Dummy Kanton St. Gallen
Canton_18	Dummy Kanton Graubünden
Canton_19	Dummy Kanton Aargau
Canton_20	Dummy Kanton Thurgau

Variable	Beschreibung
Canton_21	Dummy Kanton Ticino
Canton_22	Dummy Kanton Vaud
Canton_23	Dummy Kanton Valais
Canton_24	Dummy Kanton Neuchâtel
Canton_25	Dummy Kanton Genève
Canton_26	Dummy Kanton Jura
CommunityType_1	Dummy Städtische Gemeinden einer grossen Agglomeration
CommunityType_2	Dummy Städtische Gemeinden einer mittelgrossen Agglomeration
CommunityType_3	Dummy Städtische Gemeinden einer kleinen oder ausserhalb einer Agglomeration
CommunityType_4	Dummy Periurbane Gemeinden hoher Dichte
CommunityType_5	Dummy Periurbane Gemeinden mittlerer Dichte
CommunityType_6	Dummy Periurbane Gemeinden geringer Dichte
CommunityType_7	Dummy Ländliche Zentrums Gemeinden
CommunityType_8	Dummy Ländliche Gemeinden nahe grosser Agglomerationen
CommunityType_9	Dummy Ländliche periphere Gemeinden
TaxBurden_1	Dummy Gemeinden mit tiefer Steuerbelastung (0 bis 5,4%)
TaxBurden_2	Dummy Gemeinden mit mittlerer Steuerbelastung (über 5,4% bis 6,8%)
TaxBurden_3	Dummy Gemeinden mit höherer Steuerbelastung (über 6,8%)
TravelTimeToCenters_1	Dummy Gemeinden mit geringer Reisezeit zu Zentren (0 bis 11min)
TravelTimeToCenters_2	Dummy Gemeinden mit mittlerer Reisezeit zu Zentren (über 11 bis 18min)
TravelTimeToCenters_3	Dummy Gemeinden mit hoher Reisezeit zu Zentren (über 18min)
PublicTransportQuality_1	Dummy ÖV-Gütekategorie A
PublicTransportQuality_2	Dummy ÖV-Gütekategorie B
PublicTransportQuality_3	Dummy ÖV-Gütekategorie C
PublicTransportQuality_4	Dummy ÖV-Gütekategorie D
PublicTransportQuality_5	Dummy ÖV-Gütekategorie E
NoiseExposure_1	Dummy tiefe Lärmbelastung (0 bis 45 Dezibel)
NoiseExposure_2	Dummy mittlere Lärmbelastung (über 45 bis 52 Dezibel)
NoiseExposure_3	Dummy hohe Lärmbelastung (über 52 Dezibel)
Slope_1	Dummy geringe Hangneigung (0 bis 4 Grad)
Slope_2	Dummy mittlere Hangneigung (über 4 bis 9 Grad)
Slope_3	Dummy hohe Hangneigung (über 9 Grad)
Exposure_1	Dummy Exposition gegen Nord, Nordwest, Ost, Nordost
Exposure_2	Dummy Exposition gegen West, Südwest, Süd, Südost
LakeView_1	Dummy keine sichtbare Seefläche (0 Hektaren)
LakeView_2	Dummy wenig sichtbare Seefläche (bis 100 Hektaren)
LakeView_3	Dummy viel sichtbare Seefläche (über 100 Hektaren)
MountainView_1	Dummy 0 bis 4 sichtbare Berggipfel
MountainView_2	Dummy 5 bis 13 sichtbare Berggipfel
MountainView_3	Dummy mehr als 13 sichtbare Berggipfel
DistanceToLakes_1	Dummy 100 Meter Distanz oder weniger zum nächsten See
DistanceToLakes_2	Dummy mehr als 100 Meter Distanz zum nächsten See
DistanceToRivers_1	Dummy 100 Meter Distanz oder weniger zum nächsten Fluss
DistanceToRivers_2	Dummy mehr als 100 Meter Distanz zum nächsten Fluss
DistanceToHighVoltagePowerLines_1	Dummy 100 Meter Distanz oder weniger zur nächsten Hochspannungsleitung
DistanceToHighVoltagePowerLines_2	Dummy mehr als 100 Meter Distanz zur nächsten Hochspannungsleitung
Year_2017	Dummy Transaktionen aus dem Jahr 2017
Year_2018	Dummy Transaktionen aus dem Jahr 2018
Year_2019	Dummy Transaktionen aus dem Jahr 2019

6.3 Variablenliste Hedonisches Modell Eigentumswohnungen

Variable	Beschreibung
Ln_NetLivingArea	Natürlicher Logarithmus Nettowohnfläche
PrimaryOrSecondaryHome_Quota_1	Dummy Erstwohnungen in Gemeinden mit mehr als 20% Zweitwohnungen
PrimaryOrSecondaryHome_Quota_2	Dummy Erstwohnungen in Gemeinden mit weniger als 20% Zweitwohnungen
PrimaryOrSecondaryHome_Quota_3	Dummy Zweitwohnungen in Gemeinden mit mehr als 20% Zweitwohnungen
PrimaryOrSecondaryHome_Quota_4	Dummy Zweitwohnungen in Gemeinden mit weniger als 20% Zweitwohnungen
YearOfConstruction_1	Dummy Baujahr vor 1919
YearOfConstruction_2	Dummy Baujahr 1919 bis 1945
YearOfConstruction_3	Dummy Baujahr 1946 bis 1970
YearOfConstruction_4	Dummy Baujahr 1971 bis 1990
YearOfConstruction_5	Dummy Baujahr 1991 bis 2005
YearOfConstruction_6	Dummy Baujahr 2005 bis 2015
YearOfConstruction_7	Dummy Baujahr nach 2015
NumberOfRooms_1	Dummy 1 Zimmer
NumberOfRooms_2	Dummy 2 Zimmer
NumberOfRooms_3	Dummy 3 Zimmer
NumberOfRooms_4	Dummy 4 Zimmer
NumberOfRooms_5	Dummy 5 Zimmer
NumberOfRooms_6	Dummy 6 oder mehr Zimmer
NumberOfBathrooms_1	Dummy 1 Badezimmer
NumberOfBathrooms_2	Dummy 2 Badezimmer
NumberOfBathrooms_3	Dummy 3 Badezimmer
NumberOfBathrooms_4	Dummy 4 oder mehr Badezimmer
ConstructionQuality	Ausbaustandard
PropertyCondition	Wohnungszustand
Canton_1	Dummy Kanton Zürich
Canton_2	Dummy Kanton Bern
Canton_3	Dummy Kanton Luzern
Canton_4	Dummy Kanton Uri
Canton_5	Dummy Kanton Schwyz
Canton_6	Dummy Kanton Obwalden
Canton_7	Dummy Kanton Nidwalden
Canton_8	Dummy Kanton Glarus
Canton_9	Dummy Kanton Zug
Canton_10	Dummy Kanton Fribourg
Canton_11	Dummy Kanton Solothurn
Canton_12	Dummy Kanton Basel-Stadt
Canton_13	Dummy Kanton Basel-Landschaft
Canton_14	Dummy Kanton Schaffhausen
Canton_15	Dummy Kanton Appenzell Ausserrhoden
Canton_16	Dummy Kanton Appenzell Innerrhoden
Canton_17	Dummy Kanton St. Gallen
Canton_18	Dummy Kanton Graubünden
Canton_19	Dummy Kanton Aargau
Canton_20	Dummy Kanton Thurgau
Canton_21	Dummy Kanton Ticino
Canton_22	Dummy Kanton Vaud
Canton_23	Dummy Kanton Valais
Canton_24	Dummy Kanton Neuchâtel

Variable	Beschreibung
Canton_25	Dummy Kanton Genève
Canton_26	Dummy Kanton Jura
CommunityType_1	Dummy Städtische Gemeinden einer grossen Agglomeration
CommunityType_2	Dummy Städtische Gemeinden einer mittelgrossen Agglomeration
CommunityType_3	Dummy Städtische Gemeinden einer kleinen oder ausserhalb einer Agglomeration
CommunityType_4	Dummy Periurbane Gemeinden hoher Dichte
CommunityType_5	Dummy Periurbane Gemeinden mittlerer Dichte
CommunityType_6	Dummy Periurbane Gemeinden geringer Dichte
CommunityType_7	Dummy Ländliche Zentrumsgemeinden
CommunityType_8	Dummy Ländliche Gemeinden nahe grosser Agglomerationen
CommunityType_9	Dummy Ländliche periphere Gemeinden
TaxBurden_1	Dummy Gemeinden mit tiefer Steuerbelastung (0 bis 5,4%)
TaxBurden_2	Dummy Gemeinden mit mittlerer Steuerbelastung (über 5,4% bis 6,8%)
TaxBurden_3	Dummy Gemeinden mit höherer Steuerbelastung (über 6,8%)
TravelTimeToCenters_1	Dummy Gemeinden mit geringer Reisezeit zu Zentren (0 bis 11min)
TravelTimeToCenters_2	Dummy Gemeinden mit mittlerer Reisezeit zu Zentren (über 11 bis 18min)
TravelTimeToCenters_3	Dummy Gemeinden mit hoher Reisezeit zu Zentren (über 18min)
PublicTransportQuality_1	Dummy ÖV-Gütekategorie A
PublicTransportQuality_2	Dummy ÖV-Gütekategorie B
PublicTransportQuality_3	Dummy ÖV-Gütekategorie C
PublicTransportQuality_4	Dummy ÖV-Gütekategorie D
PublicTransportQuality_5	Dummy ÖV-Gütekategorie E
NoiseExposure_1	Dummy tiefe Lärmbelastung (0 bis 45 Dezibel)
NoiseExposure_2	Dummy mittlere Lärmbelastung (über 45 bis 52 Dezibel)
NoiseExposure_3	Dummy hohe Lärmbelastung (über 52 Dezibel)
Slope_1	Dummy geringe Hangneigung (0 bis 4 Grad)
Slope_2	Dummy mittlere Hangneigung (über 4 bis 9 Grad)
Slope_3	Dummy hohe Hangneigung (über 9 Grad)
Exposure_1	Dummy Exposition gegen Nord, Nordwest, Ost, Nordost
Exposure_2	Dummy Exposition gegen West, Südwest, Süd, Südost
LakeView_1	Dummy keine sichtbare Seefläche (0 Hektaren)
LakeView_2	Dummy wenig sichtbare Seefläche (bis 100 Hektaren)
LakeView_3	Dummy viel sichtbare Seefläche (über 100 Hektaren)
MountainView_1	Dummy 0 bis 4 sichtbare Berggipfel
MountainView_2	Dummy 5 bis 13 sichtbare Berggipfel
MountainView_3	Dummy mehr als 13 sichtbare Berggipfel
DistanceToLakes_1	Dummy 100 Meter Distanz oder weniger zum nächsten See
DistanceToLakes_2	Dummy mehr als 100 Meter Distanz zum nächsten See
DistanceToHighVoltagePowerLines_1	Dummy 100 Meter Distanz oder weniger zur nächsten Hochspannungsleitung
DistanceToHighVoltagePowerLines_2	Dummy mehr als 100 Meter Distanz zur nächsten Hochspannungsleitung
Year_2017	Dummy Transaktionen aus dem Jahr 2017
Year_2018	Dummy Transaktionen aus dem Jahr 2018
Year_2019	Dummy Transaktionen aus dem Jahr 2019

© BFS 2020

6.4 Geoinformationen

Kantone und Grossregionen

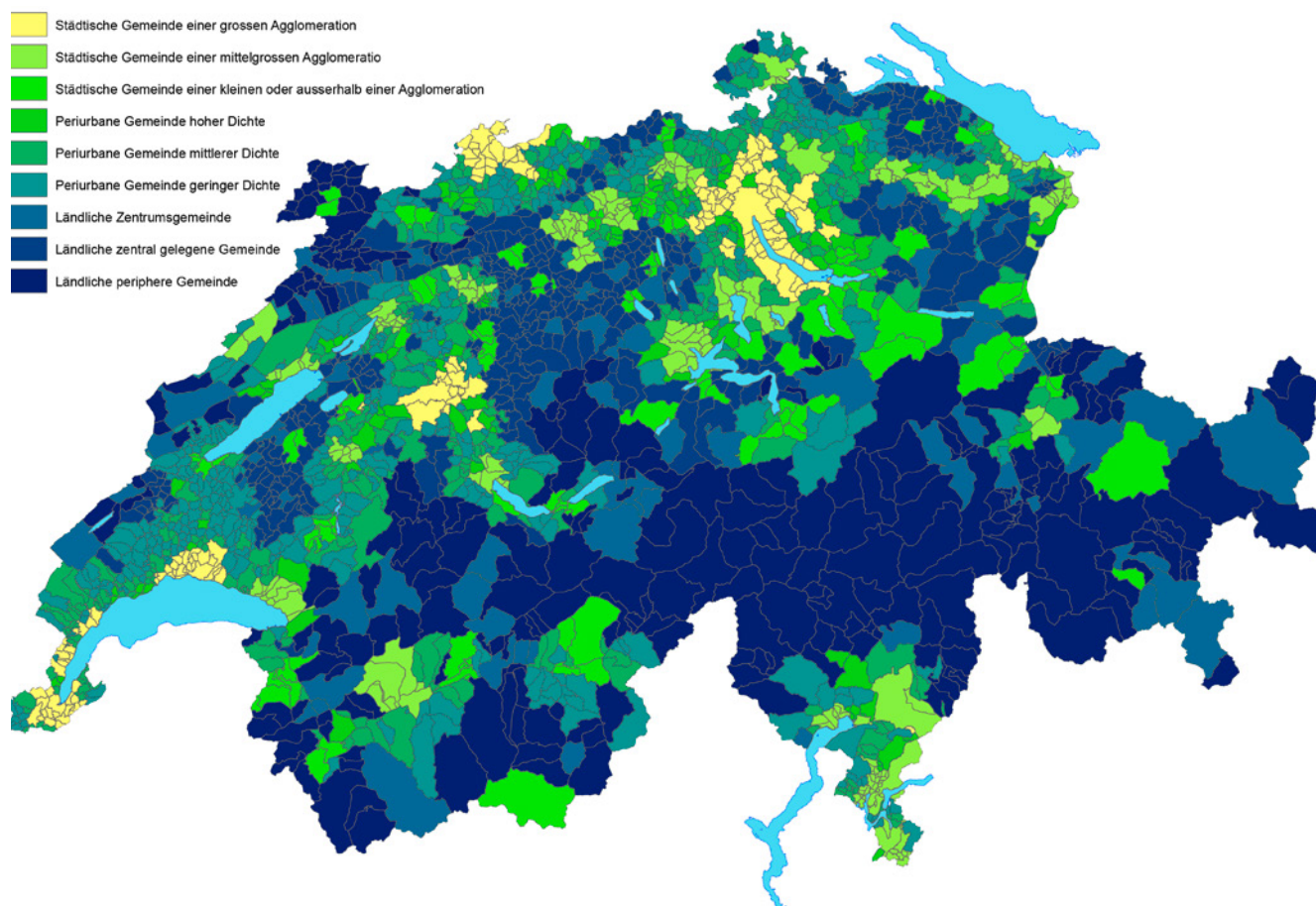
G13



© BFS 2020

Gemeindetypen

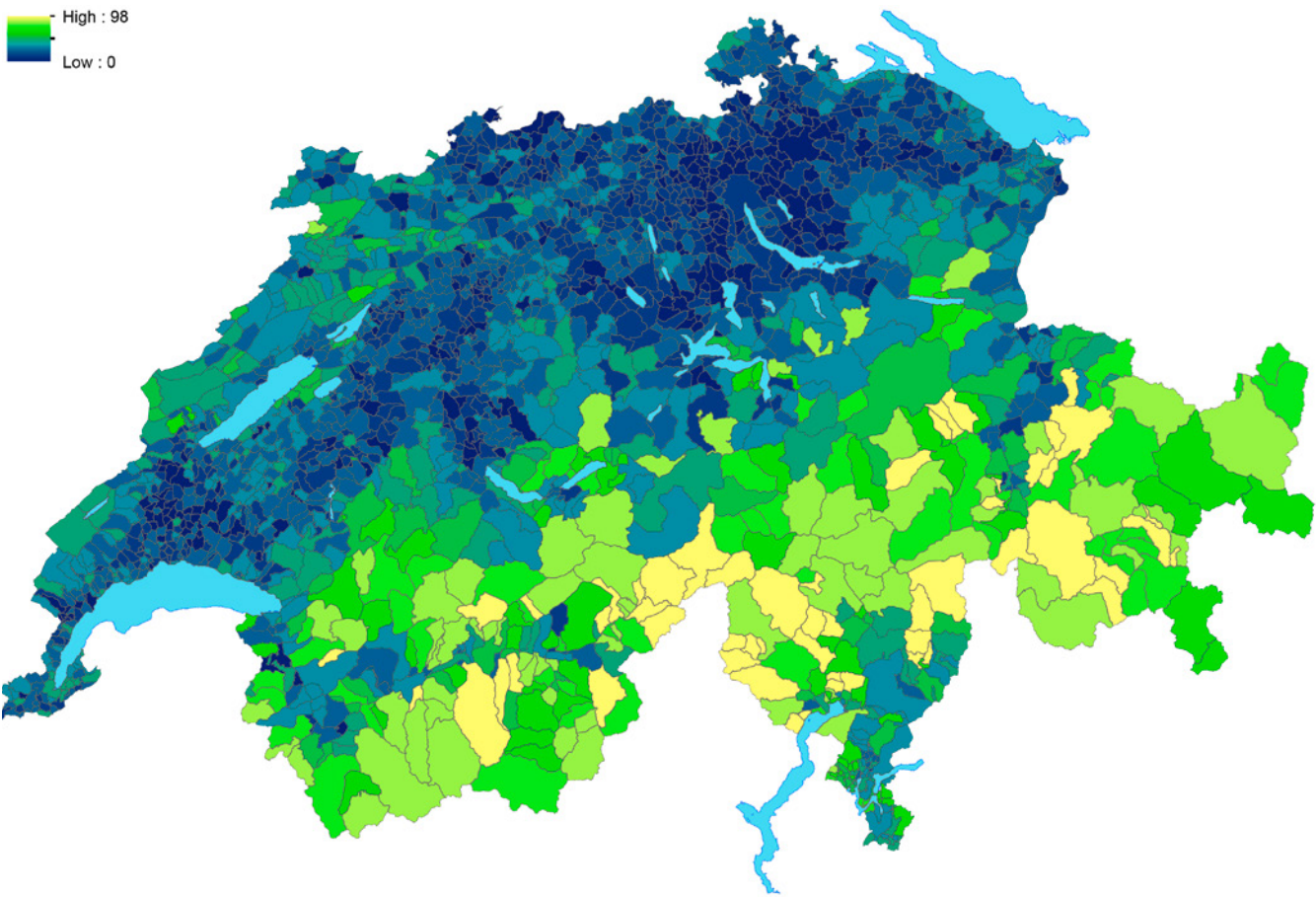
G14



© BFS 2020

Zweitwohnungsquote in Prozent

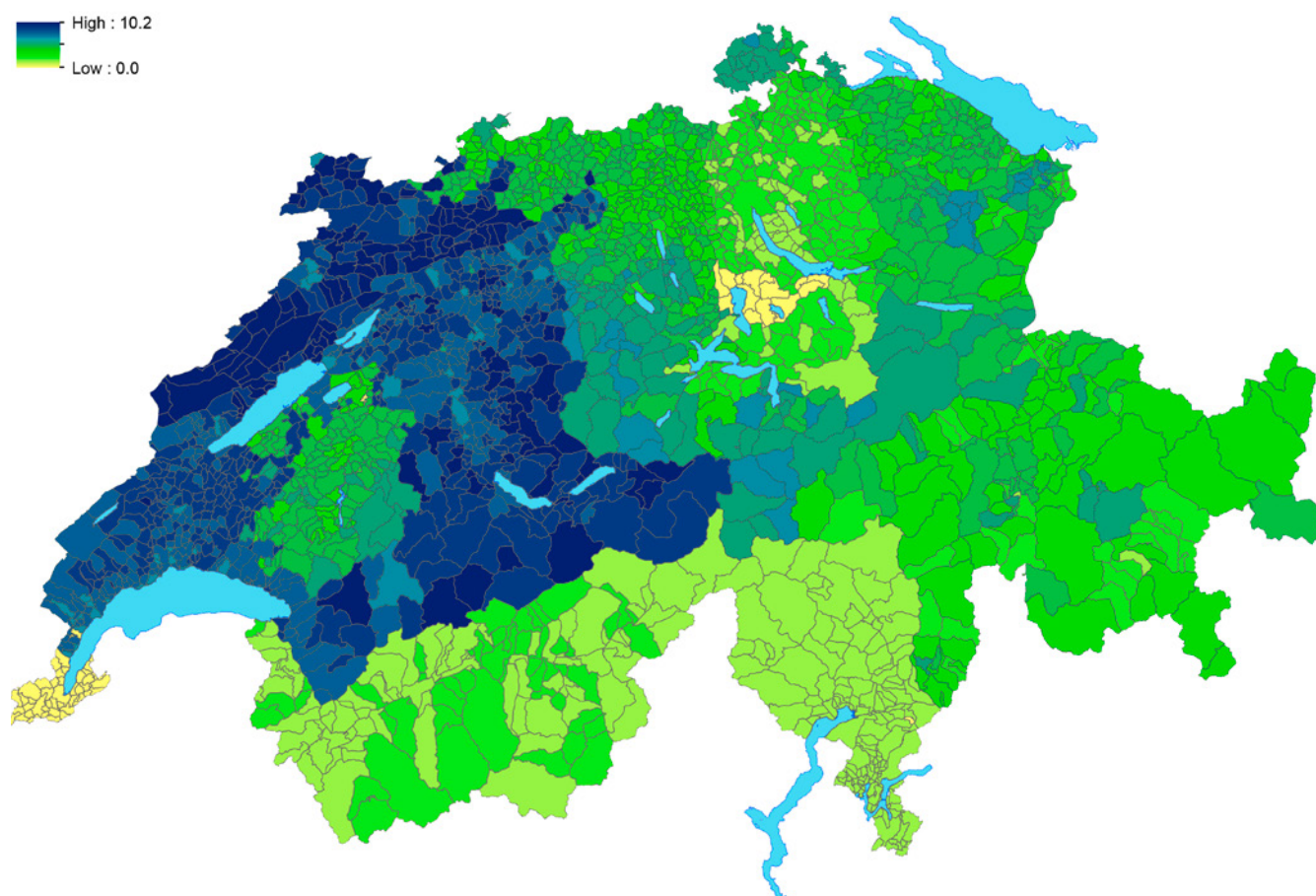
G15



© BFS 2020

Steuerbelastung in Prozent

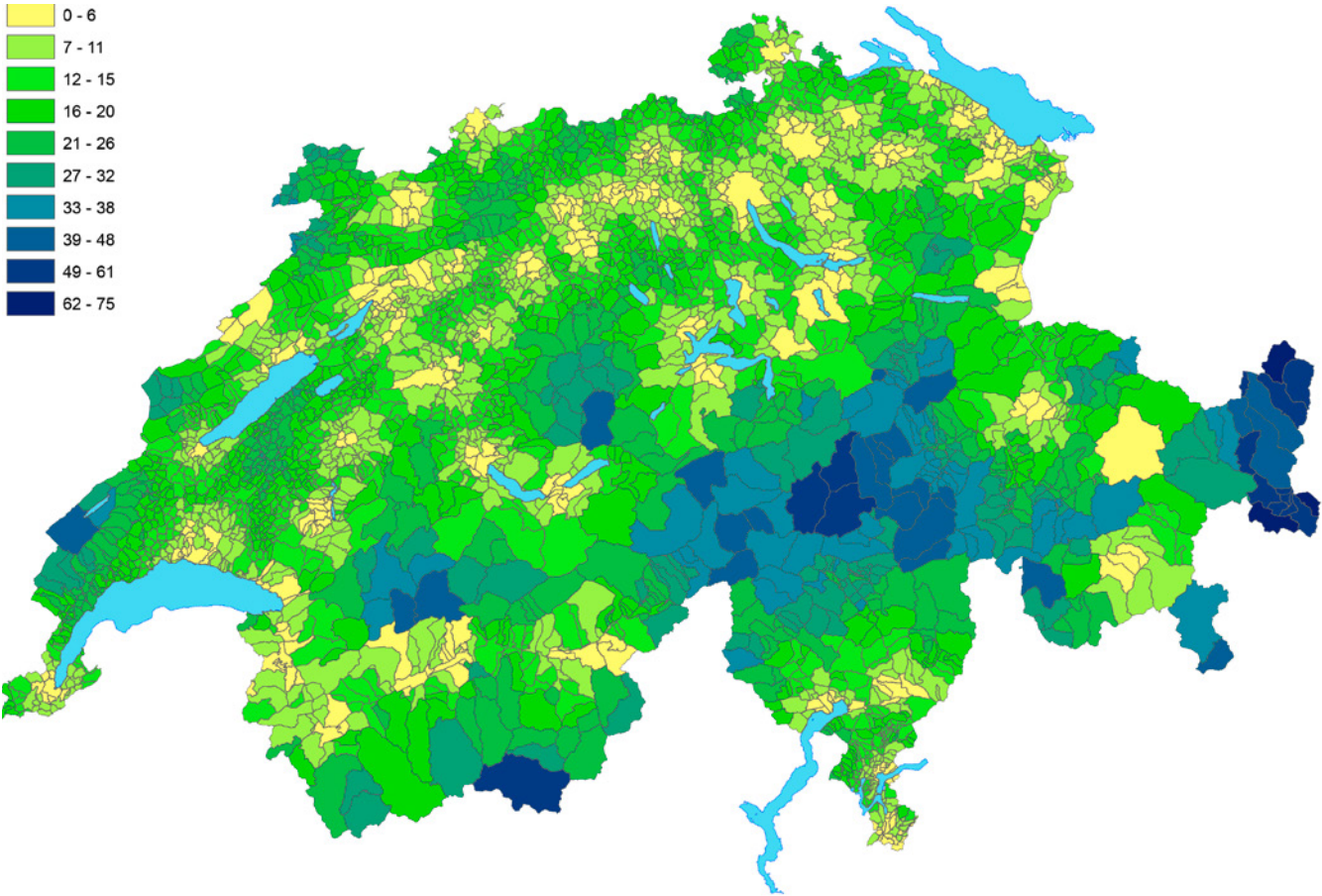
G16



© BFS 2020

Reisezeit MIV zu Agglomerationszentren in Minuten

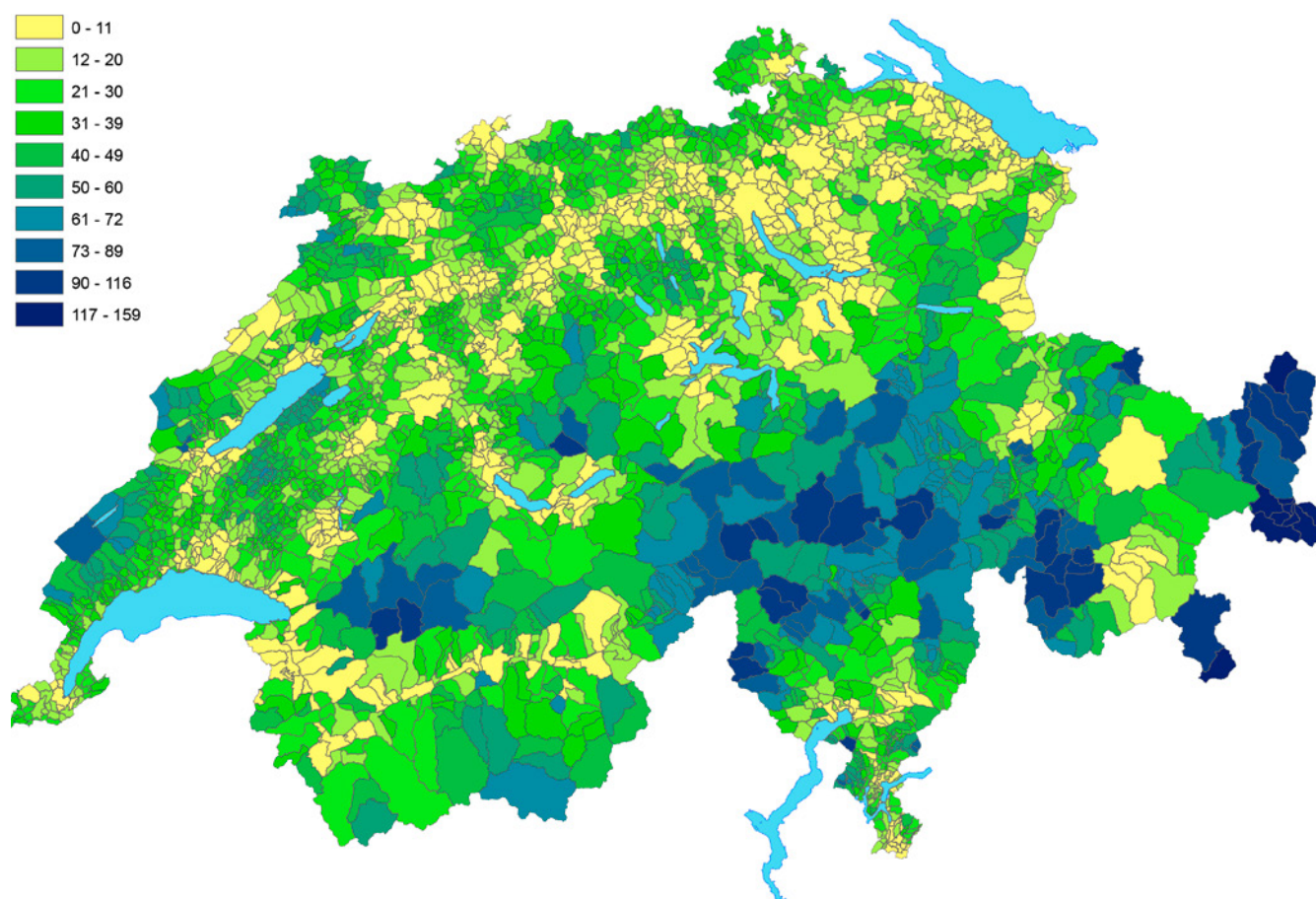
G17



© BFS 2020

Reisezeit ÖV zu Agglomerationszentren in Minuten

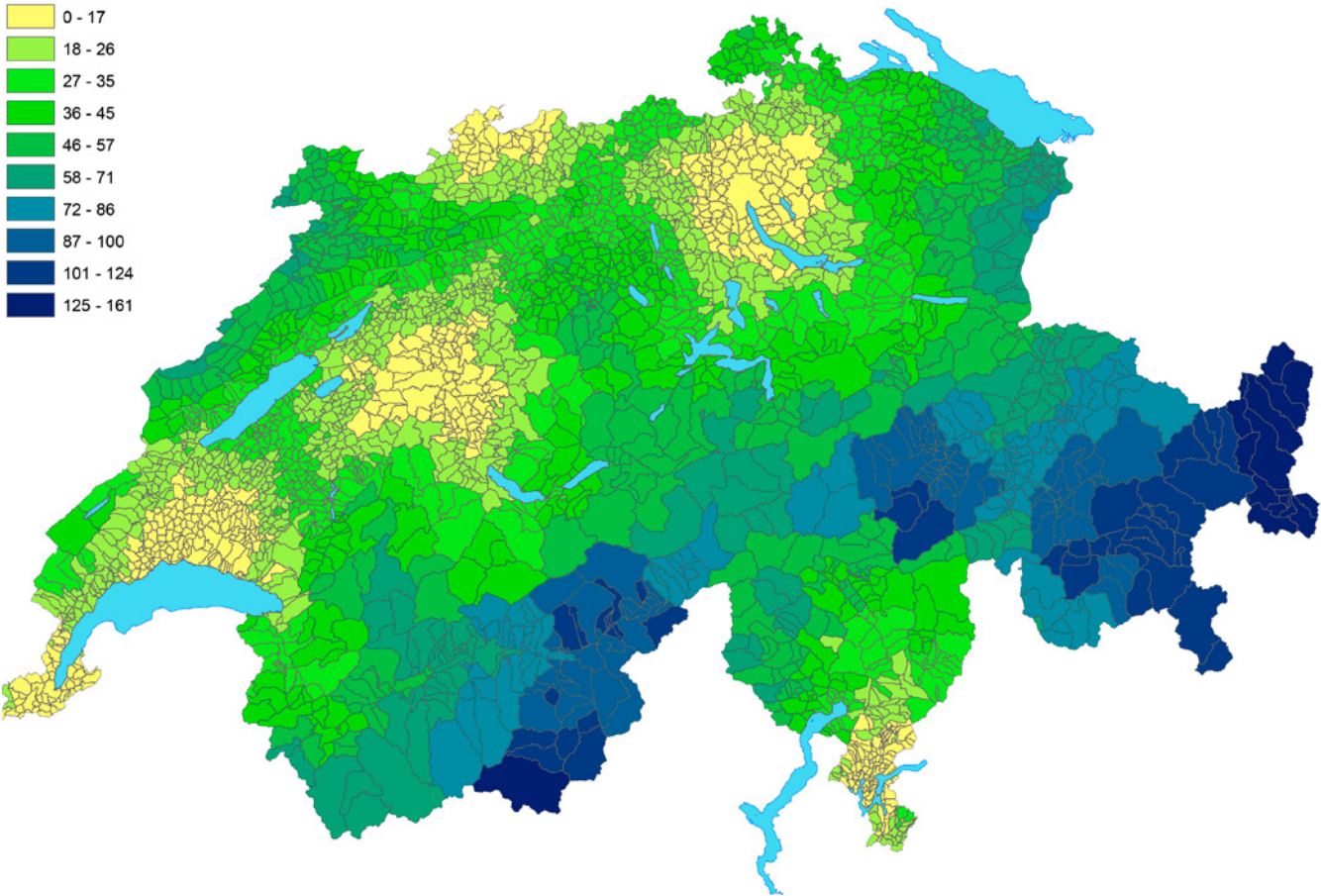
G18



© BFS 2020

Reisezeit MIV zu Kernstädten in Minuten

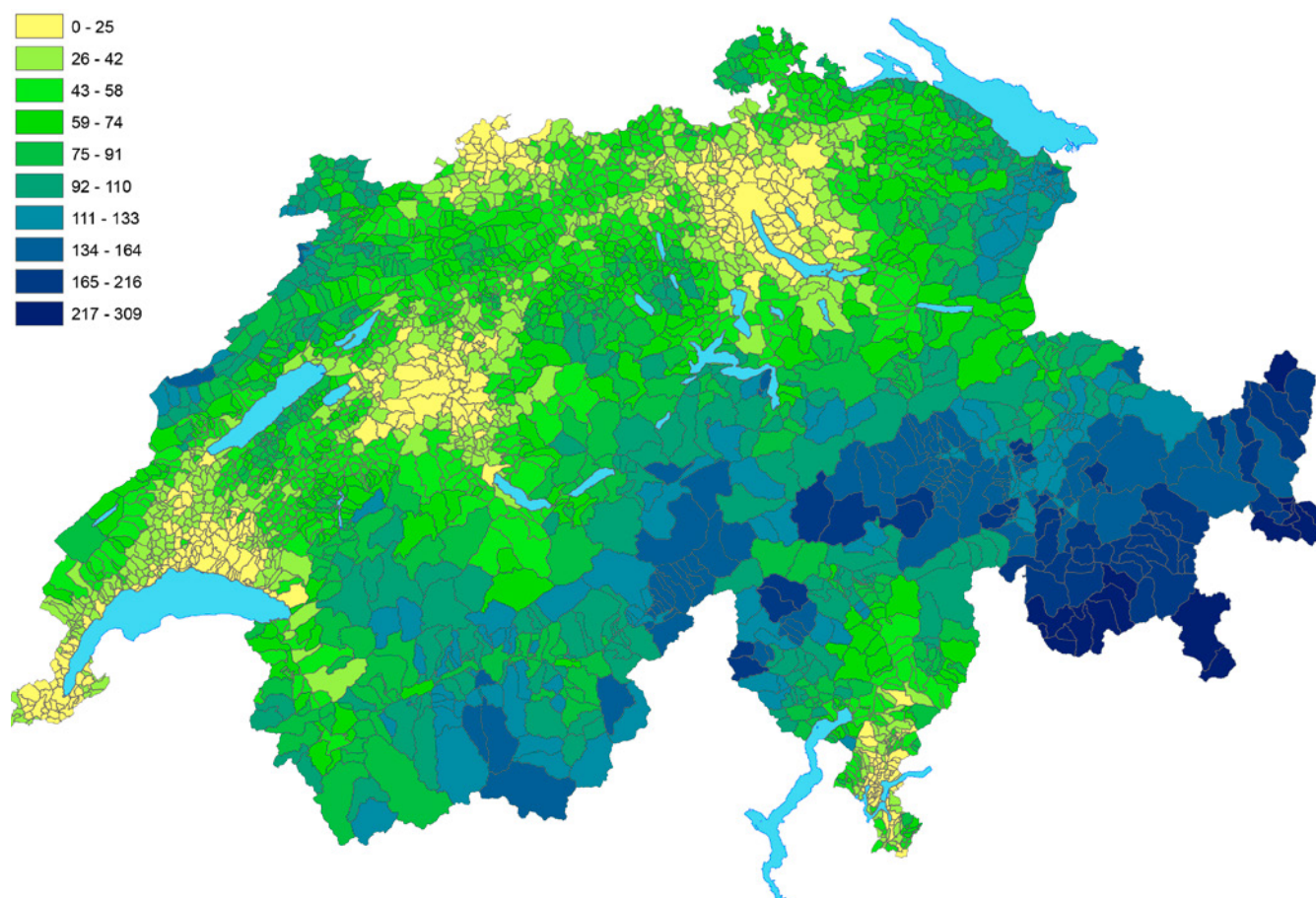
G19



© BFS 2020

Reisezeit ÖV zu Kernstädten in Minuten

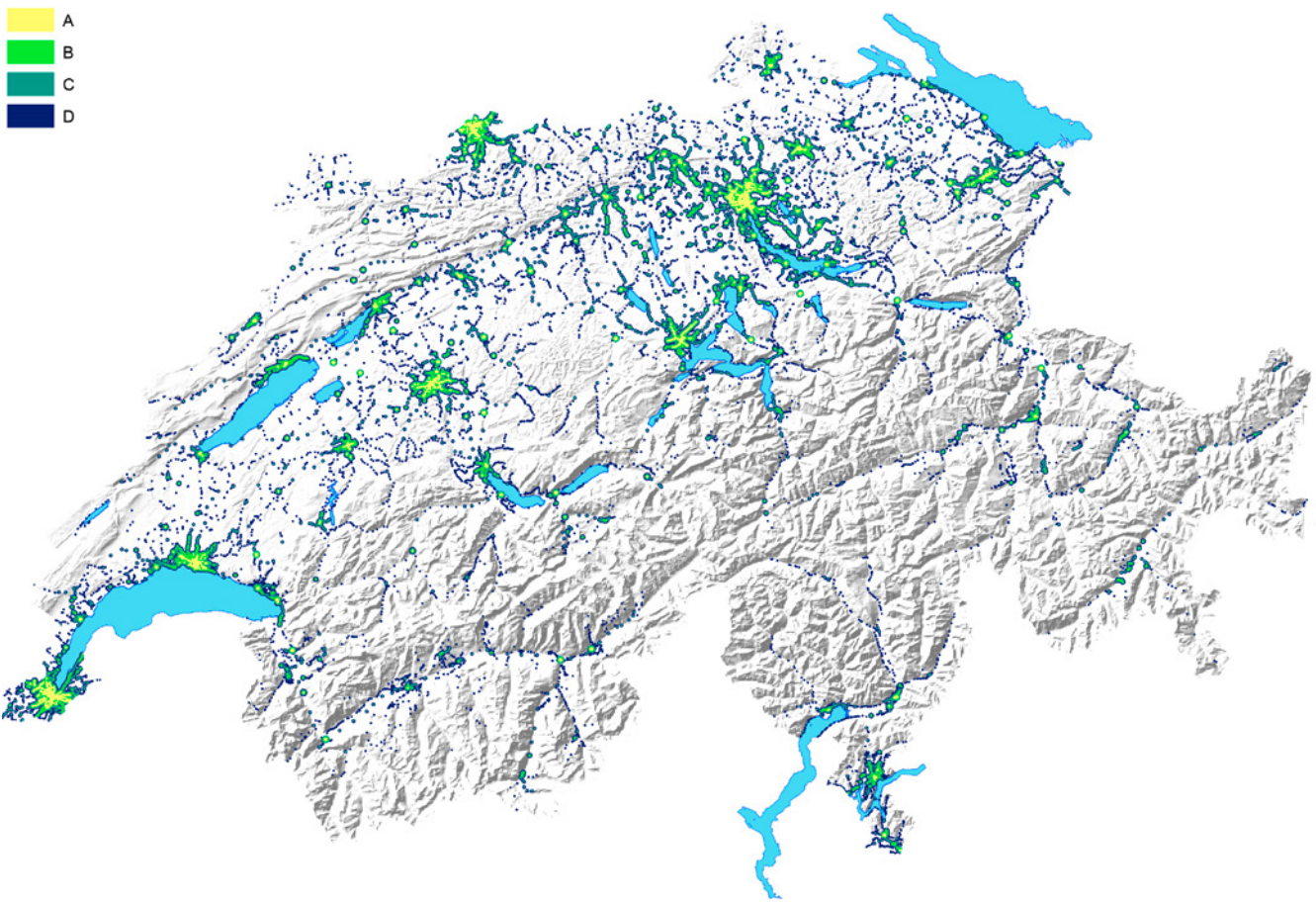
G20



© BFS 2020

ÖV-Güteklassen

G21

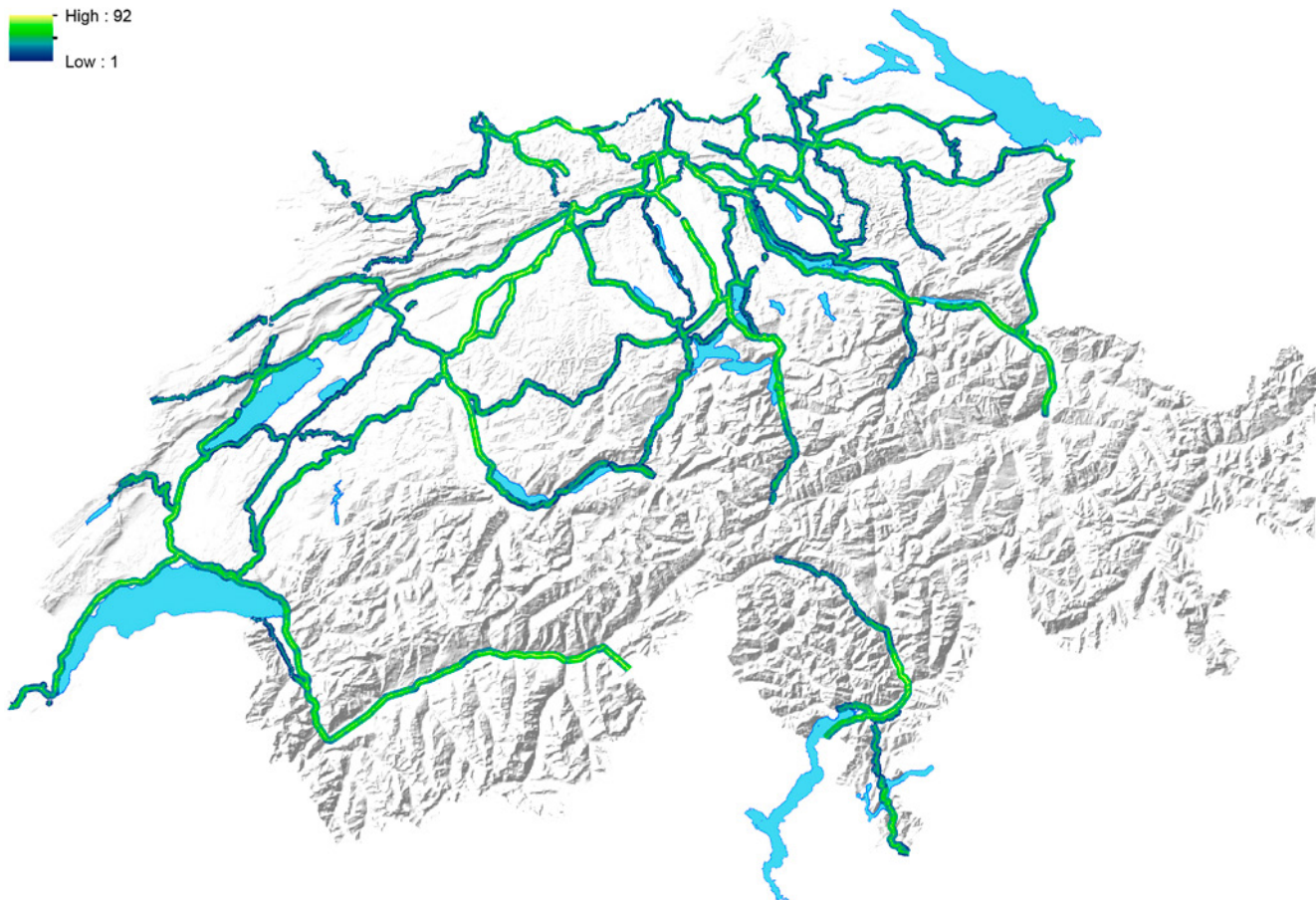


© BFS 2020

Eisenbahnlärm Tag in Dezibel

G22

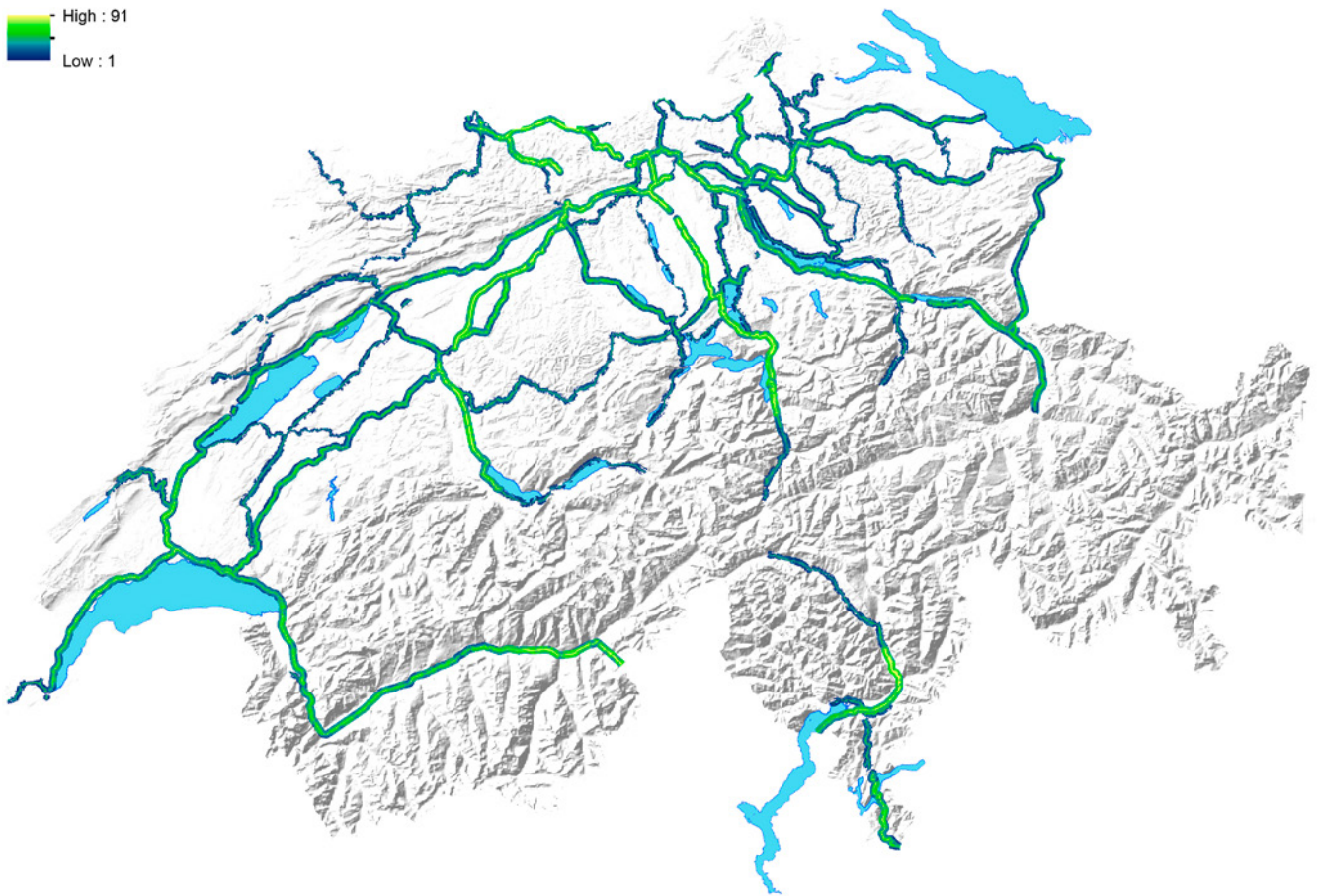
High : 92
Low : 1



© BFS 2020

Eisenbahnlärm Nacht in Dezibel

G23



© BFS 2020

Strassenlärm Tag in Dezibel

G24



© BFS 2020

Strassenlärm Nacht in Dezibel

G25



© BFS 2020

Hangneigung in Grad

G26



© BFS 2020

Exposition in Grad Abweichung von der Nordrichtung

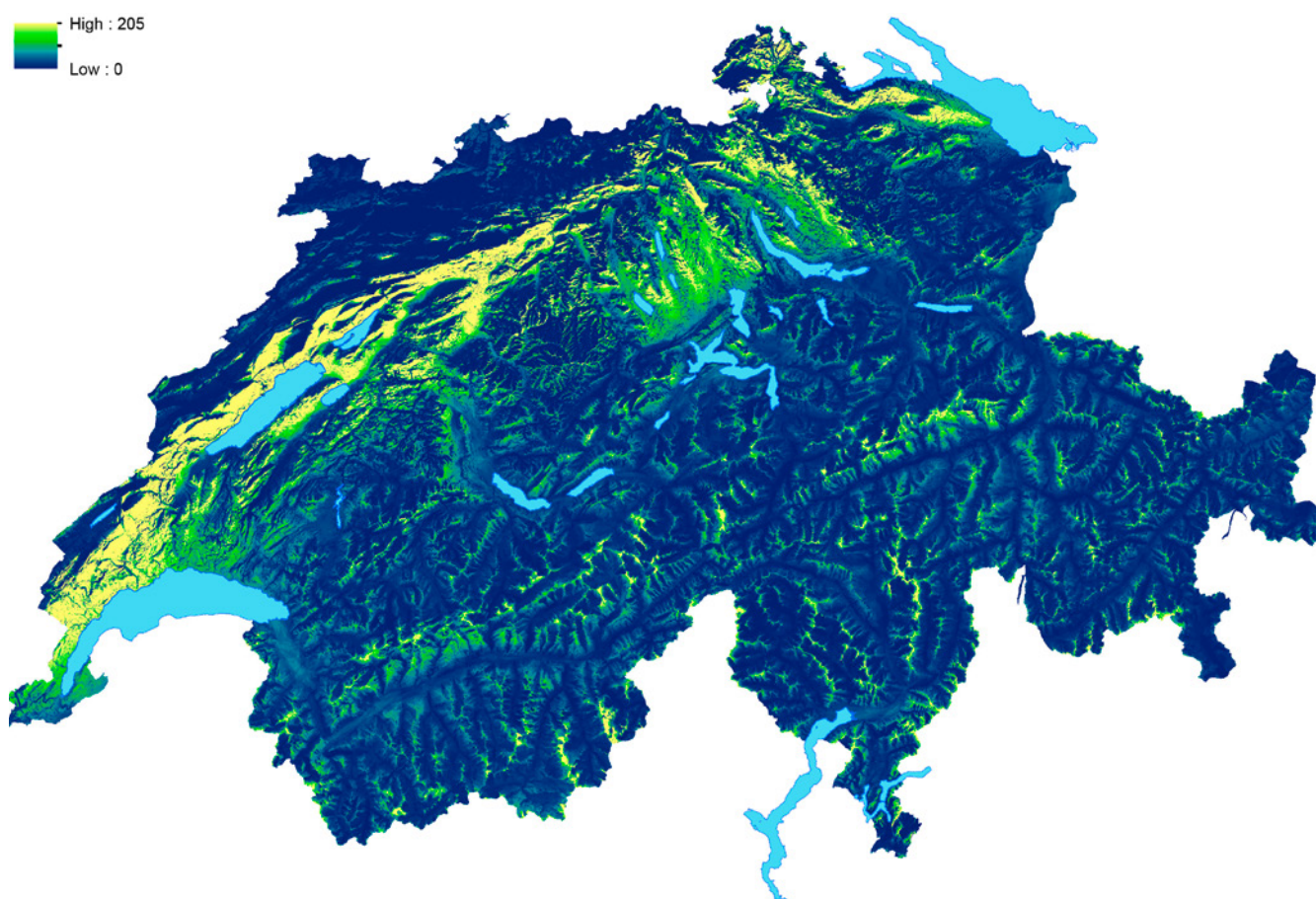
G27



© BFS 2020

Anzahl sichtbare Berggipfel

G28



© BFS 2020



© BFS 2020

Flüsse

G30



© BFS 2020

Hochspannungsleitungen

G31



© BFS 2020

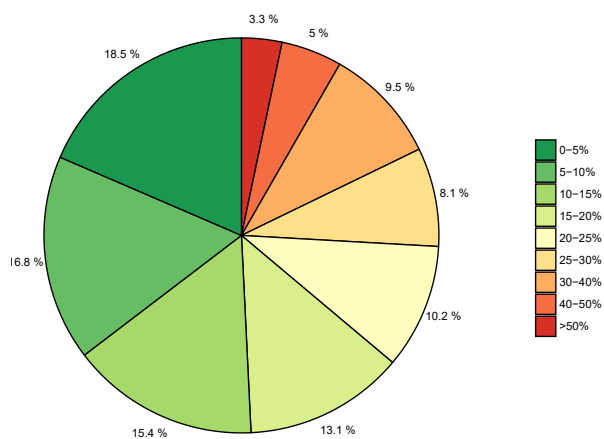
6.5 Schätzgenauigkeit in den Straten

6.5.1 Einfamilienhäuser aufgeteilt nach Gemeindetypen

Städtische Gemeinden einer grossen Agglomeration

%-uale Abweichung der Transaktions- von den Schätzpreisen

G32

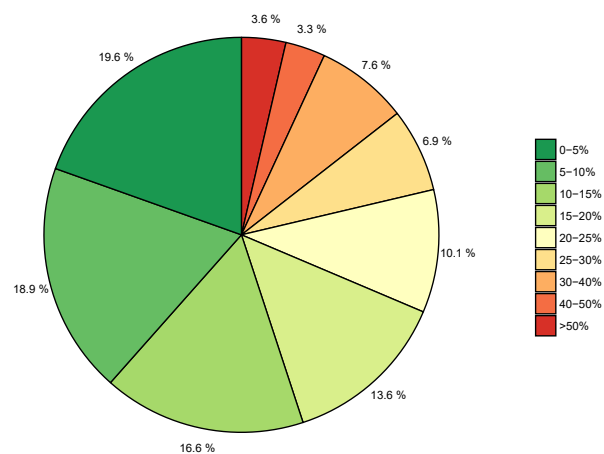


© BFS 2020

Städtische Gemeinden einer mittelgrossen Agglomeration

%-uale Abweichung der Transaktions- von den Schätzpreisen

G33

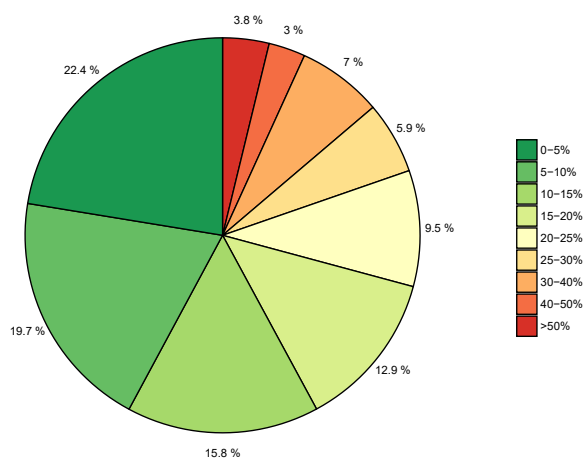


© BFS 2020

Städtische Gemeinden einer kleinen oder ausserhalb einer Agglomeration

%-uale Abweichung der Transaktions- von den Schätzpreisen

G34

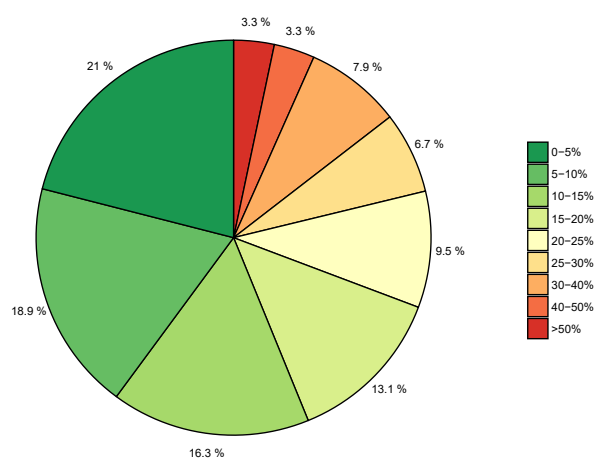


© BFS 2020

Intermediäre Gemeinden

%-uale Abweichung der Transaktions- von den Schätzpreisen

G35

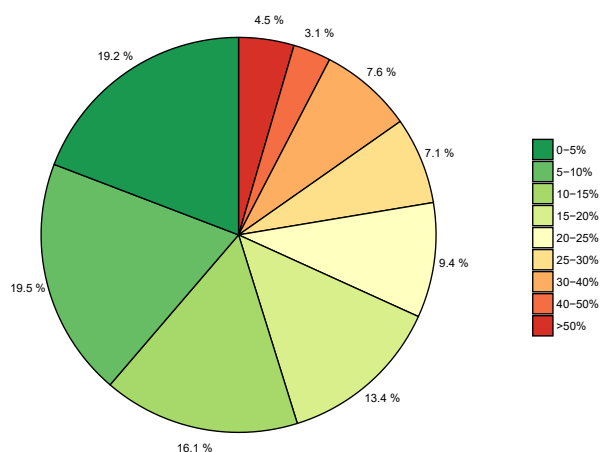


© BFS 2020

Ländliche Gemeinden

%-uale Abweichung der Transaktions- von den Schätzpreisen

G36



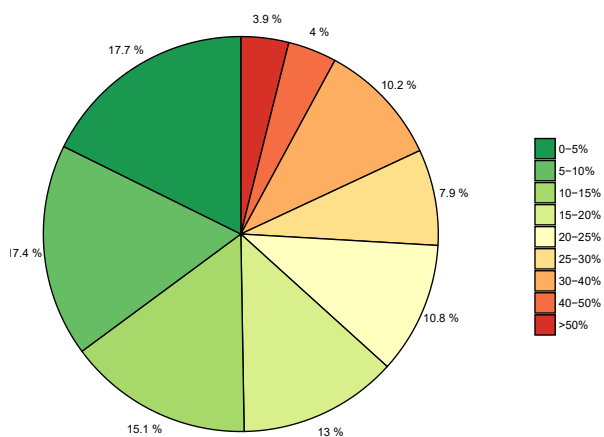
© BFS 2020

6.5.2 Eigentumswohnungen aufgeteilt nach Gemeindetypen

Städtische Gemeinden einer grossen Agglomeration

%-uale Abweichung der Transaktions- von den Schätzpreisen

G37

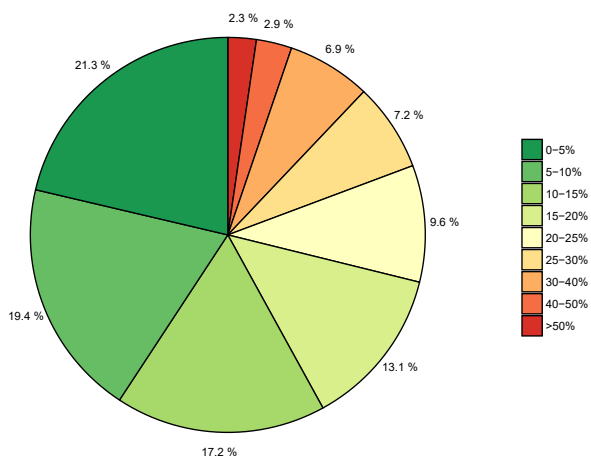


© BFS 2020

Städtische Gemeinden einer mittelgrossen Agglomeration

%-uale Abweichung der Transaktions- von den Schätzpreisen

G38

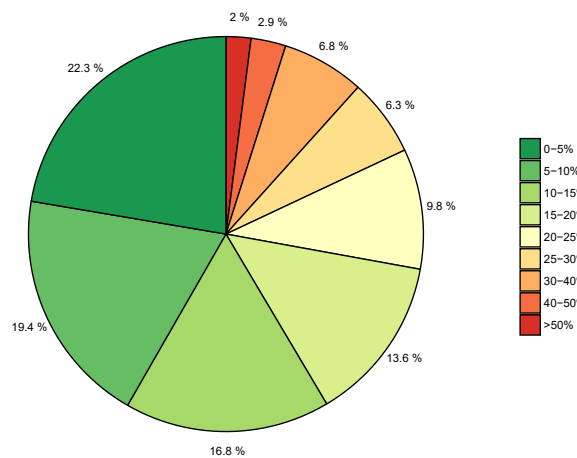


© BFS 2020

Städtische Gemeinden einer kleinen oder ausserhalb einer Agglomeration

%-uale Abweichung der Transaktions- von den Schätzpreisen

G39

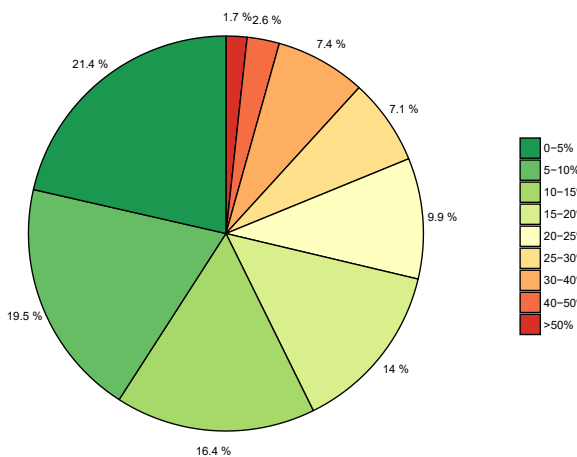


© BFS 2020

Städtische Gemeinden einer kleinen oder ausserhalb einer Agglomeration

%-uale Abweichung der Transaktions- von den Schätzpreisen

G40

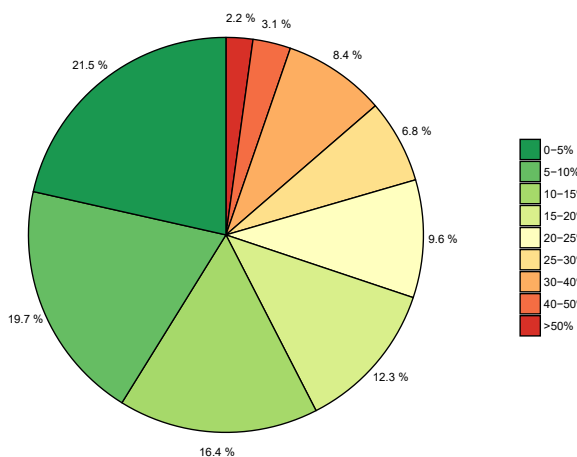


© BFS 2020

Ländliche Gemeinden

%-uale Abweichung der Transaktions- von den Schätzpreisen

G41



© BFS 2020

7 Literaturverzeichnis

ARGE EPFL, Econability and HEG (2012). Machbarkeitsstudie für die statistische Erfassung der Immobilienpreise.

Bera, A. K. and C. M. Jarque (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. In *Economic Letters*. Nr.3, p. 255–259.

Bollen, K. A. and R. W. Jackman (1990). Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases. *Modern methods of data analysis*, p. 257–291.

Breusch, T.S. and A. R. Pagan (1980). A simple test for heteroscedasticity and random coefficient variation. In: *Journal of the Econometric Society*, *Econometrica*, p. 817–838.

Bundesamt für Landestopografie (swisstopo), Digitales Höhenmodell swissALTI3D, 20.10.2020, https://shop.swisstopo.admin.ch/de/products/height_models/alti3D.

Bundesamt für Raumentwicklung, ÖV-Güteklassen, 20.10.2020, <https://www.are.admin.ch/are/de/home/mobilitaet/grundlagen-und-daten/verkehrserschliessung-in-der-schweiz.html>.

Bundesamt für Raumentwicklung, Reisezeiten, 20.10.2020, <https://www.are.admin.ch/are/de/home/mobilitaet/grundlagen-und-daten/verkehrserschliessung-in-der-schweiz.html>.

Bundesamt für Raumentwicklung, Zweitwohnungen, 20.10.2020, <https://www.are.admin.ch/are/de/home/raumentwicklung-und-raumplanung/raumplanungsrecht/zweitwohnun-gen.html>.

Bundesamt für Statistik (2016). Detailkonzept Projekt Immobilienpreisindex.

Bundesamt für Statistik, Erhebungstool Immobilienpreisindex, 20.10.2020, <https://github.com/bfs-preis/impi/wiki/Einführung>.

Bundesamt für Statistik, Gemeindetypologie und Stadt/Land-Typologie 2012, 20.10.2020, <https://www.bfs.admin.ch/bfs/de/home/statistiken/querschnittsthemen/raeumliche-analysen/raeumliche-gliederungen/raeumliche-typologien.assetdetail.2543323.html>.

Bundesamt für Umwelt, Lärmbelastung, 20.10.2020, <https://www.bafu.admin.ch/bafu/de/home/themen/laerm/zustand.html>.

Bundesamt für Zivilluftfahrt, Lärmbelastungskataster, 20.10.2020, <https://www.bazl.admin.ch/bazl/de/home/politik/umwelt/laerm/laermbelastungskataster.html>.

Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics* Vol. 19 (1), p. 15–18.

Eidgenössische Steuerverwaltung, Steuerbelastung in den Gemeinden, 20.10.2020, <https://www.estv.admin.ch/estv/de/home/allgemein/steuerstatistiken/fachinformationen/steuer-belastungen/steuerbelastung.html>.

European Environment Agency, Digital surface model EU-DEM, 20.10.2020, <https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-eu-dem>.

eurostat, Nomenclature of territorial units for statistics (NUTS), 20.10.2020, <https://ec.europa.eu/eurostat/web/nuts/background>.

O'Hanlon, N. (2011). Constructing a National House Price Index for Ireland. *Journal of the Statistical and Social Inquiry Society of Ireland*, Vol XL, Central Statistics Office and Centre for Policy Studies.

Silver, M. (2020). Report to the Swiss Federal Statistical Office (FSO) on two hedonic models: the residential property (house) price index (RPPI) and the rent price index (RPI).

Wooldridge, J. M. (2013), *Introductory Econometrics, a modern approach*, 5th edition, South-Western, chap. 8.3, p. 280.

Publikationsprogramm BFS

Das Bundesamt für Statistik (BFS) hat als zentrale Statistikstelle des Bundes die Aufgabe, statistische Informationen zur Schweiz breiten Benutzerkreisen zur Verfügung zu stellen. Die Verbreitung geschieht gegliedert nach Themenbereichen und mit verschiedenen Informationsmitteln über mehrere Kanäle.

Die statistischen Themenbereiche

- 00 Statistische Grundlagen und Übersichten
- 01 Bevölkerung
- 02 Raum und Umwelt
- 03 Arbeit und Erwerb
- 04 Volkswirtschaft
- 05 Preise
- 06 Industrie und Dienstleistungen
- 07 Land- und Forstwirtschaft
- 08 Energie
- 09 Bau- und Wohnungswesen
- 10 Tourismus
- 11 Mobilität und Verkehr
- 12 Geld, Banken, Versicherungen
- 13 Soziale Sicherheit
- 14 Gesundheit
- 15 Bildung und Wissenschaft
- 16 Kultur, Medien, Informationsgesellschaft, Sport
- 17 Politik
- 18 Öffentliche Verwaltung und Finanzen
- 19 Kriminalität und Strafrecht
- 20 Wirtschaftliche und soziale Situation der Bevölkerung
- 21 Nachhaltige Entwicklung, regionale und internationale Disparitäten

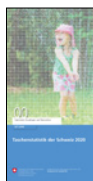
Die zentralen Übersichtspublikationen

Statistisches Jahrbuch der Schweiz



Das vom Bundesamt für Statistik (BFS) herausgegebene Statistische Jahrbuch ist seit 1891 das Standardwerk der Schweizer Statistik. Es fasst die wichtigsten statistischen Ergebnisse zu Bevölkerung, Gesellschaft, Staat, Wirtschaft und Umwelt des Landes zusammen.

Taschenstatistik der Schweiz



Die Taschenstatistik ist eine attraktive, kurzweilige Zusammenfassung der wichtigsten Zahlen eines Jahres. Die Publikation mit 52 Seiten im praktischen A6/5-Format ist gratis und in fünf Sprachen (Deutsch, Französisch, Italienisch, Rätoromanisch und Englisch) erhältlich.

Das BFS im Internet – www.statistik.ch

Das Portal «Statistik Schweiz» bietet Ihnen einen modernen, attraktiven und stets aktuellen Zugang zu allen statistischen Informationen. Gerne weisen wir Sie auf folgende, besonders häufig genutzte Angebote hin.

Publikationsdatenbank – Publikationen zur vertieften Information

Fast alle vom BFS publizierten Dokumente werden auf dem Portal gratis in elektronischer Form zur Verfügung gestellt. Gedruckte Publikationen können bestellt werden unter der Telefonnummer 058 463 60 60 oder per Mail an order@bfs.admin.ch.
www.statistik.ch → Statistiken finden → Kataloge und Datenbanken → Publikationen

NewsMail – Immer auf dem neusten Stand



Thematisch differenzierte E-Mail-Abonnemente mit Hinweisen und Informationen zu aktuellen Ergebnissen und Aktivitäten.
www.news-stat.admin.ch

STAT-TAB – Die interaktive Statistikdatenbank



Die interaktive Statistikdatenbank bietet einen einfachen und zugleich individuell anpassbaren Zugang zu den statistischen Ergebnissen mit Downloadmöglichkeit in verschiedenen Formaten.
www.stattab.bfs.admin.ch

Statatlas Schweiz – Regionaldatenbank und interaktive Karten



Mit über 4500 interaktiven thematischen Karten bietet Ihnen der Statistische Atlas der Schweiz einen modernen und permanent verfügbaren Überblick zu spannenden regionalen Fragestellungen aus allen Themenbereichen der Statistik.
www.statatlas-schweiz.admin.ch

Individuelle Auskünfte

Zentrale Statistik Information

058 463 60 11, info@bfs.admin.ch

Bei der Berechnung von Preisindizes müssen nicht nur die Preise, sondern auch die Merkmale der Güter verglichen werden. Bei Immobilien ist ein besonders hoher Anteil am beobachteten Preisunterschied zwischen den Objekten auf ihre unterschiedlichen Beschaffenheiten zurückzuführen.

Immobilien gelten als äusserst heterogenes Gut. Häuser und Wohnungen unterscheiden sich in zahlreichen Eigenschaften und in weitaus höherem Mass voneinander als Konsum- und Industriegüter. Entsprechend gross ist die Gefahr einer Verzerrung in einem Immobilienpreisindex, wenn den Qualitätsdifferenzen zwischen den Gütern, die in zwei aufeinander folgenden Perioden gehandelt wurden, nicht ausreichend Rechnung getragen wird.

Um die Preise von verschiedenen Liegenschaften vergleichbar zu machen, wendet das Bundesamt für Statistik (BFS) ein Qualitätsbereinigungsverfahren an. Mittels diesem filtert es die Preisdifferenz heraus, die durch die abweichende Qualität der einzelnen Objekte erklärbar ist. Beim Schweizerischen Wohnimmobilienpreisindex (IMPI), um den es in dieser Publikation geht, kommen eine Stratifizierung sowie ein hedonisches Modell des Typs «Hedonic Repricing» zur Anwendung. Diese Broschüre erläutert das Vorgehen des BFS bei der Herleitung des Modells.

Online

www.statistik.ch

Print

www.statistik.ch
Bundesamt für Statistik
CH-2010 Neuchâtel
order@bfs.admin.ch
Tel. 058 463 60 60

BFS-Nummer

2071-2002

ISBN

978-3-303-05764-3

**Statistik
zählt für Sie.**

www.statistik-zaehlt.ch