

Cordula Schöneich

hat Politikwissenschaft studiert und ist seit 2011 im Statistischen Bundesamt tätig, zunächst im Bereich Zensus. Seit 2014 arbeitet sie im Referat „Agrar-, Bau- und Immobilienpreise“ und befasst sich vor allem mit der Weiterentwicklung der Immobilienpreisstatistiken.

Markus Teske

studierte Volkswirtschaftslehre an der Universität Rostock und arbeitete anschließend am ZEW Mannheim. Seit März 2017 ist er als wissenschaftlicher Mitarbeiter im Referat „Agrar-, Bau- und Immobilienpreise“ des Statistischen Bundesamtes tätig und befasst sich mit der Weiterentwicklung des Häuserpreisindex.

REGIONALISIERUNG DES HÄUSERPREISINDEX

Cordula Schöneich, Markus Teske

↳ **Schlüsselwörter:** Wohnimmobilien – Preisindizes – Immobilienpreisstatistik – Preise für Wohneigentum – regionale Immobilienpreise

ZUSAMMENFASSUNG

Der Preisindex für Wohnimmobilien (kurz: Häuserpreisindex) misst vierteljährlich die Preisentwicklung der in Deutschland von Privatpersonen erworbenen Ein- und Zweifamilienhäuser und Eigentumswohnungen. Als Teil eines Indikatorensets dient er unter anderem der Beurteilung der Finanzstabilität. Seit Juni 2019 werden neben der bundesweiten Preisentwicklung auch Indizes für siedlungsstrukturelle Kreistypen veröffentlicht. Im Zuge der dazu nötigen Arbeiten wurden die Berechnungen methodisch verbessert und alternative hedonische Methoden getestet. Der vorliegende Aufsatz stellt sowohl die grundlegende Berechnungsweise als auch die wesentlichen methodischen Verbesserungen dar.

↳ **Keywords:** residential property – price indices – real property price statistics – residential property prices – regional real property prices

ABSTRACT

On a quarterly basis, the residential property price index (in short: house price index) measures the price development of single-family and two-family houses and freehold dwellings purchased by private individuals in Germany. As part of a set of indicators it is used, for example, for assessing financial stability. Since June 2019 the price development for Germany as a whole has been supplemented by the publication of indices for district types in terms of settlement structure. As part of the work required, the calculations were improved methodologically and alternative hedonic methods were tested. This article describes the basic calculation approach as well as the most important methodological improvements.

1

Einleitung

Die Immobilienpreise in Deutschland sind seit der Finanzkrise 2008 stark angestiegen. Zwischen 2010 und 2018 erhöhten sich die Preise für Eigentumswohnungen und Ein- und Zweifamilienhäuser bundesweit um 45 %.

Mit dieser Entwicklung und ihren potenziellen makroökonomischen und sozialen Auswirkungen¹ hat auch das Interesse an regelmäßigen und zuverlässigen Zahlen zur Entwicklung des Immobilienmarkts zugenommen. Um das Marktgeschehen adäquat zu analysieren, bedarf es verschiedener Informationen. Ein wichtiger Indikator ist die Entwicklung der Immobilienpreise. Das Statistische Bundesamt ist durch eine europäische Rechtsverordnung² seit 2013 verpflichtet, quartalsweise einen Preisindex für Wohnimmobilien, im Folgenden Häuserpreisindex (HPI), sowie dessen Teilindizes für neu erstellte und bestehende Wohnimmobilien zu berechnen. Der Häuserpreisindex bildet die bundesweite Preisentwicklung der von Privatpersonen erworbenen Wohnimmobilien ab.

Die Preisentwicklung für Wohnimmobilien kann regional sehr unterschiedlich sein. Um Ergebnisse für verschiedene Regionstypen bereitzustellen – insbesondere, um die Unterteilung nach der Dimension „Stadt – Land“ zu ermöglichen –, wurde der Häuserpreisindex weiterentwickelt.³ Im Juni 2019 hat das Statistische Bundesamt erstmals Preisindizes für Eigentumswohnungen sowie Ein- und Zweifamilienhäuser getrennt nach siedlungsstrukturellen Kreistypen veröffentlicht. Damit ist es möglich, die Entwicklung von Preisen in städtischen und ländlichen Regionen getrennt zu beobachten. Für die regelmäßige Veröffentlichung von Ergebnissen auf Ebene der Kreistypen waren methodische Anpassungen und Analysen der bisherigen Berechnung notwendig.

- 1 Für eine Darstellung der Zusammenhänge unterschiedlicher Faktoren auf den Immobilienmärkten siehe Sachverständigenrat (2018, Seite 332 ff.), Deutsche Bundesbank (2019) sowie Buch (2019).
- 2 Verordnung (EU) Nr. 93/2013 der Kommission vom 1. Februar 2013 zur Festlegung von Durchführungsbestimmungen zu der Verordnung (EG) Nr. 2494/95 des Rates über harmonisierte Verbraucherpreisindizes im Hinblick auf die Schaffung von Preisindizes für selbst genutztes Wohneigentum (Amtsblatt der EU Nr. L 33, Seite 14).
- 3 Für eine Beschreibung des methodischen Vorgehens während der Pilotphase der Statistik siehe zum Beispiel Dechent (2008) und Dechent (2011).

Diese wiesen auch auf Verbesserungspotenzial für die Berechnung des bundesweiten Häuserpreisindex hin, sodass eine Revision der Ergebnisse rückwirkend bis 2016 umgesetzt wurde.

Das folgende Kapitel 2 gibt zunächst einen Überblick über die HPI-Datengrundlage. Die allgemeine Methodik zur Berechnung des Häuserpreisindex, die Abbildung der Lagequalität in der Regression sowie die Verbesserung und weitere Automatisierung der Plausibilisierung beschreibt Kapitel 3. Es zeigt zudem unter Anwendung alternativer Methoden Testrechnungen, die zum Ziel hatten, auch für Teilindizes mit weniger Datensätzen stabile Ergebnisse zu berechnen.

Kapitel 4 stellt erste Ergebnisse des regionalisierten Häuserpreisindex vor, das abschließende Kapitel 5 gibt einen Ausblick auf weitere Herausforderungen und mögliche Verbesserungen auf dem Feld der Wohnimmobilienpreisbeobachtung.

2

Datengrundlage

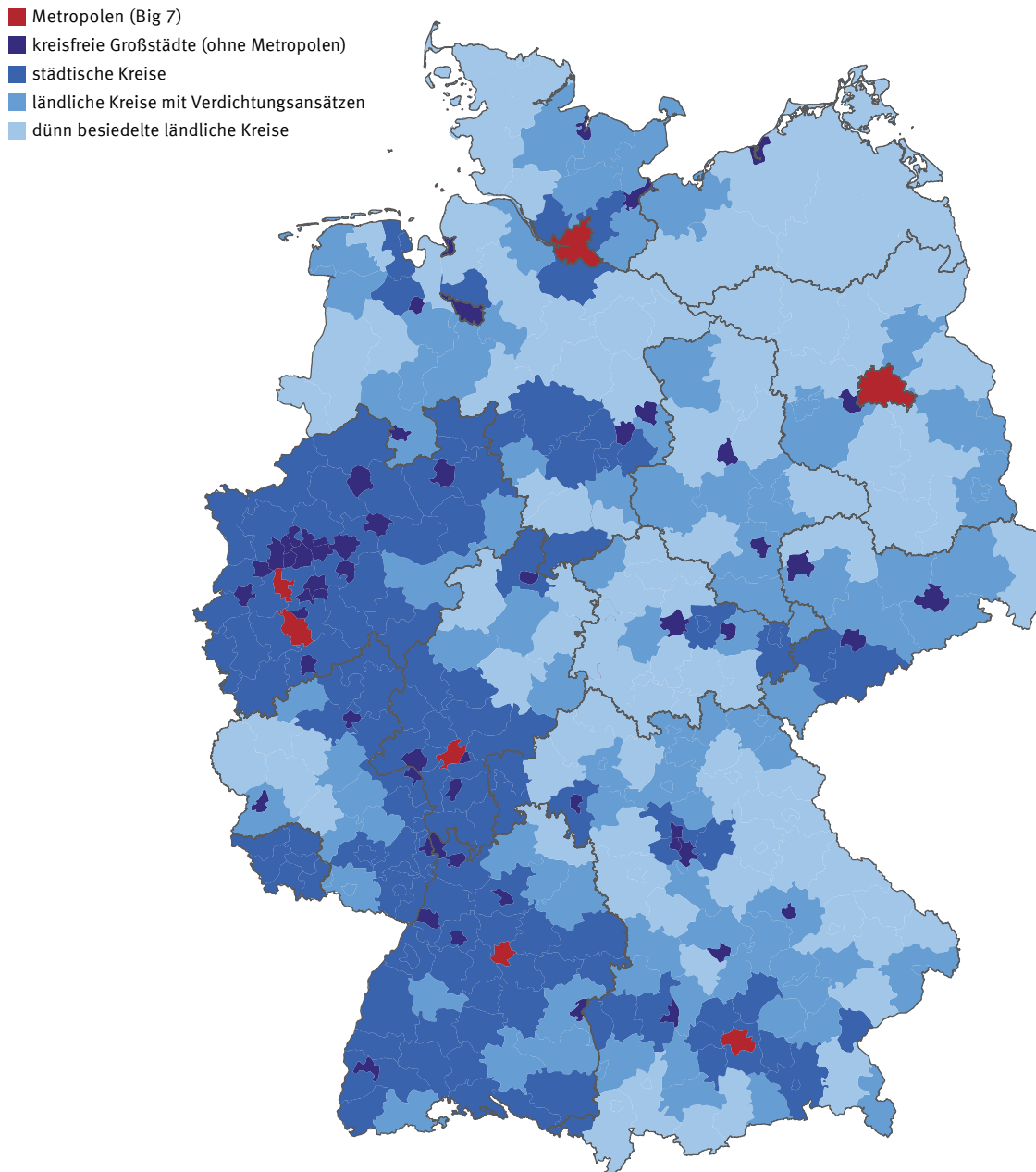
Der Häuserpreisindex wird auf Grundlage von Transaktionsdaten der Gutachterausschüsse für Grundstückswerte berechnet.⁴ In Deutschland sind Notare dazu verpflichtet, die Kaufverträge über den Erwerb/Verkauf einer Immobilie an den lokal zuständigen Gutachterausschuss zu übermitteln. Die Gutachterausschüsse übertragen die in dem Vertrag enthaltenen Informationen in ihre Kaufpreissammlungen; soweit erforderlich erfragen sie zusätzliche Informationen bei den Käufern, wie etwa das Alter der Immobilie. Die Gutachterausschüsse übermitteln quartalsweise Daten über Käufe von Wohnimmobilien, wie Preise und Eigenschaften der verkauften Immobilien, online an das Statistische Bundesamt.

Für die Datenlieferung zur Berechnung des Häuserpreisindex werden sämtliche Gutachterausschüsse in Deutschland herangezogen, mit Ausnahme derjenigen in Baden-Württemberg. Aufgrund der kleingliedrigen Orga-

- 4 Die (Oberen) Gutachterausschüsse für Grundstückswerte sind gemäß §§ 192 ff. Baugesetzbuch unabhängige Gremien, unter anderem mit der Aufgabe, Transparenz auf dem deutschen Grundstücksmarkt zu schaffen. Sie sind in der Regel auf Ebene der Landkreise beziehungsweise kreisfreien Städte organisiert.

Grafik 1

Siedlungsstrukturelle Kreistypen des Bundesinstituts für Bau-, Stadt- und Raumforschung



© GeoBasis-DE / BKG 01.01.2016

2020 - 01 - 0059

nisation der Gutachterausschüsse in diesem Bundesland – in der Regel auf Ebene der Gemeinden – sind hier nur die Gutachterausschüsse zur Meldung verpflichtet, in deren Zuständigkeitsbereich Gemeinden mit mehr als

20 000 Einwohnerinnen und Einwohnern liegen. Abgesehen von Baden-Württemberg beruht der Häuserpreisindex somit prinzipiell auf sämtlichen marktüblichen bei den Gutachterausschüssen registrierten Transaktionen

von Eigentumswohnungen sowie Ein- und Zweifamilienhäusern in Deutschland.

Für die Regionalisierung des Häuserpreisindex wurde die Ebene der siedlungsstrukturellen Kreistypen des Bundesinstituts für Bau-, Stadt- und Raumforschung (BBSR) gewählt. Das BBSR gruppiert die Landkreise beziehungsweise kreisfreien Städte Deutschlands anhand ihrer Siedlungsstruktur, unter anderem der Einwohnerdichte, in vier Typen: kreisfreie Großstädte, städtische Kreise, ländliche Kreise mit Verdichtungsansätzen und dünn besiedelte ländliche Kreise. Für die Berechnung des Häuserpreisindex wurden darüber hinaus die sieben größten deutschen Städte (Berlin, Hamburg, München, Köln, Frankfurt am Main, Stuttgart und Düsseldorf) als gesonderter Kreistyp ausgegliedert. Diese Ausgliederung, die sogenannten Big 7, ist als Analyseebene auf dem deutschen Immobilienmarkt weit verbreitet.

↳ Grafik 1

Die siedlungsstrukturellen Kreistypen ermöglichen eine Unterteilung nach städtischen und ländlichen Regionen, deshalb wurden sie für die Berechnung regionaler Indizes gewählt. Ein weiterer Vorteil ist die Vergleichsmöglichkeit mit anderen Statistiken, die auf Kreisebene oder Kreistypenebene vorliegen. So werden beispielsweise die Mieten als Teil des Verbraucherpreisindex seit Februar 2019 ebenfalls für Kreistypen ausgewiesen. Zudem sind die siedlungsstrukturellen Kreistypen zeitlich relativ stabil. Das Bundesinstitut für Bau-, Stadt- und Raumforschung aktualisiert sie je nach Notwendigkeit etwa alle ein bis zwei Jahre, wobei normalerweise nur wenige Kreise neu zugeordnet werden. Somit ist auch über längere Zeiträume hinweg eine Interpretation der Entwicklung der Indizes möglich. Für die Berechnung der regionalen Indizes wurden die Kreistypen zum Stand 2015 genutzt.

3

Methodik

Die Preisstatistik soll Preisentwicklungen in einem Index so abbilden, dass qualitative Unterschiede der Produkte den Preisindex nicht beeinflussen (reiner Preisvergleich). Da jede Immobilie einzigartig ist, kann im Gegensatz zur üblichen Herangehensweise der amtlichen Preisstatistik in der Regel kein direkter Preisver-

gleich desselben Produkts erfolgen. Zum einen wird ein und dieselbe Immobilie teilweise Jahrzehnte gehalten, bevor es zu einer erneuten Transaktion kommt. Zum anderen werden Preise beziehungsweise Verkehrswerte von Immobilien zwischen Transaktionen nur selten gutachterlich ermittelt und stehen damit der amtlichen Statistik nicht zur Verfügung.

Vor diesem Hintergrund sind spezielle Methoden notwendig, um Preisindizes für Wohnimmobilien ermitteln zu können. Diese werden in den folgenden Abschnitten erläutert. Zunächst wird dargestellt, wie Preisindizes für das heterogene Gut „Wohnimmobilien“ allgemein berechnet werden können, anschließend, wie dabei die Lage der Immobilie als wichtiges preisbestimmendes Merkmal integriert wird. Schließlich wird die Plausibilisierung der Daten beschrieben, bevor abschließend experimentelle Indizes unter Anwendung alternativer hedonischer Methoden gegenübergestellt werden.

3.1 Hedonisches Regressionsmodell

Zur Berechnung von Preisindizes für Wohnimmobilien können unterschiedliche Methoden verwendet werden (siehe zum Beispiel Hill und andere, 2018). Eurostat, das Statistische Amt der Europäischen Union, benennt als zu favorisierende Methode die Hedonik.⁵

Das Statistische Bundesamt nutzt die Methode der hedonischen doppelten Imputation. Ziel ist es, für jede zu einem Zeitpunkt verkaufte Immobilie zu schätzen, was diese Immobilie zu einem anderen Zeitpunkt gekostet hätte. Grundidee dabei ist, dass jedes Objektmerkmal einen Einfluss auf den Preis hat und dass man diese Einflüsse quantifizieren kann, um Immobilien mit ähnlicher Ausstattung und in ähnlichem Zustand miteinander zu vergleichen. Bei der doppelten Imputation wird zunächst auf Basis der Transaktionen einer Periode eine Regression geschätzt, in welcher der Preis einer Immobilie durch ihre Eigenschaften erklärt wird. Dasselbe Vorgehen wird für die Vergleichsperiode anhand der in dieser Periode gehandelten Immobilien wiederholt. Auf Basis der nun ermittelten Regressionskoeffizienten für beide Perioden kann anhand der Objektmerkmale für jede verkaufte Immobilie bestimmt werden, wie hoch ihr geschätzter Preis in der jeweils anderen Periode wäre.

⁵ Für eine Diskussion der unterschiedlichen Methoden siehe Eurostat (2013) oder Eurostat (2017).

So ist es möglich, für jede Immobilie einen Preis in beiden Perioden zu imputieren.¹⁶ Für die Berechnung der Elementarindizes werden die Preisverhältnisse aller Immobilien in den beiden betrachteten Perioden als geometrisches Mittel aggregiert (Jevons-Index).

Für den Häuserpreisindex für Deutschland werden für Häuser folgende Regressionsmodelle berechnet¹⁷:

$$\ln(p_{H\ddot{a}user}) = \beta_0 + \beta_1 \cdot \ln(Wohn) + \beta_2 \cdot \ln(Grund) + \beta_3 \cdot Alter + \beta_4 \cdot Alter^2 + \beta_5 \cdot Alter^3 + \sum_{i=1}^p \beta_{i+5} \cdot d_{Lage_i} + \varepsilon$$

Hierbei steht „Wohn“ für die Wohnfläche, „Grund“ für die Grundstücksfläche und „Alter“ für das Alter der jeweiligen Immobilie in Jahren. Die Variablen „d_Lage“ stellen Lagedummies dar, die je nach Kreistyp aus Bodenrichtwert-Klassen, Lagequalität-Klassen oder einer Kombination aus beiden bestehen (siehe Abschnitt 3.2).

Bei Wohnungen wird ein ähnliches Modell genutzt, hier allerdings ohne die Grundstücksfläche:

$$\ln(p_{Wohnungen}) = \beta_0 + \beta_1 \cdot \ln(Wohn) + \beta_2 \cdot Alter + \beta_3 \cdot Alter^2 + \beta_4 \cdot Alter^3 + \sum_{i=1}^p \beta_{i+4} \cdot d_{Lage_i} + \varepsilon$$

Diese Regressionsmodelle entsprechen in ihrer Struktur jenen, die auch für die regionalisierten Häuserpreisindizes gerechnet werden.

3.2 Abbildung der Lage

Ein wichtiges preisbestimmendes Merkmal einer Immobilie ist ihre Lage. Ihre adäquate Berücksichtigung bei der Berechnung von Immobilienpreisindizes birgt sowohl methodische als auch praktische Herausforderungen.

6 Die doppelte Imputation wird in der Literatur der einfachen Imputation vorgezogen. Grund dafür ist, dass Verzerrungen durch möglicherweise preisbeeinflussende, nicht im Regressionsmodell berücksichtigte Variablen (sogenannte Omitted Variable Bias – Verzerrung durch ausgelassene Variablen) eher vermieden werden können (Hill, 2013).

7 Die Variable „Alter“ wurde vor Juni 2019 logarithmiert in das Regressionsmodell aufgenommen, jedoch zeigte sich, dass ein Polynom zu einer besseren Anpassung führt. Zudem wurden zuvor für Häuser und Wohnungen jeweils getrennte Regressionen für die Segmente Neubau und Bestand geschätzt, während nun nur eine Regression für Neubau und Bestand zusammen ermittelt wird.

Die Lage wird hauptsächlich über die Bodenrichtwerte definiert. Die Gutachterausschüsse für Grundstückspreise erstellen nach gesetzlichen Vorgaben (§196 Baugesetzbuch) Bodenrichtwerte für zusammengefasste, ähnliche Grundstücke mit weitgehend gleicher Nutzungsart. Der Bodenrichtwert gibt für hypothetisch unbebauten Boden den durchschnittlichen Lagewert in Euro je Quadratmeter an. Durch gesetzliche Vorgaben und Verordnungen – zum Beispiel die Bodenrichtwertrichtlinie (BRW-RL) – sind Bodenrichtwerte prinzipiell bundesweit vergleichbar und eignen sich daher als Lagemerkmale für die Regressionen zur Ermittlung des Häuserpreisindex. Die Gutachterausschüsse sind verpflichtet, die Bodenrichtwerte mindestens alle zwei Jahre zu aktualisieren.

Dazu werten die Gutachterausschüsse die Kauffälle innerhalb einer Bodenrichtwertzone beziehungsweise Kauffälle ähnlicher Grundstücke aus. Folglich ist die Entwicklung der Bodenrichtwerte zu einem gewissen Grad abhängig von der Entwicklung der Immobilienpreise. Wird in den Regressionen zu verschiedenen Zeitpunkten eine Lage unterschiedlich bewertet, könnte dies zu einer Unterschätzung der Preisentwicklung führen. Um dies zu vermeiden, werden die Bodenrichtwerte für die Berechnung des Häuserpreisindex klassiert. Die Anzahl der Klassen und die Klassengrenzen variieren je nach Kreistyp und Teilindex.

↳ Exkurs

Die grundsätzliche Idee der Hedonik ist es, Qualitätsunterschiede von Produkten aus deren Preis herauszurechnen. Bei der Lage oder genauer der Lagequalität einer Immobilie handelt es sich jedoch um eine besondere Qualität. Diese kann zum Beispiel beschrieben werden durch vorhandene Infrastruktur, die soziodemografische Struktur, die vorhandenen Grünanlagen und andere Merkmale. Die Erwartung an einen Preisindex für Wohnimmobilien ist jedoch eher, dass er die Preisentwicklung eines geografisch-physischen Orts beschreibt und dabei zunächst unberücksichtigt lässt, ob sich der Ort verändert hat. Man möchte beschreiben können, dass Preise gestiegen sind, obgleich beispielsweise ein Viertel einer Großstadt sich „verbessert“ hat. Daher sollte das Ziel eines hedonisch bereinigten Preisindex für Wohnimmobilien sein, Qualitätsveränderungen der Lage zumindest mittelfristig nicht aus der Preisentwicklung „herauszurechnen“.

Bei den Arbeiten zur Regionalisierung des Häuserpreisindex zeigte sich, dass die klassierten Bodenrichtwerte nicht immer optimal geeignet sind, die Lage einer Immobilie im Regressionsmodell abzubilden. Insbesondere in den Metropolen und Großstädten sind in den vergangenen Jahren die Bodenrichtwerte stark angestiegen. Dies konnte dazu führen, dass geografisch gleiche Lagen in einer Periode einer anderen Bodenrichtwertklasse zugeordnet wurden als in der vorausgegangenen Periode – und zwar dann, wenn die Gutachterausschüsse die Bodenrichtwerte neu bewerten.¹⁸

Eine intuitive Lösung des Problems bestünde darin, die Grenzen zwischen den Bodenrichtwert-Klassen regelmäßig zu aktualisieren. Da jedoch keine flächendeckenden Informationen über die zeitliche Entwicklung der Bodenrichtwerte zur Verfügung stehen, können die Klassengrenzen höchstens schätzungsweise angepasst werden. Um diese Probleme zu beheben, wurde geprüft, ob auf eine alternative Abbildung der Lage ausgewichen werden kann.

Die Gutachterausschüsse vieler Großstädte bestimmen für ihr Zuständigkeitsgebiet sogenannte Wohnlagen oder Lagequalitäten. Diese beziehen neben den Bodenrichtwerten – je nach ermittelndem Gutachterausschuss – weitere Kriterien ein, beispielsweise die Anbindung an den öffentlichen Personennahverkehr. Die sachverständige Einschätzung der ortskundigen Mitarbeiterinnen und Mitarbeiter der jeweiligen Gutachterausschüsse spielt dabei ebenso eine wichtige Rolle. Die Lagequalität unterliegt im Gegensatz zu den Bodenrichtwerten in der Regel keinen regelmäßigen, flächendeckenden Änderungen in kurzen Zeitabständen und ist folglich eine stabilere Größe. Die Gutachterausschüsse übermitteln für den Häuserpreisindex vier Lagequalitäten von „einfacher Lage“ bis „Spitzenlage“. In Metropolen und in einigen Großstädten ist dieses Merkmal in den Meldungen der Gutachterausschüsse sehr gut ausgefüllt und konnte daher als alternatives Lagemerkmal getestet werden.

Vor diesem Hintergrund erschien die Lagequalität als Lagemerkmal in Städten besser geeignet als der Bodenrichtwert. Deshalb wurden in Testrechnungen im Kreis-

typ „Metropolen“ keine Bodenrichtwert-Klassen, sondern Lagequalität-Klassen als Dummyvariablen in den Regressionen verwendet. Im Kreistyp „Kreisfreie Großstädte (ohne Metropolen)“ übermitteln nicht alle Gutachterausschüsse Lagequalitäten und die Transaktionszahlen je Lagequalität sind innerhalb einer Stadt oft zu niedrig, um verlässliche Koeffizienten für die entsprechenden Dummyvariablen zu berechnen. Daher wurden für diesen Kreistyp wo möglich Lagequalität-Klassen genutzt und Bodenrichtwert-Klassen beibehalten, sofern erstere nicht verfügbar waren. Die Lagequalität ist nicht bundesweit standardisiert. Daher wurden je Metropole/Großstadt separate Dummyvariablen für Lagequalität-Klassen gebildet. Klassen wurden daher wo nötig innerhalb einer Metropole/Großstadt zusammengefasst, um eine ausreichende Zahl von Datensätzen je Lage in der Regression nutzen zu können.

Der Häuserpreisindex für Deutschland wies unter Verwendung von Bodenrichtwertklassen in den ersten Quartalen der Jahre 2017 und 2018 leichte Rückgänge gegenüber dem Vorquartal auf. Bei Nutzung von Lageklassen zeigen sich hingegen Anstiege in diesen Quartalen. Bei Testrechnungen für den Kreistyp Metropolen wurden diese Unterschiede ebenfalls festgestellt.

Diese Ergebnisse wurden in der Gesamtschau als plausibler eingestuft, da sie zum einen die theoretischen Überlegungen zur Verwendung von Bodenrichtwerten und Lagequalitäten bestätigen. Zum anderen ließ das allgemeine Marktumfeld Preisrückgänge, die nur in den ersten Quartalen auftraten, insbesondere in den Metropolen nicht plausibel erscheinen. Folglich wurden für die entsprechenden Metropolen und Großstädte die Lagequalität-Klassen auch für die Berechnung des Häuserpreisindex für Deutschland rückwirkend ab 2016 eingeführt und der Index revidiert.¹⁹

Vor der Revision wurde für das erste Quartal 2017 ein Preisrückgang im Vergleich zum Vorquartal von 1,0% ausgewiesen, für das erste Quartal 2018 ein Preisrückgang von 0,2% zum Vorquartal. Nach der Revision resultieren für diese beiden Quartale Preisanstiege von 0,5 beziehungsweise 0,9%. Zudem ergeben sich für die meisten übrigen Quartale im Zeitraum vom ersten Quartal 2016 bis zum vierten Quartal 2018 höhere Veränderungsraten als vor der Revision.

¹⁸ Eine zusätzliche Problematik bei der Verwendung von Bodenrichtwerten kann insbesondere im städtischen Raum durch die Geschossflächenzahl zustande kommen. Nicht alle Gutachterausschüsse liefern auf eine fixe Geschossflächenzahl normierte Bodenrichtwerte. Diese Problematik soll hier nicht weiter vertieft werden.

¹⁹ In den übrigen Kreistypen ist das Merkmal Lagequalität in den Meldungen der Gutachterausschüsse kaum befüllt, sodass hier weiterhin auf die Bodenrichtwerte zurückgegriffen wird.

Die Lagequalität von Immobilien stellt derzeit eine gute Möglichkeit dar, Lagen innerhalb von Metropolen und einigen Großstädten adäquat zu berücksichtigen. Es wäre dennoch wünschenswert, auf ein standardisiertes Kriterium zur Darstellung der Lage einer Immobilie zurückgreifen zu können. Auch die Variable Lagequalität kann mittelfristig Anpassungen unterliegen. Derzeit liegen dem Statistischen Bundesamt keine verwertbaren Anschrifteninformationen unterhalb der Gemeindeebene vor. Um auch künftig die Qualität der Ergebnisse sicherzustellen, wäre von Vorteil, künftig auch Anschrifteninformationen und Geokoordinaten zu erheben und für die Berechnung zu testen. Denkbar wäre beispielsweise, mithilfe von Geokoordinaten einerseits räumliche Abhängigkeiten strukturell in den hedonischen Modellen zu berücksichtigen. Andererseits könnten Geokoordinaten mit frei verfügbaren Points of Interest verschnitten werden, um ein alternatives Lagemerkmal zu konstruieren.

Die Verwendung von Geokoordinaten für die Berechnung von Immobilienpreisindizes wird in der internationalen Literatur diskutiert. Ob Geokoordinaten gegenüber Dummyvariablen zum Beispiel für administrative Gebietseinheiten, Postleitzahlgebiete oder – wie für den deutschen Häuserpreisindex – Lageklassen/Bodenrichtwertklassen vorzuziehen sind, ist nicht pauschal zu beantworten.¹⁰ Umfangreiche Testrechnungen wären nötig, um die Umsetzbarkeit zu prüfen und Ergebnisse bewerten zu können.

3.3 Plausibilisierung der Daten

Neben der Auswahl der Variablen für das Regressionsmodell zur Berechnung des Häuserpreisindex spielt auch die Identifikation von Ausreißern in den Daten eine wichtige Rolle für valide Ergebnisse. Ausreißer beziehungsweise einflussreiche Beobachtungen können große Auswirkungen auf die geschätzten Koeffizienten und somit letztlich auf die imputierten Preise haben (Silver, 2018). Dies ist besonders wichtig bei der Berechnung regionaler Indizes, da aufgrund der geringeren Anzahl an Kauffällen wenige Ausreißer die Ergebnisse

stärker beeinflussen können.¹¹ Um dies zu vermeiden wurde die Ausreißerbereinigung verbessert und weiter automatisiert.

Die Plausibilisierung und Aufbereitung der Daten erfolgt in vier Schritten: 1. Bereinigung um offensichtliche Fehler und Inkonsistenzen, 2. Anwendung fester jährlich aktualisierter Abschneidegrenzen, 3. Prüfung auf mehrdimensionale Ausreißer durch Cook's Distance und Least Trimmed Squares und schließlich 4. eine visuelle Kontrolle von Streudiagrammen.

Die von den Gutachterausschüssen gelieferten Daten werden zunächst auf offensichtliche Eingabefehler und Inkonsistenzen geprüft und korrigiert. Etwa 45 % der übermittelten Datensätze können nicht in die Berechnung eingehen, da sie fehlende Werte in einem der Kernmerkmale aufweisen.

Bereits vor den Arbeiten zur Regionalisierung des Häuserpreisindex wurden Plausibilitäts-Grenzen nach Kreistypen unterschieden, jedoch einige Besonderheiten zugelassen, wie etwa sehr hochpreisige Häuser auf Sylt, das im Kreis Nordfriesland liegt und damit dem ländlichsten Kreistyp zugeordnet ist. Sie fallen im Bundesergebnis kaum ins Gewicht, sind jedoch für den Kreistyp „dünn besiedelte ländliche Kreise“ sehr untypisch und beeinflussen diesen Index zum Teil übermäßig stark. Das war insbesondere dann der Fall, wenn die Zahl der Datensätze aus diesen Regionen mit außergewöhnlichen Preisen zwischen Beobachtungs- und Referenzzeitpunkt stark variierte. Die Extremwerte innerhalb eines Kreistyps werden nun streng als Ausreißer behandelt und von der Berechnung ausgeschlossen. Zu diesem Zweck werden jeweils alle Kauffälle für Wohnungen und Häuser eines gesamten Jahres betrachtet. Für die Merkmale, die als abhängige oder unabhängige Variablen in die Regression eingehen sollen¹², wurden das 0,5%- und das 99,5%-Quantil bestimmt¹³. Die so ermittelten Grenzwerte werden ein Jahr konstant gehalten.

10 Hill/Scholz (2018) zeigen beispielsweise für Sydney, dass sich Immobilienpreisindizes kaum unterscheiden, wenn die Lage über Postleitzahlgebiete oder Geokoordinaten abgebildet wird. Die Autoren führen dies auch darauf zurück, dass Postleitzahlgebiete in Sydney jeweils relativ kleine Gebiete umfassen.

11 Aus verschiedenen Gründen werden die Kreistypenergebnisse derzeit mit getrennten Regressionen berechnet – insbesondere aber aufgrund der unterschiedlichen Abbildung der Lage der Immobilien in den unterschiedlichen Kreistypen. Damit einher geht die Entscheidung, kompliziertere Interaktionsterme der Lagevariablen mit den anderen erklärenden Variablen zu vermeiden.

12 Kaufpreis, Wohnfläche, Bodenrichtwert sowie Grundstücksfläche bei Häusern.

13 Eurostat schlägt als „Standardlösung“ vor, einfach jeweils das 1%- beziehungsweise das 99%-Quantil als Abschneidegrenze anzusetzen.

ten und bilden für die quartalsweise Plausibilisierung der Daten die Abschneidegrenzen. Gegenüber einer Abschneidegrenze auf Basis der obersten/untersten Quantile eines Quartals bietet das den Vorteil, dass in einem Quartal fehlerhaft gemeldete Daten eines Gutachterausschusses die Grenzwerte nicht beziehungsweise weniger beeinflussen können.

Diese univariate Ausreißerbereinigung wird ergänzt um Prüfungen der Robustheit der Ergebnisse anhand von Cook's Distance. Cook's Distance misst den Einfluss einer Beobachtung auf die Regressionskoeffizienten, indem sie bestimmt, welchen Effekt das Auslassen dieser Beobachtung auf das Regressionsergebnis hat. Bei diesem Verfahren wird allerdings nur der Einfluss einzelner Beobachtungen auf die Regression berücksichtigt. Häufig ist jedoch der Fall, dass mehrere Kauffälle gemeinsam ein Ausreißercluster bilden. In diesen Fällen stößt Cook's Distance an seine Grenzen. Das Verfahren wird daher nur flankierend für die Berechnung des Häuserpreisindex angewendet. Nur bei einer deutlichen Differenz des resultierenden Index werden die Fälle genauer betrachtet und gegebenenfalls von der Berechnung ausgeschlossen.

Um auch Gruppen von (multidimensionalen) Ausreißern automatisiert erkennen und eliminieren zu können, wurde Least Trimmed Squares (LTS) als zusätzliches Verfahren eingeführt (Rousseeuw, 1984). Grundidee von LTS ist das wiederholte Ziehen von Zufallsstichproben aus der Gesamtheit der zu berücksichtigenden Datensätze (n), für die jeweils eine OLS-Regression mit denselben Variablen wie im eigentlichen Zielmodell berechnet wird. Minimiert wird

$$\sum_{i=1}^h e_i^2,$$

wobei $e_1^2 \leq \dots \leq e_n^2$ die geordneten quadrierten Residuen sind und $h = (3n + p + 1)/4$, mit p als Anzahl der Koeffizienten.¹⁴ Zufallsstichproben, die keine Ausreißercluster enthalten, weisen eine geringere Summe quadratischer Abweichungen auf als andere.¹⁵

14 In der Literatur wird meist eine Stichprobengröße von $h = (n + p + 1)/2$ angegeben, in der statistischen Auswertungssoftware SAS ist jedoch standardmäßig der oben genannte Umfang implementiert.

15 Um nicht für alle möglichen Zufallsstichproben Regressionen schätzen zu müssen, ist in SAS ein sogenannter C-step implementiert. Den Grundgedanken erläutern Rousseeuw/van Driessen (2000).

Ausgehend von dem Modell für die Zufallsstichprobe, bei der die quadratischen Abweichungen minimiert sind, werden standardisierte Residuen für alle n Beobachtungen eines Quartals berechnet. Beobachtungen, deren standardisiertes Residuum um mehr als vier Standardabweichungen vom Mittelwert abweicht, werden als Ausreißer definiert und für die Regression und anschließende doppelte Imputation nicht berücksichtigt.

3.4 Alternative hedonische Methoden

Der Häuserpreisindex wird auf Bundesebene getrennt nach Neubau und Bestand veröffentlicht. Ziel war es, entsprechende Ergebnisse für Kreistypen zu berechnen. Die Indizes für den Neubau, die auf deutlich weniger Datensätzen beruhen als der Bestand, erwiesen sich auf Ebene der Kreistypen zum Teil als sehr volatil. Daher wurde getestet, ob sich die Volatilität verringern lässt, indem alternative hedonische Methoden angewendet werden. Einige hedonische Methoden beziehen mehr Kauffälle ein als nur die eines einzelnen Quartals, um Qualitätsunterschiede der gehandelten Immobilien zu berücksichtigen. Zwei Verfahren werden in der Literatur empfohlen, wenn eine geringe Zahl an Datensätzen für die Berechnung zur Verfügung steht.¹⁶ Diese Verfahren werden im Folgenden zunächst skizziert, bevor damit berechnete Indizes der hedonischen doppelten Imputation gegenübergestellt werden.

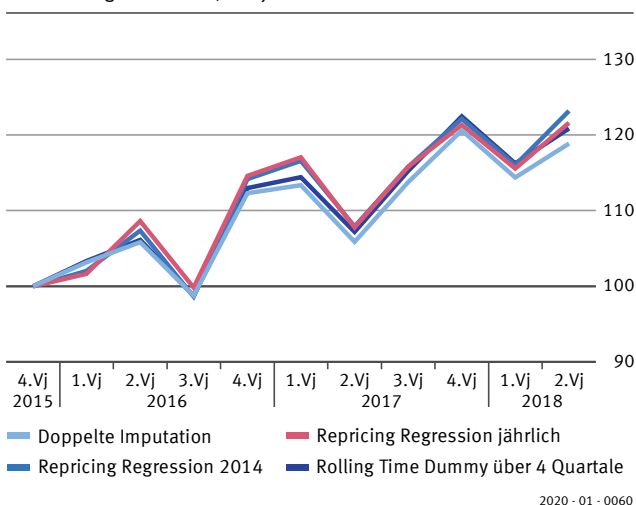
Bei der Rolling-Time-Dummy-Methode (RTD) wird für die Kauffälle mehrerer Quartale gemeinsam ein Regressionsmodell geschätzt. Dabei werden zusätzlich Dummy-Variablen für jedes Quartal eingeführt. Aus den Koeffizienten der Dummy-Variablen lassen sich direkt die durchschnittlichen Preisveränderungsraten im Vergleich zum ältesten Quartal ableiten. Anschließend wird das Zeitfenster sukzessive bis zum aktuellen Rand rollierend verschoben. Die Wahl der Größe des Zeitfensters beziehungsweise der in einer Regression berücksichtigten Quartale ist eine Ermessensentscheidung. Das Zeitfenster sollte allerdings nicht zu groß sein, da sich die Koeffizienten der Objektmerkmale im Zeitverlauf verändern können. In der Praxis anderer Mitgliedstaaten der Europäischen Union (EU) werden beispielsweise häufig zwei oder vier benachbarte Quartale kombiniert.

16 Siehe Hill und andere (2018) sowie Silver (2018).

Die Repricing-Methode stellt auf die Berechnung einer nicht qualitätsbereinigten Preisveränderungsrate und einen Qualitätsanpassungsfaktor ab. Erstere ist das Verhältnis der geometrischen Mittelwerte der Kaufpreise zweier zu vergleichender Quartale. Für die Bestimmung des Qualitätsanpassungsfaktors werden wieder Regressionsmodelle ähnlich denen in Abschnitt 3.1 geschätzt. Hierbei können wie bei der RTD-Methode Transaktionen mehrerer benachbarter Quartale verwendet werden. Die daraus ermittelten Koeffizienten für die Objektmerkmale werden genutzt, um die Qualitätsunterschiede der gehandelten Immobilien verschiedener Quartale zu berücksichtigen. Dazu wird je Quartal für eine hypothetische, anhand ihrer einzelnen Objektmerkmale (zum Beispiel Wohnfläche, Alter)¹⁷ durchschnittliche Immobilie der Preis mittels der Koeffizienten geschätzt. Die Regressionskoeffizienten für den Qualitätsanpassungsfaktor werden dabei für einen selbst gewählten Zeitraum fix gehalten. Allerdings sollte regelmäßig ein Regressionsmodell, beruhend auf aktuelleren Transaktionen, geschätzt werden, damit über die Zeit variierenden impliziten Preisen der Objektmerkmale Rechnung getragen wird.

↳ Grafik 2 stellt die hedonische doppelte Imputation als Standardberechnungsmethode der RTD- und der Repricing-Methode gegenüber. Beispielhaft werden

Grafik 2
Experimentelle Häuserpreisindizes unter Verwendung verschiedener hedonischer Methoden
Neu erstellte Häuser im Kreistyp "Ländliche Kreise mit Verdichtungsansätzen", 4. Vj 2015 = 100



die Berechnungen für neu erstellte Häuser im Kreistyp „ländliche Kreise mit Verdichtungsansätzen“ dargestellt. Die alternativen Berechnungsmethoden führen trotz breiterer Datengrundlage für die Regressions-schätzungen zu ähnlich volatilen Indexverläufen wie die doppelte Imputation. Hill und andere (2018) ermitteln auf Basis von Transaktionsdaten für Tokio und Sydney unter Verwendung unterschiedlicher hedonischer Methoden ebenfalls relativ ähnlich verlaufende Indizes.

Die deutliche Volatilität des Index kann auch andere Ursachen als die reine Preisentwicklung haben. Die geringen Transaktionszahlen je Quartal können dazu führen, dass zufällige Änderungen in der Zusammensetzung der Kauffälle von einem Quartal zum nächsten eher auftreten und die Abbildung der Preisentwicklung potenziell beeinflussen können. Dies gilt insbesondere dann, wenn relevante preisbestimmende Objektmerkmale nicht in der Regression berücksichtigt werden können. Ein Indiz dafür ist das relativ geringe Bestimmtheitsmaß in den Regressionen für neu erstellte Ein- und Zweifamilienhäuser im Kreistyp „ländliche Kreise mit Verdichtungsansätzen“. Es beträgt zwischen dem vierten Quartal 2015 und dem zweiten Quartal 2018 nur durchschnittlich etwa 50% je Quartal, während es beispielsweise für Häuser im deutschlandweiten Häuserpreisindex im Durchschnitt bei 75% liegt. Dies legt nahe, dass für neue Häuser im ländlichen Raum preisbestimmende Objektmerkmale derzeit unberücksichtigt sind. Daher soll für eine angedachte künftige Überarbeitung des Merkmalskatalogs geprüft werden, ob zum Beispiel der Ausstattungsstandard erhoben werden kann. Zudem wurde in weiteren Testrechnungen für Kreistypenergebnisse eine gemeinsame Regression für Neubau und Bestand berechnet und anschließend die geschätzten Preise getrennt nach Neubau und Bestand aggregiert, analog dem aktuellen Vorgehen beim bundesweiten Häuserpreisindex. Es resultierten jedoch ähnlich volatile Indizes.

Im Zusammenspiel dieser Betrachtungen und anderen Testrechnungen¹⁸ wurde entschieden, für neue Häuser und Wohnungen keine Indizes auf Ebene der Kreistypen zu berechnen und zu veröffentlichen. Stattdessen werden für neue und bestehende Häuser beziehungsweise Wohnungen gemeinsame Regressionen geschätzt und Indizes berechnet.

17 Bei Dummyvariablen wird der durchschnittliche Anteil an Transaktionen verwendet, für die die Dummyvariable den Wert 1 annimmt.

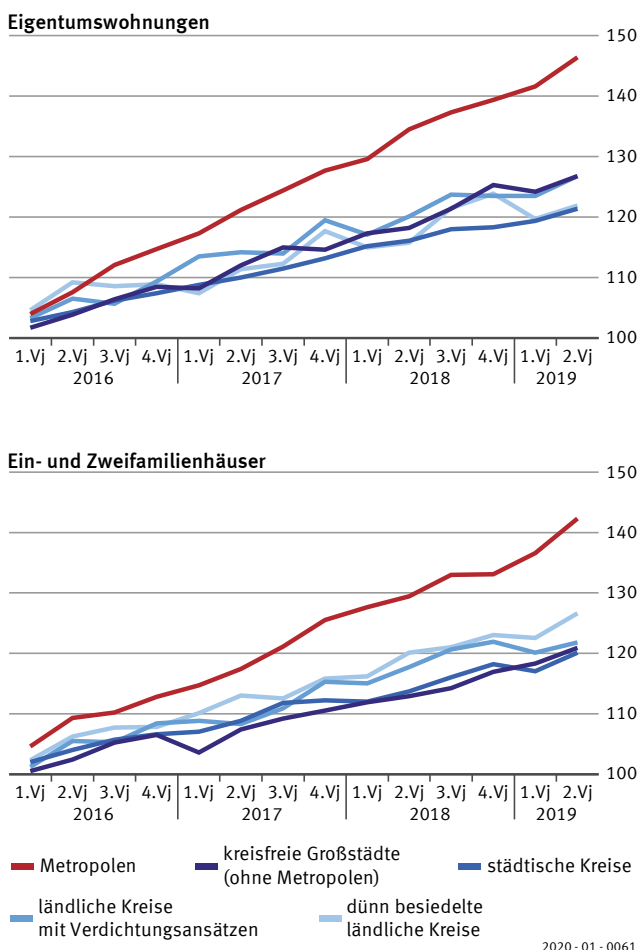
18 Ein Beispiel ist die Anwendung strengerer Grenzwerte bei den Plausibilitätsprüfungen (siehe Abschnitt 3.3).

4

Ergebnisse

In [Grafik 3](#) sind die Indizes für die siedlungsstrukturellen Kreistypen jeweils für Eigentumswohnungen sowie für Ein- und Zweifamilienhäuser dargestellt. In den sieben größten Metropolen stiegen die Preise am stärksten – zwischen 2016 und 2018 für Eigentumswohnungen um 23%, für Ein- und Zweifamilienhäuser um 20%. Über den gesamten verfügbaren Zeitraum, also vom vierten Quartal 2015 bis zum zweiten Quartal 2019, sind die Preise in den Metropolen für Ein- und Zweifamilienhäuser um 42% gestiegen, die für Eigentumswohnungen um 46%. In den übrigen Kreistypen

Grafik 3
Häuserpreisindizes für siedlungsstrukturelle Kreistypen
4. Vj 2015 = 100



– auch in den zwei ländlicheren Kreistypen – stiegen die Preise in diesem Zeitraum ebenfalls deutlich, allerdings weniger stark als in den Metropolen.

5


Fazit und Ausblick

Das Statistische Bundesamt stellt seit Juni 2019 regelmäßig regionalisierte Ergebnisse des Häuserpreisindex auf Ebene der siedlungsstrukturellen Kreistypen des Bundesinstituts für Bau-, Stadt- und Raumforschung bereit. Dadurch kann nun die Preisentwicklung von Wohnimmobilien in ländlichen und städtischen Räumen beobachtet werden. Mit der Berechnung dieser Ergebnisse erfolgten methodische Änderungen, die auch den bundesweiten Häuserpreisindex weiter qualitativ verbesserten. So wird insbesondere die Lage von Immobilien nun besser berücksichtigt und die Plausibilisierung der Daten verfeinert.

Die beschriebenen Arbeiten führten zu einer Verbesserung und Erweiterung der Ergebnisse, allerdings gibt es noch weiteres Entwicklungspotenzial. Die Abbildung der Lagequalität von Immobilien in der Regression zur Berechnung des Häuserpreisindex muss weiter beobachtet werden. Auf internationaler Ebene wird der Einsatz von Geokoordinaten bei der Berechnung von Immobilienpreisindizes analysiert. Auch für den Häuserpreisindex ist wünschenswert, Geokoordinaten in den Erhebungskatalog aufzunehmen und entsprechende Analysen durchzuführen.

Anzustreben wären zudem noch aktuellere Ergebnisse, die auf einer breiten Datenbasis berechnet werden können. Die arbeitsaufwendige Dateneingabe und Recherche bei den Gutachterausschüssen führt dazu, dass regelmäßig Daten für die Berechnung des Häuserpreisindex nach Veröffentlichung der ersten Ergebnisse nachgemeldet werden. Regelmäßige Revisionen ermöglichen es, diese Daten nachträglich zu berücksichtigen.

Die Arbeit der Gutachterausschüsse könnte erleichtert werden, würden Informationen aus den Kaufverträgen von Immobilien in digitaler und strukturierter Form von den Notaren an die Gutachterausschüsse übermittelt.

Dieses Ziel wird aktuell im Rahmen des Projekts „eNoVA“ (Elektronischer Notariat-Verwaltungs-Austausch) angestrebt (Nationaler Normenkontrollrat, 2019). Eine Umsetzung dieser Idee könnte mittelfristig zu schneller verfügbaren, auswertbaren Informationen über den Immobilienmarkt führen und so die Markttransparenz weiter verbessern. 

LITERATURVERZEICHNIS

Buch, Claudia. *Die Rolle des Immobiliensektors für die Finanzstabilität*. Rede beim DVFA Immobilienforum. 2019. [Zugriff am 13. Januar 2020]. Verfügbar unter: www.bundesbank.de

Dechent, Jens. *Häuserpreisindex – Projektfortschritt und erste Ergebnisse für bestehende Wohngebäude*. In: *Wirtschaft und Statistik*. Ausgabe 1/2008, Seite 69 ff.

Dechent, Jens. *Preisindizes für Wohnimmobilien*. In: *Wirtschaft und Statistik*. Ausgabe 11/2011, Seite 1126 ff.

Deutsche Bundesbank. *Die Preise für Wohnimmobilien in Deutschland im Jahr 2018*. In: *Monatsbericht* Februar 2019, Seite 55.

Europäische Kommission (Eurostat). *Technical Manual on Owner-Occupied Housing and House Prices Indices*. Luxemburg 2017. [Zugriff am 13. Januar 2020]. Verfügbar unter: ec.europa.eu/eurostat

Eurostat. *Handbook on Residential Property Prices Indices (RPPIs)*. Luxemburg 2013. [Zugriff am 13. Januar 2020]. Verfügbar unter: ec.europa.eu/eurostat

Hill, Robert J. *Hedonic Price Indexes for Residential Housing: A Survey, Evaluation and Taxonomy*. In: *Journal of Economic Surveys*. 2013. Jahrgang 27. Nr. 5, Seite 879 ff.

Hill, Robert J./Scholz, Michael. *Can geospatial data improve house price indexes? A hedonic imputation approach with splines*. In: *The Review of Income and Wealth*. Jahrgang 64. Ausgabe 4/2018, Seite 737 ff. [Zugriff am 13. Januar 2020]. Verfügbar unter: onlinelibrary.wiley.com

Hill, Robert J./Scholz, Michael/Shimizu, Chihiro/Steurer, Miriam. *An evaluation of the methods used by European countries to compute their official house price indices*. In: *Economie et Statistique/Economics and Statistics*. 2018. Ausgabe 500-501-502, Seite 221 ff. [Zugriff am 13. Januar 2020]. Verfügbar unter: doi.org/10.24187/ecostat.2018.500t.1953

LITERATURVERZEICHNIS

Nationaler Normenkontrollrat. *Weniger Bürokratie, bessere Gesetze*. Jahresbericht 2019. Berlin 2019. [Zugriff am 13. Januar 2020]. Verfügbar unter: www.normenkontrollrat.bund.de

Rousseeuw, Peter J. *Least Median of Squares Regression*. In: Journal of the American Statistical Association. 1984. Jahrgang 79. Nr. 388, Seite 871 ff. [Zugriff am 13. Januar 2020]. Verfügbar unter: www.researchgate.net

Rousseeuw, Peter J./van Driessen, Katrien. *An algorithm for positive-breakdown regression based on concentration steps*. In: Batagelj, Vladimir/Ferligoj, Anuška/Gaul, Wolfgang/Opitz, Otto/Schader, Martin (Herausgeber). *Data Analysis: Scientific Modeling and Practical Application*. New York 2000. Seite 335 ff. [Zugriff am 13. Januar 2020]. Verfügbar unter: www.researchgate.net

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung. *Vor wichtigen wirtschaftspolitischen Weichenstellungen. Jahresgutachten 2018/19*. Wiesbaden 2018. Hier: Kapitel 7 – Keine schnellen Lösungen in der Wohnungspolitik. Abschnitt I Herausforderung Immobilienmarkt, Seite 332 ff.

Silver, Mick. *How to measure hedonic property price indexes better*. In: EURONA – Eurostat Review on National Accounts and Macroeconomic Indicators. Ausgabe 1/2018, Seite 35 ff. [Zugriff am 13. Januar 2020]. Verfügbar unter: ec.europa.eu/eurostat

RECHTSGRUNDLAGEN

Verordnung (EU) Nr. 93/2013 der Kommission vom 1. Februar 2013 zur Festlegung von Durchführungsbestimmungen zu der Verordnung (EG) Nr. 2494/95 des Rates über harmonisierte Verbraucherpreisindizes im Hinblick auf die Schaffung von Preisindizes für selbst genutztes Wohneigentum (Amtsblatt der EU Nr. L 33, Seite 14).

Verordnung (EG) Nr. 2494/95 des Rates vom 23. Oktober 1995 über harmonisierte Verbraucherpreisindizes (Amtsblatt der EG, Nr. L 257, Seite 1).

Verordnung (EU) 2016/792 des Europäischen Parlaments und des Rates vom 11. Mai 2016 über harmonisierte Verbraucherpreisindizes und den Häuserpreisindex sowie zur Aufhebung der Verordnung (EG) Nr. 2494/95 des Rates (Amtsblatt der EU Nr. L 135, Seite 11).

Baugesetzbuch (BauGB) in der Fassung der Bekanntmachung vom 3. November 2017 (BGBl. I Seite 3634).

Richtlinie zur Ermittlung von Bodenrichtwerten (Bodenrichtwertrichtlinie – BRW-RL) vom 11. Januar 2011 (BANz Nr. 24, Seite 597).

Herausgeber

Statistisches Bundesamt (Destatis), Wiesbaden

Schriftleitung

Dr. Daniel Vorgrimler

Redaktionsleitung: Juliane Gude

Redaktion: Ellen Römer

Ihr Kontakt zu uns

www.destatis.de/kontakt

Erscheinungsfolge

zweimonatlich, erschienen im Februar 2020

Das Archiv älterer Ausgaben finden Sie unter www.destatis.de

Artikelnummer: 1010200-20001-4, ISSN 1619-2907

© Statistisches Bundesamt (Destatis), 2020

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.