# Ein Warenkorb für den regionalen Preisvergleich

Dr. Sara Bleninger, M.Sc., Gabriel König, M.Sc.,
Dipl.Soz.Univ. Dr. Christian Baier, Dipl.Kfm.Univ. Markus König

Die Sorge für die Gleichwertigkeit der Lebensverhältnisse ist ein verfassungsmäßiger Auftrag des Staates. Um diesen erfüllen zu können, sind belastbare Daten nötig, die einerseits die Haushaltseinkommen, andererseits die Preise für Waren und Dienstleistungen in zeitlicher und räumlicher Hinsicht differenziert abbilden. Während die amtliche Statistik umfassende Daten zur regionalen Verteilung der Einkommen zur Verfügung stellt, fehlen jedoch vergleichbare Daten über regionale Lebenshaltungskosten. Abhilfe schaffen würde hier ein regionaler Preisindex, anhand dessen man die Preise für Waren und Dienstleistungen zwischen Regionen systematisch vergleichen könnte. Aufgrund des erheblichen Aufwands bei der Erhebung von Verbraucherpreisen ist eine kontinuierliche Berechnung regionaler Preisvergleiche nur anhand bereits vorhandener Daten – etwa aus der Verbraucherpreisstatistik – denkbar. Der Beitrag stellt zwei konkurrierende Modellansätze zur Definition regionaler Preisvergleiche mittels dieser Daten vor, gibt einen Einblick in deren theoretische Grundlagen und zeigt die jeweiligen Vorzüge und Schwierigkeiten auf.

## Regionaler Preisvergleich

Die Gleichwertigkeit der Lebensverhältnisse hat in Bayern Verfassungsrang. So verpflichtet der 2014 neu eingeführte Artikel 3 Absatz 2 Satz 2 der Bayerischen Verfassung den Staat, „gleichwertige Lebensverhältnisse und Arbeitsbedingungen in ganz Bayern, in Stadt und Land“ zu fördern und zu sichern. Ganz offensichtlich werden die Lebensverhältnisse unter anderem vom verfügbaren Einkommen der Menschen bestimmt. Höhe, Entwicklung und regionale Verteilung der verfügbaren Einkommen der privaten Haushalte werden vom Arbeitskreis „Volkswirtschaftliche Gesamtrechnungen der Länder“ regelmäßig berechnet. Die allgemeine Lebenserfahrung zeigt aber auch, dass sich nicht nur das Einkommen, sondern auch die Lebenshaltungskosten – oder anders ausgedrückt: die Kaufkraft dieser Einkommen – regional stark unterscheiden. An dieser Stelle verfügt die amtliche Statistik aktuell nur über unzureichende Informationen. Die Beurteilung der Gleichwertigkeit der Lebensverhältnisse wird dadurch deutlich erschwert. Abhilfe schaffen würde hier ein regionaler Preisindex, mit dessen Hilfe sich die Preise für Waren und Dienstleistungen zwischen Regionen – etwa Landkreisen – und somit die Lebenshaltungskosten systematisch vergleichen ließen. Das Bayerische Landesamt für Statistik hat das Ziel, einen derartigen Index zu konzipieren und – sofern er den Qualitätsansprüchen der amtlichen Statistik genügt – regelmäßig zu veröffentlichen. Dieser Beitrag gibt einen ersten Einblick in die theoretischen Grundlagen und zeigt die Schwierigkeiten auf, die mit den unterschiedlichen, konkurrierenden Modellansätzen verbunden sind.

Regionale Preisindizes werden von Wissenschaft, Verwaltung und Wirtschaft für vielerlei Zwecke benötigt: So brauchen etwa Forscherinnen und Forscher, die sich mit Einkommen, Löhnen oder der Wohlfahrtsmessung beschäftigen, regional

tief gegliederte Informationen zum Preisniveau, um die nominalen Werte über eine Preisbereinigung in reale Kennzahlen umzurechnen. Verwaltungsorgane sind auf verlässliche Informationen über regionale Preisunterschiede angewiesen, um beispielsweise die angemessene Höhe von Fördergeldern bestimmen zu können, und Unternehmen benötigen regionale Preisindizes, um ihre Lohnstruktur an regionale Disparitäten anzupassen.

Der großen Nachfrage nach regionalen Informationen steht jedoch ein sehr begrenztes und veraltetes Angebot gegenüber. Für Deutschland gehen die letzten Untersuchungen auf Rostin (1979) und Ströhl (1994) zurück. Kawka (2009) bemüht sich um eine umfangreiche Datensammlung zum Thema regionaler Preisindex. Diese Datenerhebung fand aber über einen relativ langen Zeitraum statt und folgt eher dem Prinzip „was zu finden war", wobei methodische Ansprüche vernachlässigt wurden. Für Bayern gibt es eine Studie aus dem Jahr 2002, veröffentlicht vom Bayerischen Staatsministerium für Wirtschaft, Verkehr und Technologie (2003), welche jedoch durch von der Lippe und Breuer (2009) in Bezug auf die Methodik stark kritisiert wurde.

Der überwiegende Teil der genannten Studien eignet sich – aufgrund des damit verbundenen erheblichen Aufwands bei der Datenerhebung – nicht für eine regelmäßige Wiederholung. Die einzige Ausnahme stellt die Analyse für Bayern aus dem Jahr 2002 dar, diese leidet allerdings unter methodischen Problemen.

Im Folgenden soll es darum gehen, wie regionale Preisvergleiche mittels regelmäßig erhobener Daten konkret berechnet werden können. Dabei wird untersucht, wie man eine für den regionalen Preisvergleich geeignete Auswahl an Waren und Dienstleistungen bei einem gegebenen Datensatz zusammenstellen kann. Des Weiteren wird der grundlegenden Frage nachgegangen, welche Anforderungen an den Datensatz gestellt werden müssen, um regionale Preisvergleiche sinnvoll berechnen zu können.

## Daten für den regionalen Preisvergleich

Wie bereits weiter oben beschrieben, lässt sich die Nachfrage nach einem regelmäßig aktualisierten regionalen Preisindex durch die zitierten Studien nicht befriedigen. Eine kontinuierliche Berechnung regionaler Preisvergleiche ist aufgrund des Aufwands, der mit der Erhebung von Preisdaten verbunden ist, nur denkbar, wenn dazu bereits vorhandene Daten verwendet werden. Infrage kommen hierfür die monatlich erhobenen Daten der Verbraucherpreisstatistik.

Diese Daten werden jedoch für einen zeitlichen, nicht für einen räumlichen Preisvergleich genutzt. Es werden dabei Preise in unterschiedlichen Regionen erhoben, allerdings nicht mit dem Ziel, einen repräsentativen Querschnitt darzustellen. Die so generierten Daten werden zur Berechnung des Verbraucherpreisindex (VPI) zusammengefasst, ein Vergleich erfolgt dann über die einzelnen Erhebungszeitpunkte. Bei regionalen Preisvergleichen sollen die zu vergleichenden Preise einen reinen Preisunterschied in räumlicher Hinsicht messen. Qualitätsunterschiede dürfen keine Rolle spielen, das heißt, Differenzen zwischen den Preisen der jeweiligen zu vergleichenden Produkte sollen dabei explizit nicht durch eine unterschiedliche Produktqualität zustande kommen. Darüber hinaus muss der Warenkorb, der dem zu berechnenden Preisindex zugrunde liegt, für einen regionalen Preisniveauvergleich geeignet sein.

Ein grundlegender Unterschied zwischen der räumlichen und der zeitlichen Betrachtungsweise besteht darin, dass Regionen – anders als Zeiträume – keiner natürlichen Reihenfolge oder Ordnung unterliegen. Darüber hinaus sind Regionen teilweise sehr unterschiedlich groß und können sich auch hinsichtlich der dort konsumierten Güter deutlich voneinander unterscheiden, sodass Preise häufig viel stärker zwischen Regionen variieren als im Zeitverlauf (Greuel 2014). Im zeitlichen Verlauf werden die betrachteten Güter konstant gehalten, sodass immer wieder der Preisvergleich der exakt gleichen Güter abgebildet wird. Dies kann ein räumlicher Preisvergleich nur schwer leisten. Angesichts dieser unterschiedlichen Anforderungen an die Datengrundlage stellt sich die Frage,

ob sich die für die Verbraucherpreisstatistik erhobenen Preise auch für einen räumlichen Preisvergleich verwenden lassen.

### Der Warenkorb

In der Verbraucherpreisstatistik werden Preisänderungen bestimmt, indem die Preisänderung eines bestimmten Warenkorbs gemessen wird. Die Zusammensetzung des Warenkorbs soll die typischen Verbrauchsgewohnheiten eines deutschen Haushalts widerspiegeln.

Die genaue Zusammensetzung des Warenkorbs wird über das Wägungsschema bestimmt. Es besteht aus den Gewichten, mit denen die einzelnen Waren und Dienstleistungen in den Warenkorb eingehen. Dabei sind Warenkorb und Gewichte für einen zeitlichen Vergleich gestaltet. Das heißt, die Güter, deren Preise gemessen werden, sind eng definiert und werden bis zur alle fünf Jahre erfolgenden Revision konstant gehalten. Die räumliche Verteilung der Preisbeobachtungen spielt eine untergeordnete Rolle. Im Folgenden wird der Ausdruck „Produkte“ synonym für den Begriff Güter verwendet, der als Oberbegriff die im Warenkorb enthaltenen Waren und Dienstleistungen zusammenfasst.

**Classification of Individual Consumption by Purpose (COICOP)**

Die im Warenkorb enthaltenen Waren und Dienstleistungen werden über das System COICOP[1] klassifiziert (United Nations: Department of Economic and Social Affairs, Statistics Division 2018). Diese Klassifikation der Vereinten Nationen wurde zur Analyse von Konsumausgaben nach Verwendungsart definiert. Sie ermöglicht die Zuordnung der jeweiligen Preise zu tief gegliederten Kategorien über eine 10-stellige Nummer (COICOP-10-Steller). So umfasst beispielsweise der COICOP 03 die Kategorie „Bekleidung und Schuhe“ mit den Unterkategorien 032 für „Schuhe“, 03212 für „Schuhe für Damen“ und in tiefster Gliederungsebene 0321250200 für „Damensportschuhe“.

1 Classification of Individual Consumption by Purpose – Klassifikation der Verwendungszwecke des Individualverbrauchs, vgl. United Nations Statistics Division (2018)

Die Frage nach der Eignung der Daten aus der Verbraucherpreisstatistik für einen regionalen Preisvergleich impliziert damit auch die Frage, ob sich daraus ein Warenkorb konstruieren lässt, der für regionale Preisvergleiche geeignet ist. Denn auch bezüglich des Warenkorbs stellen zeitliche und räumliche Preisvergleiche unterschiedliche Anforderungen an die Daten. Während es für zeitliche Preisvergleiche sinnvoll ist, immer wieder die Preise der gleichen, eng definierten Produkte zu erheben, sodass alle Veränderungen als reine Preisunterschiede über einen Zeitraum interpretiert werden können, ist für regionale Preisvergleiche eine möglichst breite regionale Streuung der Preisbeobachtungen entscheidend. Wenn der Preis eines Produktes in möglichst vielen – idealerweise in allen – zu vergleichenden Regionen beobachtet wurde, kann auf dieser Datenbasis ein Preisindex berechnet werden, der für die vielen möglichen Vergleiche im regionalen Kontext sinnvolle Aussagen macht. Um dies zu erreichen, kann es sinnvoll sein, mit gröberen Produktdefinitionen zu arbeiten, damit möglichst viele Regionen anhand derselben Produktkategorie verglichen werden können.

Konkret stellt sich damit auch die Frage, ob und wie sich ein solcher Warenkorb für regionale Preisvergleiche aus den Daten, welche die amtliche Statistik zur Berechnung des VPI regulär erhebt, konstruieren lässt. Dabei ist die Entscheidung zu treffen, ob einzelne Produkte oder Klassen von ausreichend homogenen Produktkategorien verglichen werden sollen.

Im nachfolgenden Kapitel wird zunächst der Ansatz von Weinand und von Auer (2019) zur Berechnung regionaler Preisindizes vorgestellt, welcher eine absolute Vergleichbarkeit der Produkte für den regionalen Preisvergleich fordert. Nachdem beleuchtet wurde, wie gut dieser Ansatz für die Berechnung regionaler Preisvergleiche mittels der VPI-Daten geeignet ist, wird ein Alternativansatz vorgestellt, der weniger strenge Bedingungen an die Vergleichbarkeit der Produkte stellt. Danach

werden die zugrunde liegenden Daten beschrieben und zuletzt die beiden Ansätze hinsichtlich der vorliegenden Daten der Verbraucherpreisstatistik für Bayern miteinander verglichen.

### Perfect Matches Only (PMO)

Auf der Grundlage bestehender Forschung zu internationalen Preisvergleichen entwickeln Weinand und von Auer (2019) das Argument, dass eine Gruppierung von Produkten in sogenannte homogene Klassen zu „verunreinigten" Preisvergleichen führen würde. Eine homogene Produktklasse liegt vor, wenn sich die Produkte in dieser Klasse nicht allzu sehr voneinander unterscheiden, also in gewisser Weise eine homogene, das heißt gleichartige Gruppe bilden. Preisvergleiche sind nach Auffassung von Weinand und von Auer (2019) bei der Vorgehensweise nach homogenen Produktgruppen nicht mehr als „rein" zu betrachten, sondern basieren zum Teil auf der unterschiedlichen Qualität der Produkte, die zu einer Klasse zusammengefasst werden. Weinand und von Auer (2019) stellen daher das Postulat auf, „nur exakt Gleiches mit Gleichem" zu vergleichen und bezeichnen diesen Ansatz als Perfect Matches Only (PMO). Beispielsweise könnte man eine homogene Klasse „Eiernudeln" definieren. Bei einem exakten Vergleich könnte man hingegen nur „Eierspiralnudeln der Marke A mit der Packungsgröße 500 g" miteinander vergleichen.

Der Grundsatz des PMO-Ansatzes besteht also darin, regionale Preisvergleiche nur anhand exakt gleicher Produkte durchzuführen – mit dem Ziel, Verzerrungen durch Qualitätsunterschiede zwischen den verglichenen Produkten auszuschließen. In Bezug auf die oben aufgeworfene Frage nach dem geeigneten Warenkorb für regionale Preisvergleiche bedeutet dies, dass der Warenkorb sich – ähnlich wie beim zeitlichen Preisvergleich – aus möglichst eng definierten Produkten zusammensetzen soll, die über Vergleichseinheiten hinweg konstant sind.

### PMO als reiner Preisvergleich für die Berechnung des Preisindex

Zunächst gilt es nun, den Begriff „reiner Preisvergleich" zu klären, bevor der Zusammenhang zum PMO-Ansatz verdeutlicht wird. Basis hierfür sind die Grundlagen der Preisstatistik.

Ziel der klassischen Preisstatistik ist es, Preisveränderungen über die Zeit hinweg zu messen. Zum Zweck der Messung der zeitlichen Veränderungen werden Preisindizes berechnet. Das Grundgerüst aller Preisindizes ist dabei ein Quotient, der eine mit Mengen gewichtete Summe von Preisen (im Zähler) ins Verhältnis zu den jeweiligen Preisen einer Bezugsperiode (im Nenner) setzt. Beispielsweise wird in der deutschen Preisstatistik der Laspeyres Preisindex verwendet:

$$P_L^{t_0,t} = \frac{\sum_{i=1}^{n} p_i^t q_i^{t_0}}{\sum_{i=1}^{n} p_i^{t_0} q_i^{t_0}} = \sum_{i=1}^{n} \frac{p_i^t}{p_i^{t_0}} g_i^{t_0} \quad \text{Gleichung (1)}$$

mit $p_i^t$ als $i$-ter Preis (Preis des Produktes i) aus der Berichtsperiode $t$ beziehungsweise mit $p_i^{t_0}$ aus der Bezugsperiode $t_0$ und den Mengen (im Warenkorb) $q_i^{t_0}$ beziehungsweise den Gewichten $g_i^{t_0}$, jeweils aus der Bezugsperiode $t_0$. In der ersten Formulierung wird die Bepreisung eines Warenkorbs mit entsprechenden Mengen aus der Bezugsperiode deutlich. In der zweiten wird deutlich, dass es sich um den gewichteten Mittelwert von einzelnen Preisverhältnissen handelt (Vogel 2000).

Die Preisindexberechnung verfolgt das Ziel, Produkte und Wägungsanteile konstant zu halten. Die grundlegende Idee besteht darin, dass sich Qualitätsunterschiede bei exakt gleichen Produkten im Zähler und Nenner der Gleichung (1) „herauskürzen", so dass nur noch der reine Preisunterschied übrigbleibt. Der Preisindex soll damit die „reine" Preisänderung – also nominale Preisänderung ohne Qualitätsänderungen – abbilden (Reich 2000; Guckes 1979). Nach von der Lippe (1977) sollen Mittelwerte von Einzelpreisen gebildet werden, um diese miteinander zu vergleichen, da „auch für vergleichbare (gleichartige, nicht gleichwertige) Güter und Dienstleistungen sehr unterschiedliche Preise vorkommen" (von der Lippe 1977).

Das Kriterium „reiner Preisvergleich" zielt also darauf ab, Einflüsse, die den Vergleich stören, rechnerisch zu eliminieren (von der Lippe 2000). In

einer strengen Auslegung ist dies nur erreichbar, wenn stets die Preise von exakt gleichen Produkten verglichen werden. Es wird daher in der Datenerhebung für den VPI stets darauf geachtet, sämtliche Parameter, wie Berichtsstellen, Produkte und Dienstleistungen, Verpackungsgrößen etc., exakt über die Zeit hinweg konstant zu halten. Echte Änderungen soll es nur im Rahmen der alle fünf Jahre stattfindenden Revision geben, bei der auch Warenkorb und Gewichte aktualisiert werden.

Will man aber statt eines zeitlichen Preisvergleichs wie beim VPI räumliche Preisvergleiche anstellen, stößt man auf das Problem, dass die Produkte teilweise nicht vergleichbar sind, da regional unterschiedliche Produkte erhoben wurden. Dies wird bei Dienstleistungen besonders deutlich: Während es durchaus möglich ist, Monat für Monat beim selben Friseurgeschäft den Preis für einen Männerhaarschnitt zu erfragen, müssen beim regionalen Preisvergleich gezwungenermaßen verschiedene Friseurgeschäfte befragt werden, die in den jeweiligen Regionen liegen – und es stellt sich die Frage, inwiefern es sich dabei um die gleiche Dienstleistung handelt.

Dieses Problems nehmen sich Weinand und von Auer (2019) an, indem sie genau die für zeitliche Preisindizes etablierte Bedingung auch an den räumlichen Preisvergleich stellen: Nur exakt gleiche Produkte sollen verglichen werden. Der PMO-Ansatz legt die Idee des reinen Preisvergleichs also eng aus. Dies kann jedoch zu Problemen führen, wenn versucht wird, die Daten der Verbraucherpreisstatistik auf diese Weise für einen regionalen Preisindex zu nutzen: Viele Produkte werden nur in wenigen Regionen erhoben, so dass mit dem PMO-Ansatz keine ausreichende Anzahl an Vergleichsobjekten zur Verfügung steht (im Extremfall wurde das Produkt nur in einer Region erhoben und kann daher für die Analyse nicht genutzt werden). Außerdem gibt es mit dem PMO-Ansatz höchstwahrscheinlich kein Produkt, dessen Preis in allen Kreisen erhoben wurde. Somit lassen sich viele Beobachtungen nicht nutzen, was die Güte eines daraus berechneten regionalen Preisvergleichs mindert.

### PMO und CPD-Schätzung

Weinand und von Auer (2019) schlagen vor, den PMO-Ansatz in Verbindung mit einer Country-Product-Dummy (CPD)-Methode nach Silver (2009) zu schätzen. In der CPD-Methode werden sogenannte Dummy-Variablen definiert. Diese Dummy-Variablen können entweder den Wert 1 für „trifft zu“ oder den Wert 0 für „trifft nicht zu“ annehmen, wodurch implizit eine Gruppierung der Daten stattfindet. Mithilfe dieser Methodik wird der Einfluss einer Gruppierung kontrolliert, ganz konkret kann so – im Fall der räumlichen Gruppierung beziehungsweise der Dummy-Variable für Regionen – der regionale Unterschied berechnet werden. Das Zusammenspiel von PMO-Ansatz und CPD-Schätzung ist für die folgende Analyse von grundsätzlicher Bedeutung. Hierzu wird nachstehend kurz der CPD-Ansatz, wie ihn Weinand und von Auer (2019) anwenden, erklärt.

**Übersicht 1: Geschäftstypen**
In der Verbraucherpreisstatistik werden verschiedene Arten von Geschäften unterschieden.

| Schlüssel | Geschäftstyp | Charakterisierung |
|---|---|---|
| 1 | Warenhaus, Kaufhaus | Breites Sortiment; großflächig; Innenstadt oder Einkaufszentrum |
| 2 | SB-Warenhaus, Verbrauchermarkt | Breites Lebensmittelsortiment, auch Gebrauchsgüter; großflächig; Autokunden-orientiert |
| 3 | Supermarkt | Lebensmittel und Waren des täglichen Bedarfs; oft in Wohngebieten |
| 4 | Fachmarkt, Discounter | Konzentration auf eine Branche; Niedrigpreise; häufig Autokunden-orientiert |
| 5 | Fachgeschäft | Konzentration auf eine Branche; Auswahl, Qualität, Beratung, Service; oft Innenstadt oder Einkaufszentrum |
| 6 | Sonstiger Einzelhandel | Tankstellenshops, Kioske, Marktstände etc. |
| 7 | Öffentlicher oder privater Dienstleistungsbetrieb | Werkstätten, Friseursalons, Restaurants etc. |
| 8 | Versandhandel, Internethandel | Bestellung (z. B. im Internet/telefonisch) und Lieferung der Ware |

Weinand und von Auer (2019) verwenden, ähnlich dem Vorgehen bei der Berechnung des deutschen VPI, ein stufenweises Verfahren. Auf jeder Stufe findet dabei eine Gruppierung der Daten, eine Definition der entsprechenden Dummy-Variablen und anschließend eine Aggregation der Ergebnisse statt. Die CPD-Methodik wird daher mehrmals aufeinander aufbauend angewendet. Auf diese Weise werden die Einzeldaten sukzessive als Indizes gewichtet zusammengefasst. Die regionalen Preisniveaus werden dabei mithilfe eines Fixed-Effects-Modells als lineare Regression geschätzt. Der Preis $p_i^r$ des Produktes *i* in der Region *r* wird dabei durch das regionale Preisniveau $P^r$ in der Region *r* und dem produktspezifischen Preisniveau $\pi_i$ des Produktes *i* erklärt. Das lineare Regressionsmodell ist für die unterste Ebene wie folgt definiert:

$$\ln p_i^r = \sum_{s=2}^{R} (\ln P^s)\, region^s + \sum_{j=1}^{N} (\ln \pi_j)\, object_j + \ln \varepsilon_i^r$$

Gleichung (2)

mit den Dummy-Variablen

$$region^s = \begin{cases} 0, & falls\ r \neq s \\ 1, & falls\ r = s \end{cases}$$

für die Regionen und

$$object_j = \begin{cases} 0, & falls\ i \neq j \\ 1, & falls\ i = j \end{cases}$$

für die Produkte sowie dem regionalen Preisniveau $P^s$, dem produktspezifischen Preisniveau $\pi_j$ und einem Fehlerterm $\varepsilon_i^r$. Die Schätzung des regionalen Preisniveaus $P^s$ kann unabhängig vom produktspezifischen Preisniveau $\pi_j$ erfolgen, da die Einflüsse der in der Region s beobachteten Produkte herausgerechnet werden.

Die Daten für die Schätzung auf der ersten Stufe werden nach Güterart (in Form der COICOP-10-Steller) und Geschäftstyp (siehe Übersicht 1) gruppiert. Die Regression erfolgt dann auf dieser sogenannten Elementarindexebene für jede mögliche Kombination aus Güterart und Geschäftstyp separat. Wie aus Gleichung (2) ersichtlich, entsprechen dann die Koeffizienten der Regression den logarithmierten regionalen Preisniveaus $P^s$ sowie den logarithmierten produktspezifischen Preisniveaus $\pi_j$, wobei letztere für die weitere Betrachtung nicht von Interesse sind. Die Zusammenfassung auf der Ebene der COICOP-10-Steller und Geschäftstypen kommt dabei einer Mittelung über die berücksichtigten Preise auf der entsprechenden Ebene gleich. Im Endeffekt schätzen Weinand und von Auer (2019) in der ersten Stufe die mittleren regionalen Preisniveaus (je COICOP-10-Steller und Geschäftstyp) für die Regionen.

Der reine Preisvergleich kann an dieser Stelle nur auf zwei Arten erhalten bleiben: entweder mittels Schätzung auf Ebene der Produkte, die nach dem PMO-Ansatz verschieden sind, wobei jedes Produkt in (möglichst vielen) unterschiedlichen Regionen beobachtet worden sein muss. Oder aber in allen Regionen sind exakt die gleichen Produkte für die Schätzung vorhanden. Aufgrund der regional betrachtet spärlichen Besetzung der Preise ist mit den Daten des VPI beides nicht möglich. Für eine einzelne Schätzung nach Produkt liegen zu wenige Beobachtungen je Produkt vor. Für einen exakt gleichen Warenkorb je Region wurden zu wenige Produkte in allen Regionen beobachtet. In diesem Fall wäre der Warenkorb nicht mehr aussagekräftig. Deswegen ist es notwendig, die Preisbeobachtungen in Gruppen zusammenzufassen und entsprechend nur nach COICOP-10-Steller und nach Geschäftstyp zu kontrollieren. Durch diese Zusammenfassung nach Produktgruppen werden damit mehr beobachtete Datenpunkte pro Preisvektor erzeugt.

Weinand und von Auer (2019) verwenden an dieser Stelle nur Preise der Produkte, die nach dem PMO-Ansatz infrage kommen, um einen reinen Preisvergleich zu sichern. Zwar werden durch den PMO-Ansatz zunächst nur Produkte selektiert, die einen exakten Vergleichspunkt in mindestens einer anderen Region haben, sie werden aber aufgrund der großen Anzahl fehlender Werte im Zuge der Regressionsschätzung wieder zusammengefasst. Das Ziel, nur Gleiches mit Gleichem zu vergleichen, kann somit nicht durchgehend umgesetzt werden. Letztendlich führt der PMO-Ansatz dazu, dass eine große Menge der Daten, die mit einer gröberen Definition der Produkte beziehungsweise Produktgruppen in die Schätzung einfließen würden, nicht nutzbar ist, wie im Folgenden noch zu sehen sein wird.

### Homogene Güter

Prinzipiell verfolgt der reine Preisvergleich folgendes Ziel: Es sollen den Vergleich störende Faktoren – insbesondere Qualitätsunterschiede – eliminiert werden. Dieser Anspruch führt jedoch nicht zwangsläufig zu der Interpretation, dass nur exakt gleiche Produkte verglichen werden dürfen: Man könnte auch fordern, dass nur Vergleichbares verglichen werden soll. Dies führt zu der Überlegung, homogene Güter zu erzeugen, die weniger restriktiv definiert sind als die Güter im PMO-Ansatz. Ein homogenes Gut ist eine Klasse von Produkten verschiedener Marken und Hersteller, welche in sich hinreichend homogen ist, also vergleichbare Produkte enthält. Beim Ansatz Homogene Güter geht es folglich darum, geeignete Produktkategorien zu konstruieren, aus denen sich ein Warenkorb für regionale Preisvergleiche zusammensetzt.

#### Der Preisindex als Schätzer

Um ein handfestes Konzept für das Postulat der Homogenität zu entwickeln, nehmen wir Abstand von der bisher verfolgten Idee des akkuraten Messens und wenden uns der statistischen Ansicht einer Wahrscheinlichkeitsverteilung zu, mit deren Hilfe unverzerrt geschätzt wird. So lässt sich der reine Preisvergleich als entsprechende Eigenschaften einer Wahrscheinlichkeitsverteilung deuten. Es sollen Produkte miteinander verglichen werden, deren Preise der gleichen Verteilung folgen. In diesem Sinne sind die beobachteten Preise Stichprobenrealisationen und der berechnete Preisindex ein Schätzer.

Diese Auffassung folgt auch der Interpretation des reinen Preisvergleichs mit dem Ziel, störende Einflüsse zu minimieren (von der Lippe 2000). Das Konzept kommt dabei auch im ersten Teil der Gleichung (1) zum Ausdruck. Der Preisindex stellt einen Quotienten gewichteter Mittelwerte dar. Die Gewichte für Zähler und Nenner stammen aus der Bezugsperiode $t_0$, während die Preise im Zähler aus der Berichtsperiode $t$ und die im Nenner aus der Bezugsperiode $t_0$ stammen. So können die Preismessungen als Stichprobenrealisationen von stochastischen Größen mit einer gewissen Variabilität interpretiert werden, welche durch die Varianz zum Ausdruck kommt. Die Wahl des entsprechenden Produkts ist über die Stichprobenziehung definiert und damit ebenfalls eine stochastische Größe. Dies ermöglicht es, die Auswahl eines Produkts als Ausprägung einer Klasse von relativ homogenen – aber nicht exakt gleichen – Produkten zu betrachten, deren Preise einer gemeinsamen Verteilung folgen. Eine mögliche systematische Verzerrung, also Bias, wird im vorliegenden Fall durch das Stichprobenverfahren, die Bandbreite an beobachteten Produkten und auch durch die Menge an Beobachtungen abgewendet.

Somit entspricht unsere Definition der homogenen Güter auch der Art und Weise, wie die einzelnen Produkte in die Stichprobe gelangen: Die Preiserheberinnen und Preiserheber bekommen eine Produktbeschreibung und wählen selbst das spezifische Produkt vor Ort aus. Die Beschreibung der Produkte entwächst der Idee, homogene Produkte zu erheben. Das spezifische Produkt ist lediglich eine Realisation einer homogenen Produktgruppe.

Um das an einem Beispiel zu verdeutlichen: Der PMO-Ansatz fordert, dass nur die exakt gleichen Produkte (gleicher Name, gleiche Marke, gleiche Packungsgröße usw.) miteinander verglichen werden. Beispielsweise sollte nur die Sorte A der Marke B in Verpackung C mit genau dem gleichen Produkt ABC verglichen werden. Mit dem Ansatz Homogene Güter wird diese Idee aufgeweicht. Hier könnte man Produktgruppen miteinander vergleichen, etwa die Eigenmarke von Discounter X mit der Eigenmarke von Discounter Z.

#### Ebene der Elementarindizes

Es erscheint sinnvoll, sich bei der Entscheidung über die Detailtiefe, auf der die Vergleichbarkeit der Produkte betrachtet werden soll, an der aktuellen Berechnung des VPI zu orientieren. Auf der Ebene der sogenannten Elementarindizes – also gruppiert nach COICOP-10-Steller und nach Geschäftstyp – werden bei der VPI-Berechnung in einem ersten Schritt gewichtete Mittelwerte gebildet (Statistisches Bundesamt 2015). Diese Mittelwertberechnung des VPI wird durch Weinand und von Auer (2019) in der ersten Stufe der CPD-Schätzungen imitiert, indem auf dieser Ebene

einzelne Gleichungen geschätzt werden. Auf diese Weise wird ein regionales Preisniveau als Mittelwert unter Berücksichtigung der mittleren Preisniveaus je COICOP-10-Steller und Geschäftstyp geschätzt. Da die Schätzung letztendlich auf dieser Zusammenfassung der Produkte beruht, reicht es auch aus, die Produkte als homogene Güter entsprechend zu gruppieren. Eine spezifischere Betrachtungsweise – wie sie der PMO-Ansatz fordert – lohnt sich an dieser Stelle nicht, weil durch die CPD-Schätzung nur die mittleren Preisunterschiede der jeweils in den Regionen erhobenen Produkte berücksichtigt und in die weiteren Berechnungsstufen transportiert werden.

### Qualitätsunterschiede und Variabilität

Die Definition der homogenen Güter führt zu der Frage, wie die Vergleichbarkeit beziehungsweise Homogenität der Produkte in den jeweiligen Produktkategorien sichergestellt und nachgewiesen werden kann. Einen Ansatz für eine ähnlich gelagerte Problematik liefert eine Studie von Götz et al. (2009). In der Analyse regionaler Preisunterschiede in Österreich befassen sich die Autoren dabei mit der Zusammenstellung sinnvoller Teilwarenkörbe als Teilmenge des gesamten Warenkorbs und der Berechnung belastbarer Ergebnisse mit Hilfe eines Teilwarenkorbs. Als Kriterien zur Konstruktion des Teilwarenkorbs wird dabei die Variabilität zwischen den Warenkörben und die Variabilität innerhalb eines Warenkorbes genutzt.

Mittels dieser Idee lässt sich eine Methodik zur Selektion ausreichend homogener Güter ableiten. Betrachtet man die erhobenen Preise gruppiert auf Ebene der Elementarindizes (Gruppen nach COICOP-10-Steller und Geschäftstyp), lassen sich zwei Arten von Varianz berechnen: eine externe Varianz, also die Varianz zwischen den Mittelwerten der einzelnen Regionen, und eine interne Varianz, also die (durchschnittliche) Varianz innerhalb der Regionen. Um Skaleneffekte auszuschließen, werden hier anstelle der Varianzen externe und interne Variationskoeffizienten betrachtet, definiert als:

$$VarK(x) = \frac{\sqrt{Var(x)}}{E(x)}.$$

Homogene Güter müssten eine ähnliche externe und interne Variabilität haben, weil sie innerhalb der Region und zwischen den Regionen ähnlich streuen. Nicht ausreichend homogene Güter müssten intraregional (innerhalb der Region) stärker streuen als interregional (zwischen den Regionen). Es bietet sich an, den Quotienten aus beiden Maßen zu betrachten: Ist der Quotient nahe 1 beziehungsweise kleiner 1, ist das betrachtete Gut hinreichend homogen. Ist der Quotient deutlich größer als 1, so ist die intraregionale Streuung größer als die interregionale und man kann nicht mehr von einem homogenen Gut sprechen (Bleninger 2017).

### Vorliegende Daten

Nachfolgend wird kurz der verwendete Datensatz vorgestellt. Es werden die Einzeldaten der Verbraucherpreisstatistik, die über die Forschungsdatenzentren (FDZ) der Statistischen Ämter des Bundes und der Länder zur Verfügung gestellt wurden, für das Jahr 2019 betrachtet. Die Daten selbst werden als dezentrale und zum Teil als zentrale Erhebungen auf Basis des Gesetzes über die Preisstatistik (PreisStatG) in der im Bundesgesetzblatt Teil III, Gliederungsnummer 720-9, veröffentlichten bereinigten Fassung, zuletzt geändert durch Artikel 1 des Gesetzes vom 10. Dezember 2019 (BGBl. I S. 2117), erhoben. Erhebungseinheiten sind repräsentativ ausgewählte Einzelhandels- und Dienstleistungseinheiten. Mit den dort ermittelten Preisen wird der VPI für die Länder und für Deutschland monatlich berechnet (Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder 2021).

Um den VPI zu berechnen, werden in Bayern monatlich mehr als 70 000 Preise und in ganz Deutschland mehr als 300 000 Preise erhoben. Hierzu zählen neben Produkten aus dem Lebensmitteleinzelhandel auch Dienstleistungen oder langlebige Konsumgüter. Für die hier vorgenommene Betrachtung reicht es aus, die Daten eines Monats genauer zu untersuchen. Dabei ist darauf zu achten, dass ein Monat ohne saisonale Besonderheiten gewählt wird. So erscheint zum Beispiel der Monat Mai – im Gegensatz zum Dezember, der vom Weihnachtsgeschäft geprägt ist –

repräsentativ, da es in diesem Monat keine besonderen preisbildenden Vorkommnisse gibt.

Insgesamt besteht der aus dem Jahr 2019 vorliegende Datensatz aus 8 691 385 Preisreihen, von denen 868 132 in Bayern erhoben werden. Auf den Monat Mai entfallen davon insgesamt 70 943 Preisreihen. Aus diesen Daten werden weitere Beobachtungen mit fehlerhaften COICOP-Nummern gelöscht, welche wahrscheinlich durch Schreibfehler verursacht wurden, sodass sich die Beobachtungen auf 70 904 Einzelpreise reduzieren. Weiterhin werden Preisreihen mit einem Wägungsanteil von 0 selektiert (diese Beobachtungen sind für die aktuelle VPI Berechnung nicht relevant), was eine weitere Reduktion auf 50 528 Beobachtungen darstellt. Ebenfalls werden Zeilen des Datensatzes gelöscht, welche einen Preis von 0 aufweisen (dies entspricht einem fehlenden Preis), sodass 45 864 Preisbeobachtungen verbleiben. Weiterhin werden zentral erhobene Preisdaten gelöscht, da diese stellvertretend für alle Regionen und Länder erhoben werden, damit über alle Regionen hinweg gleich sind und somit für einen regionalen Preisvergleich keine Rolle spielen. Letztendlich bleiben 37 140 Preisreihen im Datensatz für den regionalen Preisvergleich erhalten.

Dem Vorschlag von Weinand und von Auer (2019) folgend, werden Mietpreise über eine hedonische Regression geschätzt, weshalb auch die Nettokaltmieten aus dem Datensatz entfernt werden. Dies hat auch den Vorteil, dass die räumliche Verteilung der Erhebungsstellen besser betrachtet werden kann, weil die Stichprobe der Nettokaltmieten viel stärker regional streut als die der restlichen Erhebungsstellen. Entfernt man alle Beobachtungen der Nettokaltmieten, so verbleiben 34 574 Preisbeobachtungen für den Mai 2019. Diese Preisbeobachtungen bieten die Grundlage für die nachfolgende Betrachtung des PMO-Ansatzes und des Ansatzes Homogene Güter.

## Vergleich PMO und Homogene Güter

Im Folgenden werden die beiden konkurrierenden Ansätze daraufhin verglichen, wie gut sie sich anhand der oben beschriebenen Daten der Verbraucherpreisstatistik realisieren lassen. Dabei geht es vor allem um die Frage, inwieweit die beiden Ansätze in der Lage sind, die Informationen in den vorhandenen Daten, die ursprünglich für zeitliche Preisvergleiche erhoben wurden, für einen regionalen Preisvergleich nutzbar zu machen. Wie bereits erläutert, besteht der Unterschied zwischen den beiden Ansätzen darin, wie sie die Produktgruppen definieren, die in den Preisvergleich eingehen. Der PMO-Ansatz arbeitet mit sehr engen Produktdefinitionen, die unter Verwendung der sogenannten Feinbeschreibungsmerkmale, die in den VPI-Daten enthalten sind, mit einigem Aufwand konstruiert werden. Die Produktgruppen des Ansatzes Homogene Güter lassen sich hingegen vergleichsweise einfach konstruieren, grundsätzlich anhand der Merkmale Güterart (Unterscheidung nach COICOP-10-Steller-Nummer) und Geschäftstyp.

### Konstruktion der Produkte nach dem PMO-Ansatz

Um den PMO-Ansatz zu realisieren, muss für jede beobachtete Preisreihe eine Produktbeschreibung definiert werden. Diese Beschreibung soll dabei zum einen feiner differenziert sein, als es über die Klassifizierung mittels der COICOP-10-Steller Kategorien möglich ist, um zum Beispiel Produkte von unterschiedlichen Marken unterscheiden zu können. Zum anderen sollen gleiche Produkte auch über eine gleiche Produktbeschreibung abgebildet werden, um deren Preise, zum Beispiel in unterschiedlichen Regionen, vergleichen zu können. Ziel ist es, ein Produkt über die Produktbeschreibung möglichst eindeutig zu identifizieren. Konkret setzt sich diese Produktbeschreibung aus den sogenannten Feinbeschreibungsmerkmalen, die für jede Preisbeobachtung vorliegen, zusammen. Die Feinbeschreibungsmerkmale werden von den Preiserheberinnen und Preiserhebern nach entsprechenden Vorgaben als „Metadaten" zu jedem beobachteten Preis notiert (Statistische Ämter des Bundes und der Länder 2021). Ihr eigentlicher Zweck liegt darin, eine Orientierungshilfe im Datenerhebungsprozess zu liefern: Mithilfe dieser detaillierten Angaben sollen Produkte, die für die Preiserhebung ausgewählt wurden, im Folgemonat leichter wiedergefunden werden[2]. Des Weiteren lassen sich dadurch etwaige Qualitätsänderungen des Produktes beurteilen. Ein Beispiel für

2 Gegebenenfalls kann es notwendig sein, bis zu zehn Feinbeschreibungsmerkmale zu erheben.

**Übersicht 2: Beispiel für Feinbeschreibungsmerkmale der Erhebungsposition „Vollmilch“**

| | Vorgabe der Feinbeschreibungsmerkmale für Produktkategorie „Vollmilch“ | Spezifische Feinbeschreibungsmerkmale | |
|---|---|---|---|
| 1. | Marke | Berchtesgadener Land | Weihenstephan |
| 2. | Produktart (Frische Milch, H-Milch) | Frische Milch | H-Milch |
| 3. | Verpackung (F=Flasche; B=Beutel) | F | B |
| 4. | Sonstiges | Laktosefrei | – |

vier Feinbeschreibungsmerkmale der Erhebungsposition „Vollmilch“ zeigt Übersicht 2.

Die Produktbeschreibung für den PMO-Ansatz wird durch die Zusammenfassung der ersten vier Feinbeschreibungsmerkmale umgesetzt. Dabei wird mittels Methoden der statistischen Textanalyse eine Bereinigung und Anpassung der Texteinträge durchgeführt, um Verzerrungen durch unterschiedliche Schreibweisen beziehungsweise Schreibfehler zu berücksichtigen. Diese Bereinigung erfolgt separat für jedes der vier Feinbeschreibungsmerkmale und unterteilt nach COICOP-10-Stellern.

Im ersten Schritt der Textbereinigung werden grundsätzliche Methoden zur Vereinfachung und Vereinheitlichung der Einträge angewendet. So werden Abkürzungen identifiziert und ausgetauscht, auf Kleinschreibung angepasst, Umlaute ersetzt (z. B. „ä“ zu „ae“) und bestimmte Sonderzeichen umgeschrieben (z. B. „&“ zu „und“) oder gelöscht. Weiterhin werden Satzzeichen und Zahlen sowie überzählige Leerräume im Text entfernt.

Nach dieser grundsätzlichen Bereinigung erfolgt in einem zweiten Schritt die Anpassung der Texteinträge. Innerhalb einer COICOP-10-Steller-Kategorie werden dabei für jedes Feinbeschreibungsmerkmal eindeutige Einträge ermittelt. Anschließend wird mithilfe der sogenannten Levenshtein-Distanz eine gewichtete Ähnlichkeit zwischen allen möglichen Wortpaarungen dieser eindeutig identifizierten Einträge berechnet. Die Levenshtein-Distanz gibt dabei die minimale Anzahl von Einfügungen, Löschungen und Ersetzungen an, die erforderlich sind, um eine Zeichenfolge in eine andere umzuwandeln. Mittels eines vorab definierten Grenzwertes in Abhängigkeit der Anzahl der Buchstaben pro Texteintrag erfolgt die Identifizierung sehr ähnlicher Einträge. In einem mehrstufigen Prozess werden selten vorkommende Einträge gegebenenfalls mit sehr ähnlichen, häufig vorkommenden Einträgen ersetzt. Konkret werden so kleinere Schreibfehler oder unterschiedliche Schreibweisen für ein Produkt in den Feinbeschreibungsmerkmalen angepasst.

Nach der Bereinigung und Anpassung der Texteinträge wird die Produktbeschreibung als Kombination der vier bereinigten Feinbeschreibungsmerkmale gebildet. Zusätzlich werden die Informationen zum jeweiligen Geschäftstyp, zur Menge, zur (Gewichts-)Einheit und zur Art des Angebots (z. B. Sonderangebot) zusammengefügt und angehängt. Die Maßeinheiten werden dabei, falls notwendig, vereinheitlicht und die Menge wird entsprechend angepasst (z. B.: „1 l“ wird zu „1 000 ml“). Die so gebildete detaillierte Produktbeschreibung definiert das jeweilige Produkt und bildet die Basis für den Preisvergleich mittels des PMO-Ansatzes: Preisvergleiche erfolgen nur zwischen Produkten mit exakt gleicher Produktbeschreibung. Insgesamt wurden für den Datensatz 18 450 unterschiedliche Produkte nach dem PMO-Ansatz gebildet.

### Konstruktion der Produkte nach dem Ansatz Homogene Güter

Zur Definition der Produktgruppen der homogenen Güter werden hauptsächlich zwei Variablen genutzt: Die COICOP-10-Steller-Nummer und die Variable „Geschäftstyp“. Insgesamt sind acht Geschäftstypen kategorisiert (siehe Übersicht 1), sodass sich theoretisch jeder COICOP-10-Steller in acht unterschiedliche Geschäftstypen unterteilen lässt. Zusätzlich wird die Variable „Maßeinheit“ zur Konstruktion genutzt. Mittels dieser Variablen wird für jede Preisbeobachtung eine Produktkategorie generiert. Diese Produktkategorien

setzen sich also aus der COICOP-10-Steller-Nummer, dem Geschäftstyp (der jeweiligen Erhebung) und der definierten Maßeinheit[3] zusammen und stellen damit die homogenen Güter dar. Die Maßeinheiten wurden dabei, wie bei der Konstruktion der PMO-Produkte, vereinheitlicht und die Menge wurde entsprechend angepasst. Zusätzlich wird in jeder so generierten Produktkategorie eine Basis-Menge definiert und der erhobene Preis je Beobachtung unter Verwendung der individuell erhobenen Menge auf diese Basis-Menge skaliert. Dieser skalierte Preis wird für die weiteren Berechnungen verwendet und stellt dadurch sicher, dass die Variabilität der Preise innerhalb einer so generierten Produktkategorie nicht durch unterschiedliche Mengenangaben verzerrt wird. Insgesamt wurden für den Datensatz 2 881 unterschiedliche Produktgruppen nach dem Ansatz Homogene Güter gebildet. Nachdem die Einteilung nach diesem Ansatz gröber ist, gibt es auch entsprechend weniger einzelne Kategorien als beim PMO-Ansatz.

Anders als beim PMO-Ansatz erfolgt beim Ansatz Homogene Güter der Preisvergleich nicht zwischen exakt gleichen Produkten, sondern zwischen Produkten derselben Produktgruppe.

### Anzahl der Güter im regionalen Warenkorb

Für den regionalen Preisvergleich können nur Produkte (PMO-Ansatz) beziehungsweise Produktgruppen (Ansatz Homogene Güter) berücksichtigt werden, die in mindestens zwei unterschiedlichen Regionen beobachtet werden. Nach dem PMO-Ansatz reduziert sich die Anzahl der unterschiedlichen Produkte daher von 18 450 auf 1 480 (das entspricht einer Reduktion auf 8 %). Nach dem Ansatz Homogene Güter reduziert sich die Anzahl der Produktgruppen von 2 881 auf 1 694 (das entspricht einer Reduktion auf 59 %). Hier zeigt sich die Problematik der sehr spezifischen Definition der Produkte über den PMO-Ansatz im regionalen Preisvergleich. Sehr viele Produkte können nicht berücksichtigt werden, da sie nur in einer Region beobachtet werden. Über die Konstruktion der homogenen Güter werden zwar grundsätzlich weniger unterschiedliche Produktgruppen definiert, jedoch sind dadurch (sowohl relativ als auch absolut) mehr Produkte für regionale Vergleiche verwendbar, weil die zugehörige Produktgruppe in mehreren Regionen vorhanden ist.

Die Auswahl der verbleibenden Produktgruppen hat damit auch einen direkten Einfluss auf die verbleibenden Beobachtungen des Datensatzes. Deswegen lohnt es sich, einen Blick auf das Ausmaß der Reduktion der einzelnen Preisreihen zu werfen. Durch die Datenaufbereitung und die grundsätzliche Datenbereinigung verringert sich der Datensatz insgesamt von 68 269 auf 34 574 Preisreihen (Nettokaltmieten wurden dabei schon bei den unbereinigten Daten ausgeschlossen, da diese nicht mit dem CPD-Ansatz, sondern mit hedonischen Methoden geschätzt werden sollen). Es bleiben also knapp 52 % der Preisreihen für den regionalen Vergleich übrig. Von diesen 34 574 Beobachtungen können über den Gruppierungsansatz Homogene Güter noch 28 118 Beobachtungen genutzt werden, im Fall des PMO-Ansatzes verbleiben nur 16 664 Beobachtungen, die in die Berechnung des Preisindex einfließen können. Das bedeutet, dass im Fall des PMO-Ansatzes weniger als ein Viertel der erhobenen Preisreihen für den regionalen Preisvergleich infrage kommen.

Abbildung 1 zeigt die absolute Anzahl der Beobachtungen für den gesamten Datensatz, den bereinigten Datensatz sowie nach Produktgruppierung mittels der Ansätze Homogene Güter und PMO. Zusätzlich sind die Beobachtungen nach Abteilungen (COICOP-2-Stellern) gruppiert, wie sie im VPI verwendet werden. Insgesamt betrachtet man damit den Warenkorb des Verbraucherpreisindex in zwölf Kategorien.

Diese starke Reduktion der Preisreihen im PMO-Ansatz hat noch einen zweiten Effekt (siehe Tabelle 1): Die Reduktion verteilt sich nicht gleichmäßig auf alle Abteilungen der COICOP-Klassifikation: Während in den Abteilungen „Wohnung ohne Mieten“, „Medizin“, „Verkehr“ und „Bildung“ relativ geringe Einbußen vom bereinigten Datensatz zum PMO-Ansatz zu finden sind, reduzieren sich andere Abteilungen wie „Nahrungsmittel“, „Möbel“ und „Post und Telefon“ wesentlich

3 Generell wird innerhalb einer COICOP-10-Steller-Kategorie dieselbe Maßeinheit verwendet, in Einzelfällen kann es jedoch zu Abweichungen kommen.

Abb. 1
**Anteil der jeweiligen COICOPs an erhobenen Preisen für alle Daten sowie nach den Ansätzen PMO und Homogene Güter (in mind. zwei Regionen)**
Anzahl der Beobachtungen

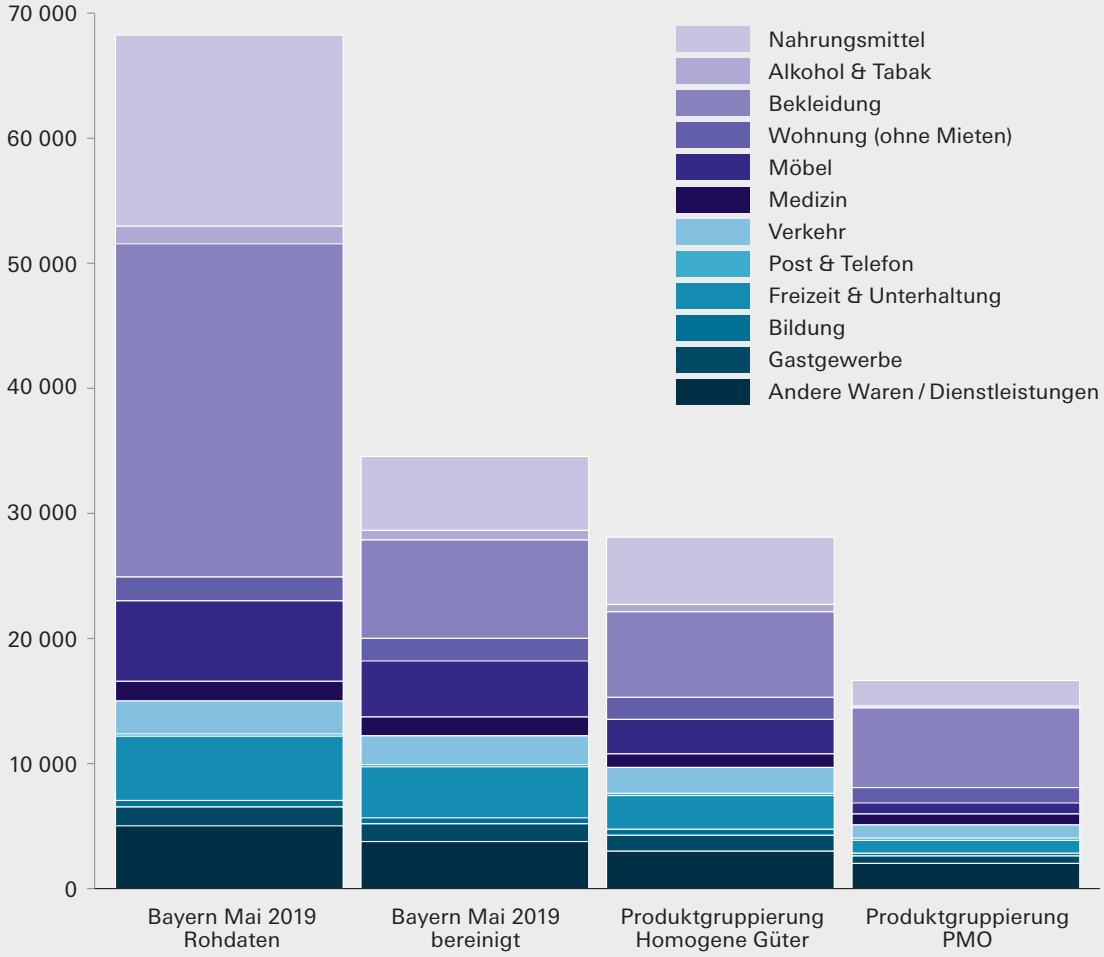



**Tabelle 1: Nutzbare Daten je COICOP für die Ansätze Homogene Güter und PMO**

| | Abteilung | In Daten erhoben | Daten bereinigt | Daten bereinigt Anteil | Homogene Güter | Homogene Güter Anteil erhalten | PMO | PMO Anteil erhalten |
|---|---|---|---|---|---|---|---|---|
| Nahrungsmittel | 1 | 15 289 | 5 910 | *38,7 %* | 5 379 | *35,2 %* | 2 053 | *13,4 %* |
| Alkohol & Tabak | 2 | 1 400 | 769 | *54,9 %* | 587 | *41,9 %* | 147 | *10,5 %* |
| Bekleidung | 3 | 26 644 | 7 867 | *29,5 %* | 6 842 | *25,7 %* | 6 366 | *23,9 %* |
| Wohnung ohne Mieten | 4 | 1 908 | 1 801 | *94,4 %* | 1 761 | *92,3 %* | 1 225 | *64,2 %* |
| Möbel | 5 | 6 432 | 4 485 | *69,7 %* | 2 757 | *42,9 %* | 885 | *13,8 %* |
| Medizin | 6 | 1 575 | 1 506 | *95,6 %* | 1 083 | *68,8 %* | 874 | *55,5 %* |
| Verkehr | 7 | 2 687 | 2 356 | *87,7 %* | 2 138 | *79,6 %* | 1 141 | *42,5 %* |
| Post & Telefon | 8 | 66 | 57 | *86,4 %* | 42 | *63,6 %* | 1 | *1,5 %* |
| Freizeit & Unterhaltung | 9 | 5 202 | 4 143 | *79,6 %* | 2 768 | *53,2 %* | 1 127 | *21,7 %* |
| Bildung | 10 | 522 | 481 | *92,1 %* | 477 | *91,4 %* | 217 | *41,6 %* |
| Gastgewerbe | 11 | 1 501 | 1 415 | *94,3 %* | 1 275 | *84,9 %* | 596 | *39,7 %* |
| Andere Waren/Dienstl. | 12 | 5 043 | 3 784 | *75,0 %* | 3 009 | *59,7 %* | 2 032 | *40,3 %* |
| **Summe** | | 68 269 | 34 574 | *50,6 %* | 28 118 | *41,2 %* | 16 664 | *24,4 %* |

stärker. Die Abteilung „Post und Telefon“ verschwindet dabei fast vollständig. Der ursprüngliche Warenkorb des VPI wird also nicht proportional in allen Bereichen gleich reduziert, sondern nicht nutzbare Beobachtungen konzentrieren sich systematisch in einzelnen Bereichen. Dabei ist vor allem problematisch, dass nur 13,4 % bei den Nahrungsmitteln für einen regionalen Vergleich nach dem PMO-Ansatz infrage kommen. Dies bewirkt, dass der so erzeugte Warenkorb nicht mehr repräsentativ für die Gebrauchsgewohnheiten in Deutschland sein kann.

Im Falle des Ansatzes Homogene Güter ist die Reduktion wesentlich geringer. Insgesamt können über 40 % der Produkte als ausreichend vergleichbar für einen regionalen Preisvergleich angesehen werden. Im Gegensatz zum PMO-Ansatz bleiben in allen Abteilungen genügend Beobachtungen erhalten. Die größten Verluste verzeichnet der Ansatz Homogene Güter in der Abteilung „Bekleidung“, wo nur ein Viertel der erhobenen Preise für einen regionalen Preisvergleich zur Verfügung steht. In der Abteilung „Nahrungsmittel“ sind noch 35,2 % der Preisbeobachtungen für die Berechnung der Preisindizes nutzbar.

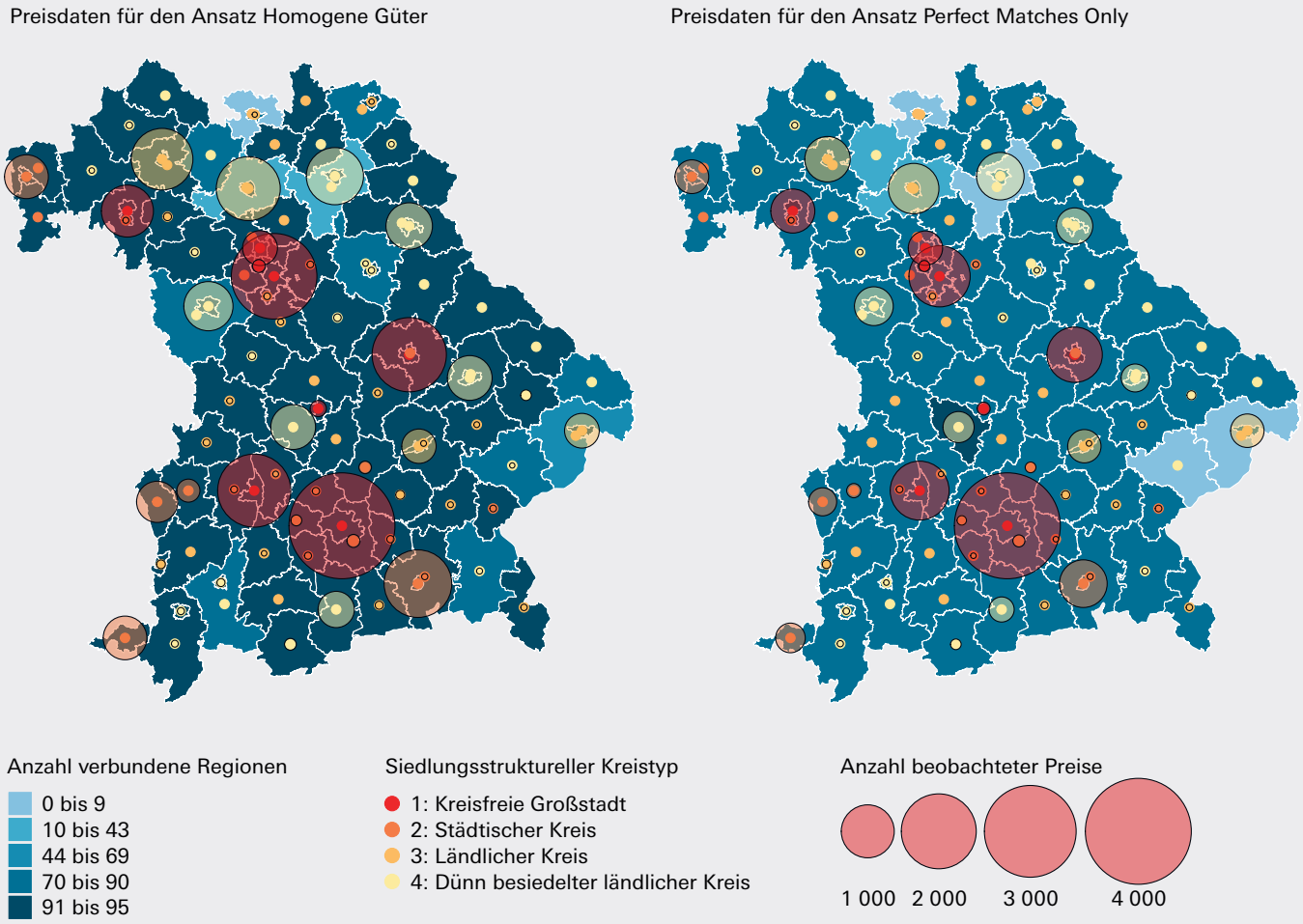
Die systematische Verzerrung des Warenkorbs im PMO-Ansatz lässt sich auch darstellen, indem man betrachtet, wie sich die Bereinigung und die Ansätze auf das Wägungsschema des VPI auswirken. Die Gewichte bestimmen den Anteil, mit dem die einzelnen COICOP-10-Steller in die Berechnung des VPI eingehen. Sie spiegeln das durch die Einkommens- und Verbrauchsstichprobe (EVS) geschätzte Verbrauchsverhalten eines typischen deutschen Haushalts wider. Dem Wägungsschema kommt damit sowohl in der Definition des Warenkorbs als auch in der Berechnung des VPI eine entscheidende Rolle zu. Kommt es hier zu Verzerrungen, ist der darauf basierende Preisindex nicht mehr repräsentativ für die Konsumgewohnheiten in Deutschland.

Die Gewichte für den Warenkorb des VPI beziehen sich auf die COICOP-10-Steller und nicht auf die Elementarindizes, also die Kombination von COICOP-10-Steller und Geschäftstyp. Deswegen vergröbert sich an dieser Stelle unsere Betrachtungsweise. Die Veränderungen der Warenkorb-Zusammensetzung auf COICOP-2-Steller-Ebene wurden für die beiden Ansätze geprüft. Das Ergebnis ist sehr erfreulich: Keine der beiden Methoden führt zu Verzerrungen hinsichtlich des Wägungsschemas, sodass nach beiden Methoden die deutschen Verbrauchsgewohnheiten bei Berücksichtigung des Wägungsschemas auch bei einem regionalen Preisindex repräsentiert werden.

Neben der Zusammensetzung des Warenkorbs an sich muss für die räumlichen Preisvergleiche auch berücksichtigt werden, wie viele Regionen sich jeweils miteinander vergleichen lassen. Die Bedingung hierfür ist, dass mindestens ein Produkt existiert, das in beiden Regionen beobachtet wurde. Gibt es ein solches Produkt, kann von einem „Match“ zwischen den Regionen gesprochen werden. Dabei zeigt sich in Bezug auf die Matches und die jeweils nutzbaren Preisbeobachtungen für die Ansätze PMO und Homogene Güter ein unterschiedliches Bild. Eine regionale Übersicht zur Anzahl der verbundenen Regionen (Matches) und der Preisbeobachtungen je Landkreis ist in Abbildung 2 dargestellt.

Es werden die Daten für die Ansätze PMO und Homogene Güter in zwei separaten Karten gezeigt. Die Anzahl der Verbindungen mit einer jeweils anderen Region ist für jeden Landkreis anhand einer farblichen Kategorisierung dargestellt. Je dunkler, desto mehr Matches hat ein Landkreis mit anderen Regionen. Zusätzlich wird die Anzahl der nutzbaren Preisbeobachtungen über die Größe des Kreises in der Region verdeutlicht. Es ist ersichtlich, dass für homogene Güter wesentlich mehr Matches in den Landkreisen existieren. Die Karte ist hier deutlich dunkler eingefärbt als die für den PMO-Ansatz. Ein Großteil der Landkreise ist beim Ansatz Homogene Güter mit 91 oder mehr der jeweils 95 verbleibenden Landkreise verbunden. Im Gegensatz dazu gibt es beim PMO-Ansatz mehr Landkreise als beim Ansatz Homogene Güter, die mit sehr wenigen (neun oder weniger) anderen Landkreisen verbunden sind. Auch die Anzahl der nutzbaren Preisreihen pro Landkreis liegt bei den homogenen Gütern über der Anzahl

Abb. 2
**Anzahl Matches und Anzahl Preisbeobachtungen je kreisfreie Stadt und je Landkreis in Bayern**

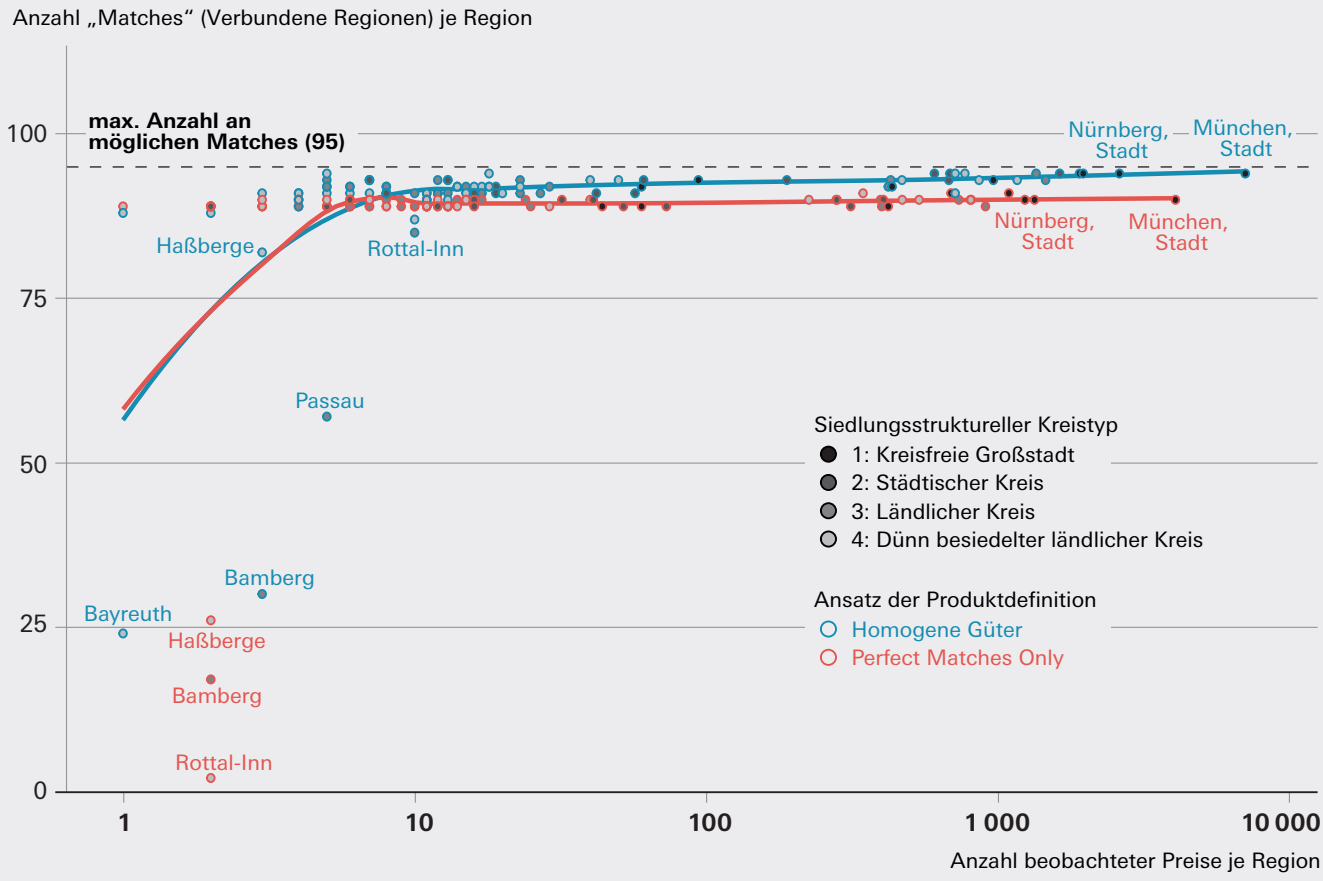



des PMO-Ansatzes. Dies wird anhand der Größe der Kreise in der jeweiligen Region ersichtlich (für Regionen mit weniger als zehn Preisbeobachtungen wird dabei kein Kreis dargestellt).

Den Zusammenhang zwischen der Anzahl an Preisbeobachtungen und der Anzahl an Matches veranschaulicht Abbildung 3. Dabei sind die Daten für den PMO-Ansatz und den Ansatz Homogene Güter entsprechend unterschiedlich markiert.

Abbildung 3 stellt die Anzahl der nutzbaren Preisbeobachtungen je Landkreis und die Anzahl der Matches mit jeweils anderen Regionen je Landkreis dar. Dabei zeigt sich ein grundsätzlicher Zusammenhang zwischen der Anzahl der beobachteten Preise je Landkreis und der Anzahl der Verbindungen mit anderen Regionen: Je mehr Preisreihen in einem Landkreis beobachtet werden (bzw. nutzbar sind), desto mehr Verbindungen hat die Region mit anderen Landkreisen. Es ist nicht verwunderlich, dass der PMO-Ansatz mit seiner strengen Vergleichbarkeitsregel bei einer vergleichbaren Anzahl von Preisbeobachtungen im Schnitt wesentlich weniger Matches mit anderen Regionen generiert als der Ansatz Homogene Güter. Deshalb liegt die geglättete Linie des PMO-Ansatzes ab sieben Preisbeobachtungen unter der des Ansatzes Homogene Güter. Unter sieben Preisbeobachtungen ist kein unterschiedlicher Zusammenhang ersichtlich. Ab einer gewissen Anzahl von Beobachtungen sind bei den Daten des Ansatzes Homogene Güter in ausreichender Menge Matches vorhanden, so dass die jeweilige Region

Abb. 3
**Vergleich der kreisfreien Städte und Landkreise nach den Ansätzen PMO und Homogene Güter**




hinsichtlich ihres Preisniveaus mit annähernd allen anderen Regionen verglichen werden kann.

Es finden sich vereinzelt Punkte beim Ursprung des Koordinatensystems. Diese Punkte stehen für Regionen, in denen es wenige Matches mit anderen Landkreisen gibt und in denen auch sehr wenige nutzbare Preisreihen erhoben wurden[4]. Im Fall von PMO handelt es sich um die drei Landkreise Rottal-Inn, Bamberg und Haßberge. Zusätzlich gibt es für den PMO-Ansatz in den zwei Landkreisen Passau und Bayreuth kein Produkt, das auch in anderen Regionen beobachtet wurde und damit auch kein Match mit anderen Landkreisen (diese zwei Landkreise sind in der Graphik für den PMO-Ansatz nicht dargestellt). Beim Ansatz Homogene Güter handelt es sich um drei Landkreise, in denen es wenige Preisbeobachtungen und Matches gibt: Bayreuth, Bamberg und Passau. Für die Regionen mit sehr wenigen nutzbaren Preisreihen basiert der regionale Preisindex, den es zu berechnen gilt, auf wenigen Datenpunkten. Deshalb ist hier die Schätzung eines regionalen Preisindex sehr unsicher. Bei allen fünf Regionen handelt es sich um ländliche oder dünn besiedelte ländliche Kreise. Die meisten Preisbeobachtungen, sowohl für den Ansatz PMO als auch für den Ansatz Homogene Güter, wurden in den kreisfreien Städten München und Nürnberg gemacht. Dies spiegelt wider, dass sich die Preiserhebung auf die regionalen Zentren konzentriert.

## Höhe der Variabilität

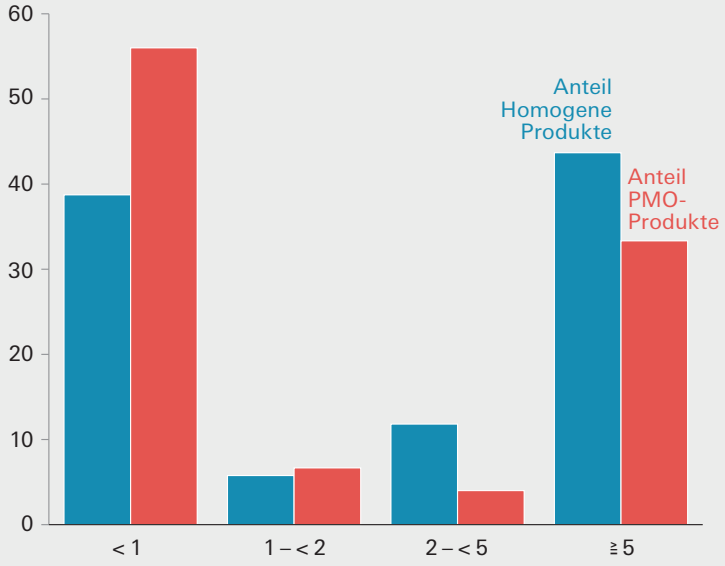
Wie bereits erwähnt, gibt eine Betrachtung der externen und internen Varianzen Auskunft über das Ausmaß an Qualitätsunterschieden einzel-

4 Die Region Coburg ist nicht enthalten, da hier im Mai 2019 keine Daten erhoben wurden.

ner Produktgruppen. Konkret wird dabei der Variationskoeffizient der internen Varianz (Streuung innerhalb der Regionen) und der Variationskoeffizient der externen Varianz (Streuung zwischen den Regionen) berechnet. Als Maß wird dann der Quotient aus internem und externem Variationskoeffizienten pro Produktgruppe gebildet. Die Ergebnisse dieses Vergleichs werden im folgenden Abschnitt vorgestellt. In Abbildung 4 werden die Quotienten der internen und externen Variationskoeffizienten nach den beiden Ansätzen anteilsmäßig für vier Kategorien dargestellt.

Die Thematik der Variabilität der Produktgruppen wurde grundsätzlich für den Ansatz Homogene Güter problematisiert. Beim PMO-Ansatz wird aufgrund der spezifischen Produktkategorisierung zunächst eine ausreichende Homogenität innerhalb der Produktgruppen unterstellt. Eine genauere Untersuchung erscheint aber auch hier sinnvoll.

Abb. 4
**Quotient der internen und externen Variationskoeffizienten nach den Ansätzen PMO und Homogene Güter**
in Prozent




Für jede Produktgruppe werden nach der entsprechenden Gruppierung (Homogene Güter oder PMO) die entsprechenden internen und externen Variationskoeffizienten berechnet und der Quotient gebildet. Eine Voraussetzung zur Berechnung des Varianzkriteriums liegt darin, dass das jeweilige Produkt sowohl in unterschiedlichen Regionen als auch mehrfach innerhalb der Regionen beobachtet werden muss. Hier zeigt sich, dass sich die Anzahl der dafür infrage kommenden Produktgruppen nochmal deutlich reduziert. Beim Ansatz Homogene Güter bleiben 728 Produkte übrig, beim PMO-Ansatz 75 Produkte. Zudem wird deutlich, dass einige Produktgruppen eine gute Vergleichbarkeit auf Grundlage des Varianzkriteriums bieten, andere diesbezüglich weniger gut abschneiden. Die Variabilität beziehungsweise ausreichende Homogenität innerhalb einer Produktgruppe kann nicht als selbstverständlich angesehen werden. Hier sollten noch genauere Untersuchungen anschließen. Dies sollte einen Einfluss auf eine bessere Konstruktion der Produktgruppen haben.

Es wird ersichtlich, dass zwar beim Ansatz Homogene Güter die Gruppen im Durchschnitt weniger homogen sind, die Unterschiede hinsichtlich der Variabilität im Vergleich zum PMO-Ansatz jedoch nicht allzu groß sind. Nachdem die Homogenität der Gruppen im PMO-Ansatz aufgrund des strikten Vergleichs sehr hoch ist, kann man aus diesem Vergleich schließen, dass auch der Ansatz Homogene Güter ausreichende Vergleichbarkeit schafft.

## Ausblick: Homogene Güter in der stufenweisen CPD-Schätzung

Für die Schätzung regionaler Preisniveaus ist die Zusammenstellung des zugrundeliegenden Warenkorbs entscheidend. Wichtig dabei ist, dass die regionalen Preisvergleiche tatsächlich auch auf Preisunterschieden und nicht auf Qualitätsunterschieden basieren.

Der von Weinand und von Auer (2019) gewählte Ansatz, der eine absolute Gleichheit der Produkte fordert, schließt zwar Qualitätsunterschiede in einem ersten Schritt aus, wird aber in der darauffolgenden CPD-Schätzung wieder aufgeweicht. Im Endeffekt führt dieses Verfahren dazu, dass sehr viele Produkte von der Analyse ausgeschlossen werden, wenige Datenpunkte je Region vor-

handen sind und die Vergleichbarkeit der Warenkörbe der einzelnen Regionen miteinander nicht steigt. Die Vergleichbarkeit ist immer nur zwischen den Regionen und nur bilateral je Produkt, nicht aber multilateral je Warenkorb gesichert. Allein aufgrund der gewählten Ebene in der CPD-Schätzung und weil mehr Preisreihen erhalten bleiben, erscheint es sinnvoll, die Vergleichsbedingung zu lockern und den Ansatz Homogene Güter zu verfolgen. Etwaige Qualitätsunterschiede können zumindest durch Wahl der Geschäftstypen in geringerem Ausmaß kontrolliert werden. Das Ideal des reinen Preisvergleichs kann in der regionalen Betrachtung in seiner absoluten Interpretation keinen Bestand haben.

Letztendlich kann festgestellt werden, dass die meisten Untersuchungen, die sich mit räumlichen Preisvergleichen beschäftigen, internationale Vergleiche ziehen (vgl. etwa Silver und Heravi 2005). Natürlich spielen bei diesen Analysen Qualitätsunterschiede und Vergleichbarkeit eine wesentlich größere Rolle, als es bei dem vorliegenden regionalen Preisvergleich innerhalb eines Bundeslands der Fall ist.

Als nächster Schritt zu einem regionalen Preisindex für Bayern wird die CPD-Methode auf den Ansatz Homogene Güter angepasst. Das von Weinand und von Auer (2019) vorgeschlagene mehrstufige Verfahren soll dabei das Grundgerüst bilden.

## Literatur

Bayerisches Staatsministerium für Wirtschaft, Verkehr und Technologie (2003): Die reale Kaufkraft in Bayern 2002.

Bleninger, Sara (2017): KriMI: A Multiple Imputation Approach for Preserving Spatial Dependencies. Imputation of Regional Price Indices using the Example of Bavaria. Bamberg: University of Bamberg Press.

Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (2021): Metadatenreport. Teil I: Allgemeine und methodische Informationen zu den Einzeldaten des Verbraucherpreisindex (EVAS-Nummer: 61111). Version 2. Hg. v. Statistische Ämter des Bundes und der Länder. Wiesbaden.

Götz, Georg / Krauskopf, Thomas / Winkler, Peter (2009): Die Bestimmung regionaler Preisindizes – Das Beispiel Österreich. Reserach Note. RatSWD.

Greuel, Marcel (2014): Internationale Kaufkraftparitäten: Methodik und empirische Umsetzung. Dissertationsschrift. Trier. Online verfügbar unter https://ubt.opus.hbz-nrw.de/opus45-ubtr/frontdoor/deliver/index/docId/632/file/Dissertation_MarcelGreuel.pdf, zuletzt abgerufen am 10.01.2021.

Guckes, Siegfried (1979): Aus der Praxis der Berechnung von Preisindizes. In: Allgemeines Statistisches Archiv 63 (3).

Kawka, Rupert (2009): Regionaler Preisindex. Hrsg. von BBSR, Bundesinstitut für Bau-, Stadt- und Raumforschung.

Reich, Utz-Peter (2000): Messung des Geldwertes: Zur Statistik und Theorie der reinen Preisbewegung. In: Allgemeines Statistisches Archiv 84, S. 461 – 487.

Rostin, Werner (1979): Zwischenörtlicher Preisvergleich des Verbraucherpreisniveaus in 31 Städten. In: Wirtschaft und Statistik 6 / 1979.

Silver, Mick (2009): The Hedonic Country Product Dummy Method ans Quality Adjustments for Purchasing Power Parity Calculations. In: IMF Working Paper (271).

Silver, Mick / Heravi, Saeed (2005): Purchasing Power Parity Measurement and Bias from Loose Item Specifications in Matched Samples: An Analytical Model and Empirical Study. In: Journal of Official Statistics 21 (3), S. 463 – 487.

Statistische Ämter des Bundes und der Länder (2021): Handbuch für die Preiserhebung in der Verbraucherpreisstatistik. Hrsg. von Statistische Ämter des Bundes und der Länder.

Statistisches Bundesamt (2015): Berechnungshandbuch: Verbundprogramm der Verbraucherpreisstatistik. Version 4.1.8, 2015.

Ströhl, Gerd (1994): Zwischenörtlicher Vergleich des Verbraucherpreises in 50 Städten. In: Wirtschaft und Statistik 6 / 1994.

United Nations: Department of Economic and Social Affairs, Statistics Division (2018): Classification of Individual Consumption According to Purpose (COICOP) 2018. Hrsg. von United Nations. Online verfügbar unter https://unstats.un.org/unsd/classifications/unsdclassifications/COICOP_2018_-_pre-edited_white_cover_version_-_2018-12-26.pdf, zuletzt abgerufen am 09.02.2022

Vogel, Friedrich (2000): Beschreibende und schließende Statistik. Formeln, Definitionen, Erläuterungen, Stichwörter und Tabellen. 12. Aufl. München, Oldenbourg Verlag.

von der Lippe, Peter Michael (1977): Wirtschaftsstatistik. 2. Aufl. Stuttgart, Gustav Fischer Verlag.

von der Lippe, Peter Michael (2000): Preisindizes mit konstanten Wertegewichten. Bemerkungen zu einem Vorschlag von Werner Neubauer. In: Jahrbücher für Nationalökonomie und Statistik 220 (6).

von der Lippe, Peter Michael / Breuer, Claus Christian (2009): Regionale Kaufkraftvergleiche in Deutschland: Bedarf, Methoden und Machbarkeit. In: AStA Wirtschafts- und Sozialstatistisches Archiv 3 (1).

Weinand, Sebastian / von Auer, Ludwig (2019): Anatomy of Regional Price Differentials. Evidence from Micro Price Data. Research Papers in Economics. Universität Trier, Trier. Online verfügbar unter https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3334287, zuletzt abgerufen am 30.11.2021.