Measuring price dynamics of package holidays with transaction data

Karola Henn[1], Chris-Gabriel Islam[1], Patrick Schwind[2] and Elisabeth Wieland[2]

This version: 08 May 2019

In Germany, package holidays, which consist of a bundle of flight and accommodation services, are an important driver of consumer prices. Several challenges arise when measuring the price development of package holidays, e.g. the quality of accomodation, the timing of the booking, the treatment of out-of-season services as well as the underlying holiday season. Statistical practices are currently based on sampling offer prices. As a possible alternative, transaction price data from the commercial booking system "Amadeus" are analysed in this study. Our data set comprises both online and offline bookings of package holidays on a daily basis which allow for a disaggregation by individual holiday destination due to their large sample size.

The paper analyses the chances and challenges in compiling a price index out of transaction data for flight package holidays. The data set raises a number of methodological issues, e.g. the grouping of unstructured text information into meaningful categories, the handling of missing information or the identification of outliers. Moreover, various index aggregation methods are analysed, which include hedonic regressions, stratification, and also a multilateral index method. Applied to six major holiday destinations for German travellers, all transaction-based methods under consideration exhibit similar price dynamics. Yet, further research is required at the micro level to assess whether the currently applied transaction-based price methods perform sufficiently well in terms of varying sample and quality adjustment.

[1] Federal Statistical Office (Destatis). [2] Deutsche Bundesbank. Corresponding author: elisabeth.wieland@bundesbank.de. This paper represents the authors' personal opinions and does not necessarily reflect the views of Destatis, Deutsche Bundesbank or its staff. The authors would like to thank for valuable comments notably from Timm Behrmann, Dorothee Blang, Edgar Brandt, Antonio Chessa, Thomas Knetsch, Michael Kuhn, Susanne Lorenz and Jens Mehrhoff. All remaining errors are ours.

Contents

1 MOTIVATION 3
2 CURRENT OFFICIAL PRACTICE IN THE GERMAN HICP 4
3 DESCRIPTION OF THE AMADEUS DATASET 6
4 METHODS OF PRICE MEASUREMENT 10
4.1 UNIT VALUE PRICE INDEX 11
4.2 HEDONIC REGRESSION METHODS 11
4.2.1 DOUBLE IMPUTATION 12
4.2.2 TIME DUMMY MODEL 14
4.3 STRATIFICATION METHODS 14
4.3.1 PRODUCT DEFINITION BY A QUANTITATIVE APPROACH 15
4.3.2 TRADITIONAL STRATIFICATION 18
4.3.3 GEKS 18
5 COMPARISON OF RESULTS 20
6 SUMMARY 26
REFERENCES 27
ANNEX 29

1 Motivation

Package holidays as a challenge for official price statistics

In traditional price collection, offer prices from pre-defined price representatives usually are collected at fixed points in time every month. The more complex a given good or service is, the more manual work is required by a National Statistical Institute (NSI) in setting up a sufficiently large selection of price representatives. This is especially true for bundles of different services, such as package holidays, which are made up of both travel and accommodation (hotel) services and have a lot of price-determining characteristics such as the category of the hotel as well as the meal type, the room or the departure airport. Moreover, travel-related prices such as the flight can fluctuate heavily within a given month.

Package holidays have a noticeable effect on the German and even the European HICP

In German price statistics, package holidays have currently a weight of 2.7 percent in the Harmonised Index of Consumer Prices (HICP) as of 2019. However, due to their high volatility and strong seasonality, package holidays have a noticeable effect on the German and even the euro area inflation rate. The Federal Statistical Office (Destatis) currently uses a global distribution system from Amadeus – as applied by travel agencies – to collect offer prices of package holidays. The sample size is limited due to the high effort required for manual price collection.[2] Therefore, it is currently not possible to publish the price development broken down by holiday regions, but by the broad sub-indices “Domestic package holidays” (ECOICOP 09.6.0.1) and “International package holidays” (ECOICOP 09.6.0.2).[3]

Transaction data on package holidays are an alternative source for price measurement

An alternative to collecting offer prices consists in actual bookings of international package holidays recorded in the Amadeus IT booking systems, which are used by online travel agencies or at traditional high street travel agencies. The aim of this paper is to investigate the chances and challenges when compiling a price index out of transaction data for flight package holidays[4], which are very heterogeneous seasonal services. Furthermore, due to its large sample size, such an experimental price index could be subdivided into relevant holiday regions, thus allowing for an economic interpretation of the underlying price movements of package holidays. The contribution of this paper is also in applying the most recent index aggregation methods, which include hedonic regressions, stratification, and also a multilateral method, to the relatively new field of measuring prices of (bundled) services by transaction data.

The paper is structured as follows: Section 2 describes the current official practice in measuring prices of package holidays by the Federal Statistical Office, which is based on offer prices. Section 3 presents the transaction dataset from Amadeus and

---

2 The Federal Statistical Office is currently extending the price collection for the package holiday index to a larger number of price representatives per destination and to a larger number of travel days per month by an automatic way with the help of the application interface of Amadeus.

3 The goods and services in the HICP are grouped according to the European Classification of Individual Consumption according to Purpose (ECOICOP). For an overview of this classification, see: https://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=COICOP_5&StrLanguageCode=EN&IntPcKey=34740023&StrLayoutCode=HIERARCHIC.

4 Note that, besides flight package holidays, the German HICP sub-index on package holidays also consists of domestic package holidays, shorter city trips to other European countries and cruises (see Section 2), which were not the subject of this study.

comments on the challenges of processing these data for the purpose of price statistics. Section 4 discusses various methods commonly used to measure prices, as well as newer index methods that have recently been developed on the basis of scanner data. Section 5 compares the price indices derived from the various methods for six major holiday destinations of German travellers. Section 6 concludes and provides an outlook on regional price indicators on package holidays.

2 Current official practice in the German HICP

HICP package holidays calculated by fixed weights method

In contrast to the transaction prices used in this paper, the German Federal Statistical Office collects offer prices to calculate the official HICP sub-index package holidays. This data represent a very detailed specified sample of trips, with the aim to ensure a pure price comparison. According to EU regulation, there are two methods that are allowed for calculating indices for package holidays: fixed weights method (also known as strict annual weights) and a class-confined seasonal weights method.[5] Before the German national CPI was revised and rebased to 2015=100 in February 2019, the class-confined seasonal weights method was used, with a different summer and winter sample. From the reporting year 2015 onwards, the official HICP sub-index package holidays is based on the fixed weights method, where missing prices of the out-of-season months are imputed.[6]

Main component: International flight package holidays (7 to 14 days)

Table 1 provides an overview of the elementary aggregates of the German HICP package holidays (09.6). The sample for the sub-index "international package holidays" consists of holidays from Germany to six holiday destinations (Balearic Islands, Canary Islands, Greece, Turkey, Egypt and the Dominican Republic) with duration of seven to 14 days and to two countries for shorter city trips. Moreover, the international aggregate includes cruises. For most holiday destinations, there exist three strata: summer, winter and whole-year strata (for four holiday destinations). Missing prices for the summer sample within a given holiday region are imputed using the winter or the whole year sample and vice versa (counter-seasonal estimation). For two countries, there is only a summer or winter sample and missing prices are imputed using all other available prices (all-seasonal estimation).

[5] See European Commission Regulation No 330/2009, Article 2, as well as Eurostat (2018), Chapter 7.1 Seasonal products and Chapter 12.5 on flights and package holidays.

[6] Switching to CPI basis 2015 and using the fixed weights methods improved the interpretability of the previous month's rate of change in April, May and November. At the same time, it increased the seasonal profile of the package holiday price index, with higher index values in the summer and lower values in the winter season.

Table 1: Elementary aggregates of German HICP sub-index package holidays (09.6)

ECOICOP	Weight of 09.6 in %	Coverage	Sample period
09.6.0.1 Domestic package holidays	5.60	Germany only, travel by train or car	summer/winter
09.6.0.2 International package holidays			
International flight package holidays (7 to 14 days)	76.95	4 holiday destinations	summer/winter/whole year
		2 holiday destinations	summer or winter only
City trips		2 holiday destinations	whole year
Cruises	17.45	Combination of flight and open-sea cruise	summer only

Collection of offer prices via START Amadeus

In German price statistics, the offer prices for international package holidays are collected from the booking system “START Amadeus”[7] via internet and cover roughly 300 price representatives. Booking codes from tour operators are used to identify a product offer with pre-defined attributes (e.g. Hotel XXXX, all inclusive, double room with sea view, for two persons and ten days, with departure flight from Frankfurt am Main). The price representatives are calculated using three offer prices (three inquiries on different points in time in advance of a given departure) for the winter/summer season or 21 offer prices (three inquiries in advance of seven departure days) for the whole-year season. In total, about 1,500 to 3,000 offer prices (depending on the timing of public holidays) are included in the price calculation of a given travel month.

Package holidays weigh heavily in German HICP

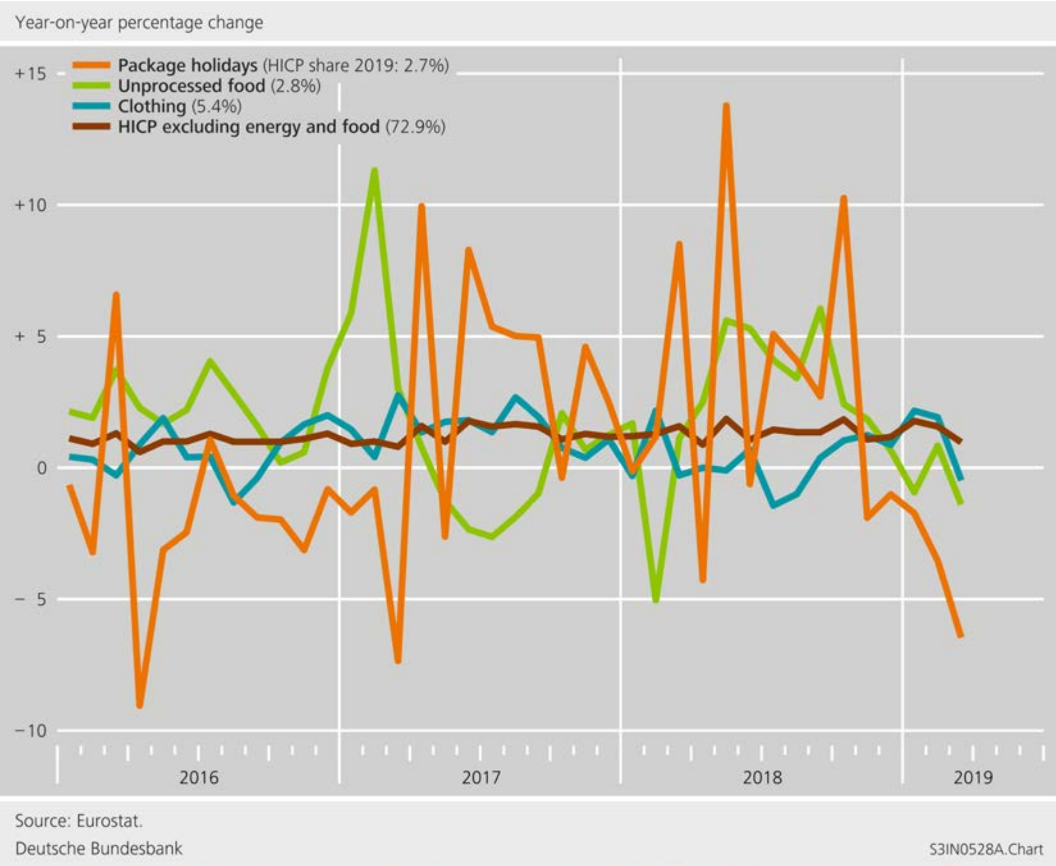
The resulting German HICP sub-index on package holidays exhibits a high volatility, as shown in Figure 1. The monthly year-on-year change in the inflation rate from 2016 onwards ranges between -9 to +14 percentage points and is therefore more volatile than other seasonal HICP components, such as clothes or unprocessed food. From a user’s point of view, a more detailed breakdown by holiday regions would be helpful in interpreting those movements.[8] From an international perspective, the weight of package holidays in the German HICP (2019: 2.7 %) is one of the highest among European countries, with higher values only observed in Iceland (6.3 %), United Kingdom (4.2 %) and Cyprus (3.2 %). Because of its weight and volatility, the challenges of measuring prices for package holidays with transaction data and how to derive prices for bundled services, which are generally more complex

[7] The company Amadeus Germany GmbH operates an IT system for sales and marketing in the field of travelling in Germany. The booking system “START Amadeus” is used by traditional high street travel agencies to handle all booking transactions for package holidays. In contrast, the offer prices for city trips are collected manually from different online travel agencies, whereas for cruises, catalogue prices are compiled.

[8] See also Deutsche Bundesbank (2017) for a comment on the impact of HICP package holidays on core inflation in Germany.

than super market goods, are very important to Germany, but may be relevant to other (European) NSIs as well.[9]

Figure 1: German HICP package holidays compared with other components



3 Description of the Amadeus dataset

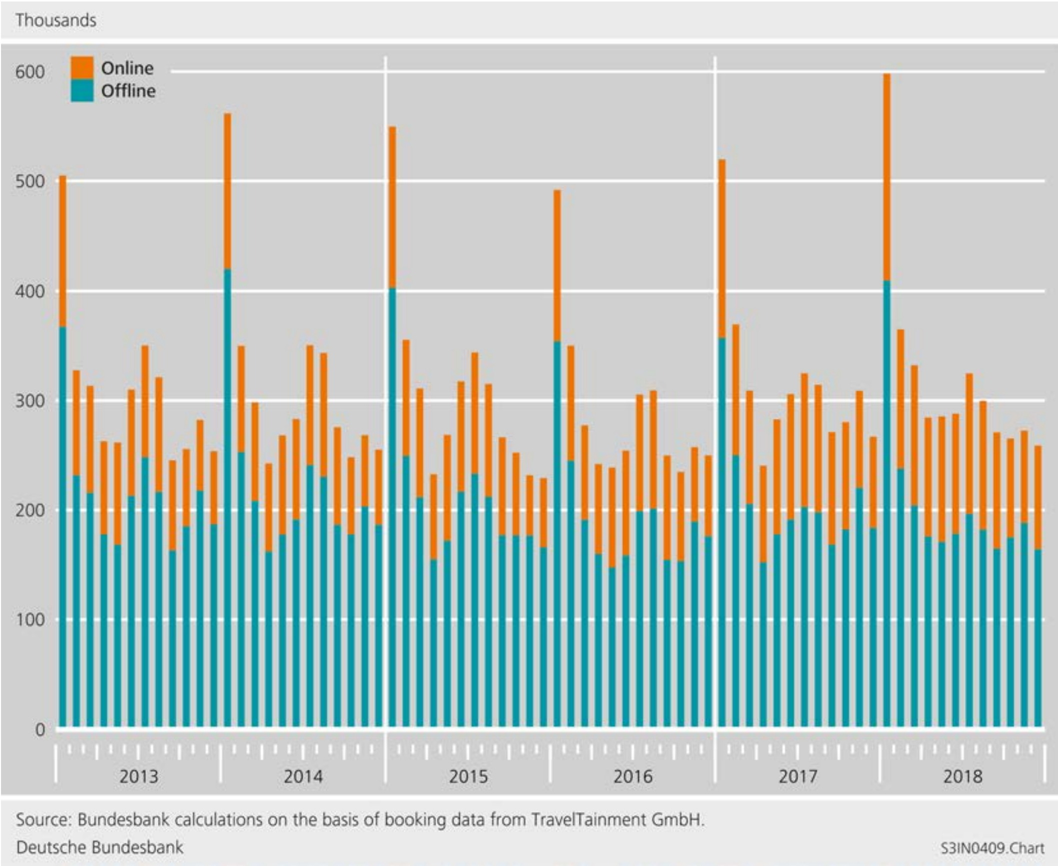
Extensive information on bookings and price determinants

Commercial IT specialist Amadeus' dataset contains around 3.7 million transaction prices per year for flight package holidays for German travellers for the period from 2013 to 2018. The data are collected via the Amadeus booking system, which is used by online travel portals as well as traditional high street travel agencies in Germany.[10] Transactions are provided by booking date and are readily available in the first calendar week after the end of each booking month. For each transaction, information on price determinants such as accommodation, holiday region and number of travellers is given.[11] The data are made up of both offline and online bookings. The offline data constitute the larger component (see Figure 2) and usually contain two to three times as many observations as online data, but they do not contain detailed information on meal type, room, car rentals and travel insurances. Given the different levels of information provided as well as the possibility of differ-

[9] To the best of our knowledge, only the Dutch NSI already implemented a transaction-data based price index for package holidays in their regular data production. For this purpose, they use a method that is very similar to the traditional stratification method in this paper (see Section 4.3.2).
[10] According to "WirtschaftsWoche" (issue 27/2018), Amadeus has a global market share of 43 %.
[11] For an overview of variables from the data provider, see Table A.1 in the Annex. Table A.2 lists the additional variables created for this paper.

ent pricing methods, it may make sense to examine the online and offline booking channel separately when measuring prices.

Figure 2: Number of offline and online transactions per booking month



Datasets that have not been compiled primarily for the purpose of price statistics may exhibit a multitude of irregularities. The transaction dataset may, for instance, be incomplete or contain incorrect entries. For example, in about 10 % of offline bookings, the holiday destination is missing. There are also cases in which the travel date ($travelDate$) is earlier than the booking date ($transactionDate$). Incorrect entries of this kind are filtered out before the start of the data analysis.[12] In a few cases, the Amadeus dataset also contains obviously incorrect information on prices and holiday duration. For this reason, outliers for prices per person per day are filtered out if they are under €27 or over €427. The bookings filtered out in this way correspond to the top and bottom 1 % of all data sorted by price. Using the same procedure, outliers for holiday duration (1 day < $duration$ < 23 days) are also filtered out. Overall, after outlier adjustment, roughly 3.4 million observations per year remain for holidays in the period from 2013 to 2018.

Data cleansing and outlier adjustment necessary for secondary price sources

In addition to data processing and outlier adjustment, it is also necessary to categorise the unstructured text information in some variables of the (more detailed) online bookings. For example, more than 100 different variations exist for the online variable $mealType$. Across the entire dataset, the number of different variations for the

Additional compression of information using string matching procedure

[12] Cancellations, which are available for offline bookings only, are not included in the analysis.

variable $roomCategory$ is even higher, at 80,000. In order to categorise this level of variety, it is necessary to use string matching techniques like substring search where the categories are defined manually in advance.[13] Identifying children's prices, for which no set definition exists across all tour operators, represents another challenge. While the offline bookings contain information on whether children are part of the booking, and if so, how many ($childrenCount$), for the online bookings an assumption must be made based on the reported ages of the travellers ($travellersAges$). In the following, children were defined as travellers less than 16 years of age.

Total revenue mostly covered by six holiday regions

Measured by total revenue in 2015 (without cruises), the most popular destinations for German travellers are Turkey (23.2 %), the Canary Islands (17.1 %), the Balearic Islands (15.9 %), Egypt (8.9 %), Greece (8.7 %) and the Dominican Republic (3.1 %), as shown in Figure 3. These six areas already account for more than three-quarters of the total revenue. For a regional disaggregation of price dynamics, it therefore makes sense to focus exclusively on these six regions.[14] The revenue shares of the nine next most visited countries are less than 2 % and are all fairly similar in size (range: 1.1 percentage points).[15] Using the transaction dataset, it is possible to derive stylised facts for the German travel market. Based on data for 2015, the typical package holidaymaker travels with one other person (64 %) and without children (80 %), flies from Düsseldorf (16 %), Frankfurt (14 %) or Munich (11 %), stays for 7 or 14 days (35 % and 19 %, respectively) in a four-star hotel (59 %), and pays an average of €92 per person per day.

Price of package holiday depends on how far in advance it is booked

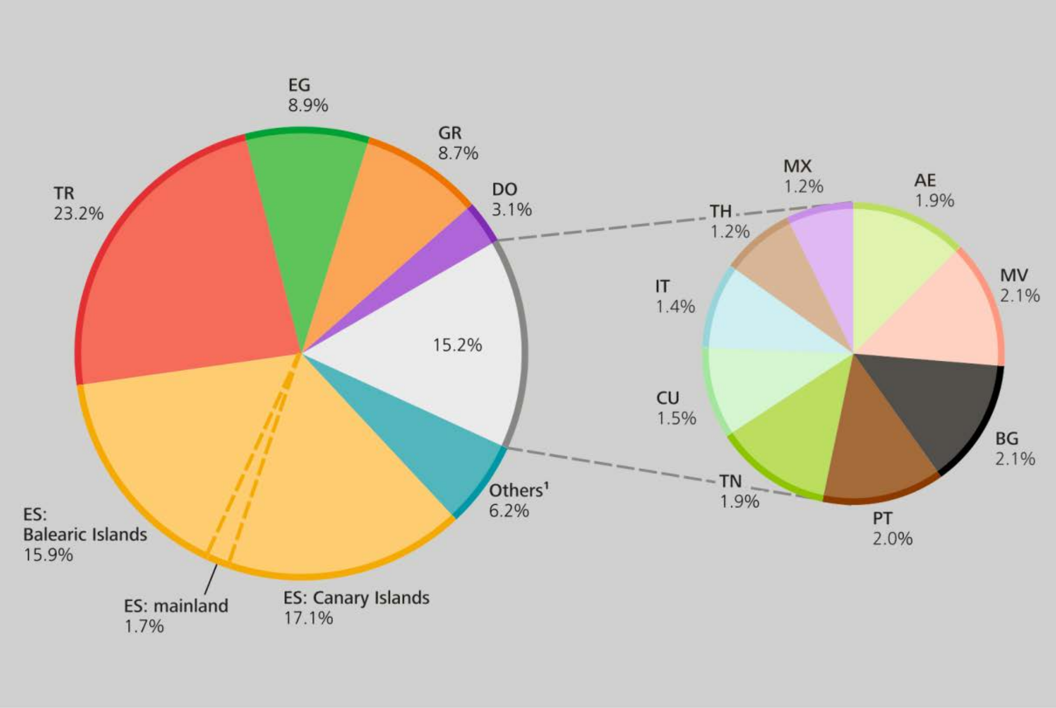
A peculiarity of the HICP package holidays is that bookings can, in principle, be made up to a year before departure and the timing of booking can have an impact on the price. For the period under review, Figure 4 shows that over 20 % of all bookings had already been made half a year prior to the travel month. On average, half of the bookings had already been made three months or more in advance. Twelve and six months before departure, the average price per person per day is 3 % more expensive than its average value of €94, whereas one month prior to departure it falls by about 3 %.

[13] See Table A.3 for a detailed description of how the variable $roomType$ is categorised via a kind of "dictionary". For the production of statistical data, such a dictionary would need to be updated from time to time.

[14] In the following these six major regions will be summarized under the variable $topArea$.

[15] Note that the six major regions' shares of total revenue shifted considerably over the observed period up to 2018. For example, Turkey's share fell by over half from 2013 to 2017, whereas the share of bookings for Greece and the Dominican Republic rose by roughly the same factor. The Spanish regions' share also rose in 2016, but declined in 2018, whereas the share of Egypt – after a severe drop in 2016 – rose by over half from 2013 to 2018. In 2018 Turkey's share recovered, whereas the share of bookings for the Dominican Republic decreased to the level of 2013.

Figure 3: Revenue shares of package holidays by destination in 2015*

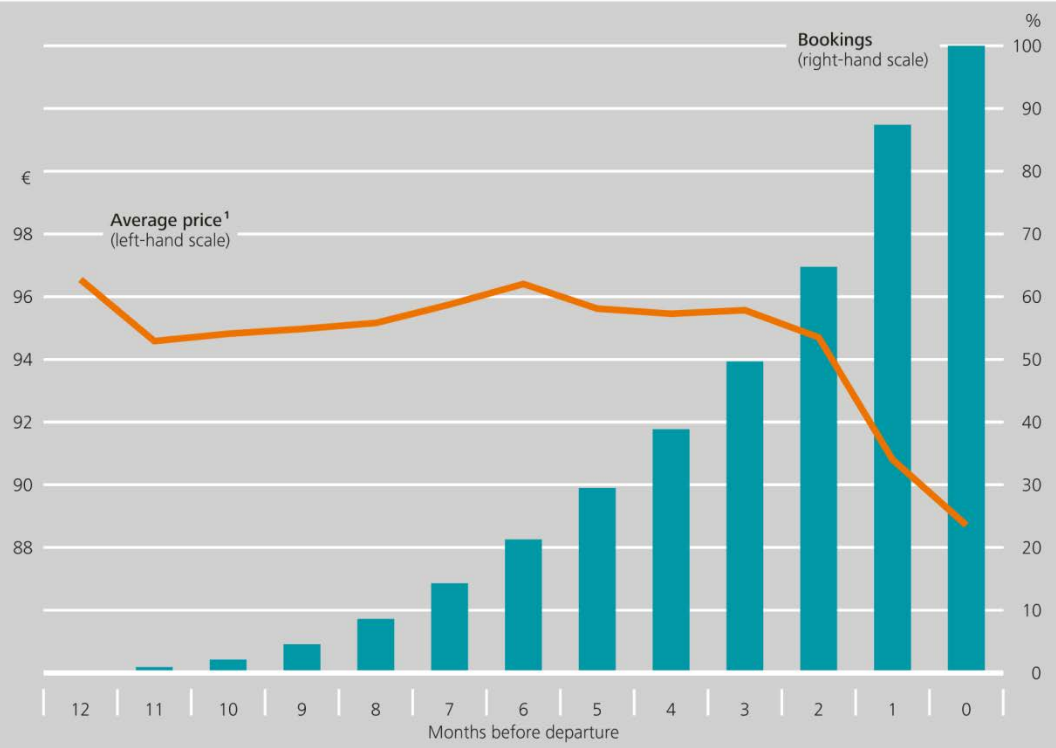


Source: Bundesbank calculations on the basis of booking data from TravelTainment GmbH. * Without cruises. **ES** Spain, **TR** Turkey, **EG** Egypt, **GR** Greece, **DO** Dominican Republic, **AE** United Arab Emirates, **MV** Maledives, **BG** Bulgaria, **PT** Portugal, **TN** Tunesia, **CU** Cuba, **IT** Italy, **TH** Thailand, **MX** Mexico. **1** Holiday regions with a transaction weight of less than 1%.

Deutsche Bundesbank S3IN0503.Chart

Figure 4: Bookings and average price by number of months before departure*

2013 – 2018



Source: Bundesbank calculations on the basis of booking data from TravelTainment GmbH. * Without cruises. **1** € per person and day.

Deutsche Bundesbank S3IN0504A.Chart

4 Methods of price measurement

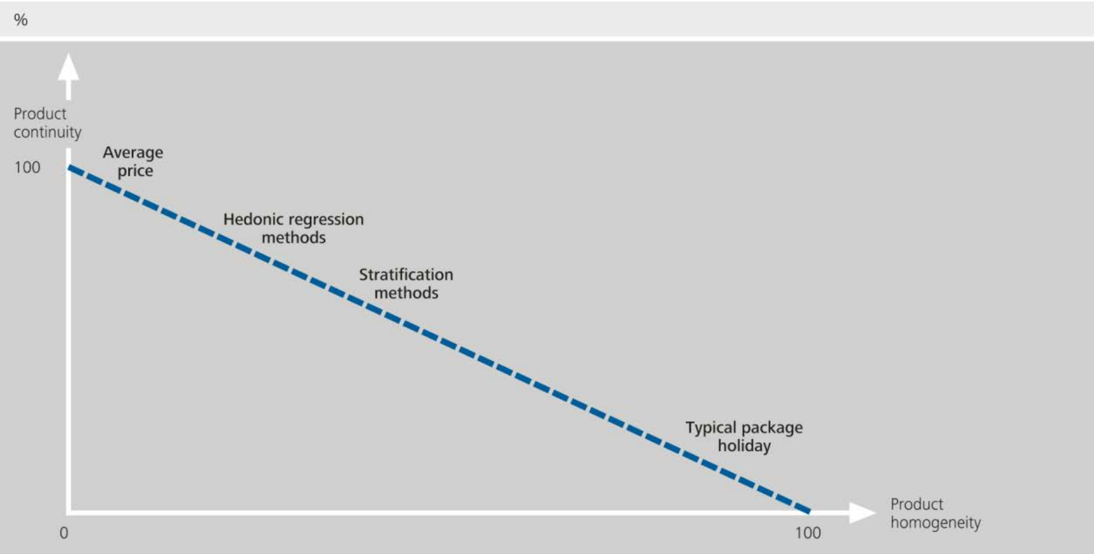
Trade-off between product continuity and homogeneity

An ideal price index would be based on a basket of goods which compares prices of exactly the same product over time. However, transaction price data typically lack a prior product mapping, leaving it to the price statistician to define similar products within a given data set. This process can be thought along the two dimensions “product continuity” and “product homogeneity”, when comparing transactions between any two periods. In the context of package holidays, two extrema are at hand (see Figure 5). A simple *average price* across all bookings would have the highest continuity, i.e. there is a high share of observations used over time. Still, it might be heavily affected by compositional changes in the underlying bookings, and therefore not provide a high degree of homogeneity in terms of comparing similar package holidays. In contrast, the price statistician could only select transactions which correspond to a (pre-defined) *typical package holiday.* This approach coincides basically with the current official practice of collecting prices only for a given price representative. When applied to transaction data, it however does not provide a high number of bookings used over time to have a sound basis for any regional disaggregation.

Two classes of models: hedonic regression and stratification

In the following, Section 4.1 will illustrate an average price with a high continuity of bookings. Consequently, two main approaches in constructing a transaction-based price index are considered, both with the aim to balance between continuity and homogeneity (see Figure 5). A first class of models will be based on *hedonic regression methods*, which estimate a price or index value by controlling for price-determining characteristics (Section 4.2). A second class of models increases the homogeneity of the bookings used by s*tratification methods* (Section 4.3). For each method the transaction data used excludes last minute bookings ($bookTime$ < 14) as well as non-German departure airports ($D(GermanAirport)$ = 0).

Figure 5: Trade-off between continuity and homogeneity



* Product continuity refers to the degree of product match in terms of observations, when comparing a given month with a base period. Product homogeneity refers to the degree of similarity of items within a given product.

Deutsche Bundesbank S3IN0537.Chart

4.1 Unit Value Price Index

The simplest approach to construct a price index is a *unit value price index*, which basically compares average prices over time. In the context of package holidays, the Price per Person per Day (PPD) as given by the variables $travellerCount$ and $duration$ is computed for each transaction. Consequently, the average PPD in a given holiday region is defined by:

Tracking the average price per person per day over time

$$\overline{PPD}_t = \frac{1}{N_t} \cdot \sum_{i=1}^{N_t} \frac{\frac{totalPrice_{i,t}}{travellerCount_{i,t}}}{duration_{i,t}}, \qquad (1)$$

where $i = 1, \ldots, N$ denotes the number of transactions in period t. For comparison purposes, the series of average prices are rebased to $2015 = 100$. The resulting *Unit Value Price index*, I_t^{UV}, in period t is given by:

$$I_t^{UV} = \frac{\overline{PPD}_t}{\overline{PPD}_{2015}} \cdot 100. \qquad (2)$$

The *unit value price index* is often applied in the context of export and import price indices and is suitable for aggregating identical, homogeneous products (see IMF, 2009). However, for more complex or heterogeneous products, this index would suffer from a *selection bias* related to compositional changes in the underlying basket of goods. As an example, suppose for a given holiday region, there would be more (costly) bookings of five-star hotel rooms in period 1 than in base period 0. Even in the case of constant prices, a *unit value price index* would signal a price increase in period 1 simply related to the compositional changes in the hotels booked between both periods. Note that such a scenario would be plausible due to exchange rate movements, for instance, leading to an increase in purchasing power of German travellers. Although the *unit value price index* is not the most sophisticated approach, it is relatively easy to implement for statistical production and uses most of the transaction data, as illustrated in Figure 5.

Approach suffers from the well-known unit value bias

4.2 Hedonic regression methods

Hedonics is a group of regression techniques, which describe the price of a given good or service as a function of several (observed) attributes, each having a marginal contribution to the overall price. In official statistics, hedonics is widely used in order to estimate a quality-adjusted price, e.g. in the context of residential house prices (see Eurostat, 2013). In the following, two different hedonic methods are tested with transaction data of package holidays. The first method is *double imputation* (see Section 4.2.1), where the prices are estimated for the base period as well as the comparison period. The second method is the *time dummy model* (see Section 4.2.2), where the index is directly derived by the coefficient of a time dummy variable in the regression.

Hedonic methods to estimate prices or directly the price index

4.2.1 Double imputation

Concept of double imputation applied to package holidays

Hedonic regression techniques can be used to estimate prices for products which are available in the base period 0 but are no longer available in the comparison period t. Specifically for package holidays, it is hardly possible to observe a booking with exactly the same characteristics in two successive periods. In German official price statistics, *double imputation*[16] is already used for the house price index[17] and some electronic goods such as notebooks or smartphones, where the life cycle of innovative products typically is only a few months. Similarly, package holidays have a high churn, because they rarely can be observed with exactly the same attributes in two successive periods, not least because of the seasonality of holiday destinations. Consequently, the *double imputation* for package holidays is performed by using both observations from the base year 2015 and a given comparison month t to estimate prices.[18] The index is calculated by estimating regression coefficients for the base year and for month t. Consequently, the observations of month t are used to estimate prices for the base year (using the regression coefficients of the base year) and prices for month t (using the regression coefficients of month t). In contrast to electronic goods, for instance, the underlying regression model of package holidays is regarded as stable over a longer period of time, since the price-determining variables rarely change.[19]

Setting up the regression model

The Amadeus dataset contains a lot of price-determining variables, as listed in Table A.1. In a first step, the dataset was examined and a regression model was set up. Model selection was done by analysing adjusted R^2 and its minimum and maximum range. To avoid multicollinearity, the variance inflation factor (VIF) and significance of variables was checked. Another requirement is that the coefficients must be stable and plausible, e.g., a coefficient on a four-star hotel should be smaller than on a five-star hotel, holding all other things equal.[20] Various combinations of characteristics were tested[21] and the following variables gave the best results: $travellerCount, duration, bookTime, channel, star$ and $isHoliday$ (Easter, Pentecost, and Christmas).[22] The adjusted R^2 indicates the explanatory content of the regression model, with an average adjusted R^2 per country between 0.704 and 0.785 the regression models seem to fit well.[23] For $travellerCount$, $duration$ and

[16] Typically, the starting point for the concept of *double imputation* is an A-, B- and C-sample, where the B-sample contains all products that are present in both base period 0 and comparison period t, and products of A- or C-sample are not present in either the base period (C-sample) or the comparison period (A-sample). However, the concept of the A-, B- and C-sample is not applicable for package holidays, since there is no B-sample available. For further details on *double imputation* by the Federal Statistical Office, see Linz et al. (2004).

[17] See Eurostat (2017), section 6.1.2.

[18] In contrast, for electronic goods, January is chosen as a base period and the index is chain-linked annually. This allows an annual adjustment of the regression model to integrate new price-determining features. See Destatis (2009), p. 261.

[19] A change in the regression model for package holidays would for example be necessary if the data provider changes the variables listed in Table A.1.

[20] For a check of the stability of coefficients, see Figure A.1.

[21] A regression model for each holiday destination was set up, but it was also tried to use a regression model for all holiday destinations including dummy variables for all holiday destinations. Using one regression model with joint dummy variables for holiday destinations (e.g. Greece and the Balearic Islands) would make it possible to calculate a whole-year index for Greece as well. However, the results were more plausible when using a single regression model for each holiday destination.

[22] The explanatory variables $depAirport$ and $weekday$ were not significant.

[23] For further evaluations of the adjusted R^2 and other indicators of regression, see Table A.4.

$bookTime$ three transformations were tested (continuous, logarithmised, and grouped by classes), pointing out that the best option is to use logarithmic values for all three variables. Moreover, in estimating a price properly, the *double imputation* method requires to capture the additional effect of public holidays during a given travel month – besides the typical holiday season – on the total price. Therefore, a dummy variable ($isHoliday$) is generated that equals 1 if a public holiday falls between the travel date and the return of a given package holiday and 0 otherwise.[24] Note that $mealType$ and $roomCategory$ are also price-determining variables, but are available for package holidays booked online only. The overall regression model for both online and offline transactions is defined as:

$$\begin{aligned}
&\ln\left(totalPrice_{i,t}\right) = \\
&\beta_0 + \beta_1 \ln\left(travellerCount_{i,t}\right) + \beta_2 \ln\left(duration_{i,t}\right) + \beta_3 \ln\left(bookTime_{i,t}\right) + \\
&\beta_4 \,\mathrm{D}\left(channel_{i,t}\right) + \beta_5 \,\mathrm{D}\left(star_{one_{i,t}}\right) + \beta_6 \,\mathrm{D}\left(star_{two_{i,t}}\right) + \beta_7 \,\mathrm{D}\left(star_{three_{i,t}}\right) + \\
&\beta_8 \,\mathrm{D}\left(star_{five_{i,t}}\right) + \beta_9 \,\mathrm{D}\left(isHoliday_{i,t}\right) + \varepsilon_{i,t}, \qquad (3)
\end{aligned}$$

where equation (3) is estimated for the base year 2015 and each comparison month t separately. Consequently, the Jevons formula is used for index calculation, i.e. the geometric mean of the estimated price relative of period t and base period 0, such that the index value for hedonic regression, I_t^{DI}, reads as follows:

$$I_t^{DI} = \left(\prod_{i=1}^{N} \frac{\hat{p}_{i,t}}{\hat{p}_{i,0}} \right)^{\frac{1}{N}}. \quad (4)$$

More detailed regression specification

As a robustness exercise, a more detailed regression specification was estimated for the *double imputation* method, based on online transactions only. The latter also include information on the meal category such as "all inclusive" and a description of the room type. Consequently, the regression model for online transactions is given as follows:

$$\begin{aligned}
&\ln\left(totalPrice_{i,t}\right) = \\
&\beta_0 + \beta_1 \ln\left(travellerCount_{i,t}\right) + \beta_2 \ln\left(duration_{i,t}\right) + \beta_3 \ln\left(bookTime_{i,t}\right) + \\
&\beta_4 \,\mathrm{D}\left(star_{one_{i,t}}\right) + \beta_5 \,\mathrm{D}\left(star_{two_{i,t}}\right) + \beta_6 \,\mathrm{D}\left(star_{three_{i,t}}\right) + \\
&\beta_7 \,\mathrm{D}\left(star_{five_{i,t}}\right) + \beta_8 \,\mathrm{D}\left(seaView_{i,t}\right) + \\
&\beta_9 \,\mathrm{D}\left(highStandard_{i,t}\right) + \beta_{10} \,\mathrm{D}\left(lowStandard_{i,t}\right) + \beta_{11} \,\mathrm{D}\left(allInclusive_{i,t}\right) + \\
&\beta_{12} \,\mathrm{D}\left(breakfastOnly_{i,t}\right) + \beta_{13} \mathrm{D}\left(isHoliday_{i,t}\right) + \varepsilon_{i,t}, \qquad (5)
\end{aligned}$$

[24] For example, the coefficient on $isHoliday$ is 0.28 for Canary Islands in December 2015, thus the price of a package holiday is about 28 % more expensive in December than in the base year 2015.

where additional dummy variables on the room and meal category were included.[25] The result of this exercise is evaluated in Section 5.

4.2.2 Time Dummy Model

Index value is derived from the time dummy coefficient

The second hedonic method is the *time dummy model*, which also constitutes a regression approach.[26] In contrary to *double imputation*, no prices are estimated, but the index is derived directly from the time dummy coefficient. For the *time dummy model*, the same regression model as in equation (3) is taken, except for $isHoliday$, since the effect of the latter is already included in the time dummy variable. The adapted regression model is as follows:

$$\ln\left(totalPrice_{i,t}\right) = \beta_0 + \beta_1 \ln\left(travellerCount_{i,t}\right) + \beta_2 \ln\left(duration_{i,t}\right) + \beta_3 \ln\left(bookTime_{i,t}\right) + \beta_4 \, \mathrm{D}\left(channel_{i,t}\right) + \beta_5 \, \mathrm{D}\left(star_{one_{i,t}}\right) + \beta_6 \, \mathrm{D}\left(star_{two_{i,t}}\right) + \beta_7 \, \mathrm{D}\left(star_{three_{i,t}}\right) + \beta_8 \, \mathrm{D}\left(star_{five_{i,t}}\right) + \gamma D_{i,t} + \varepsilon_{i,t}, \qquad (6)$$

where $D_{i,t}$ denotes the time dummy which equals 0 for the base period and 1 for the comparison month t. The regression is estimated using all observations from the base period (January) and month t. The *time dummy model* index, I_t^{TD}, is directly derived from exponentiating the coefficient of the time dummy, γ, such that:

$$I_t^{TD} = e^{\gamma}. \qquad (7)$$

The final index series is chain-linked in January by applying the growth rate to the previous index value.[27]

4.3 Stratification methods

An alternative to setting up a regression model consists in dividing a sample into homogeneous strata and to consequently compute an average price within a given stratum. The following sections are dedicated to this *stratification approach.* As a first step, Section 4.3.1 deals with the definition of homogeneous strata or products in the context of package holidays by a quantitative approach. In a next step, Section 4.3.2 presents a traditional bilateral stratification approach based on a comparison of two periods, whereas Section 4.3.3 presents a multilateral approach, the *GEKS* method recently applied to supermarket scanner data, which compares several periods in computing a price index.

[25] See Table A.2 for the list of new defined variables and Table A.3 on the categorisation of the unstructured text information in the variable $roomType$.
[26] See Destatis (2009), p. 259.
[27] Hill (2011) suggests using a correction factor in the index calculation, because of a bias in the price index, which results from the fact that $E[e^{\gamma}] \neq e^{\gamma}$.However, in the present application, the effect of the correction factor was quite small so that the factor was not included in the model.

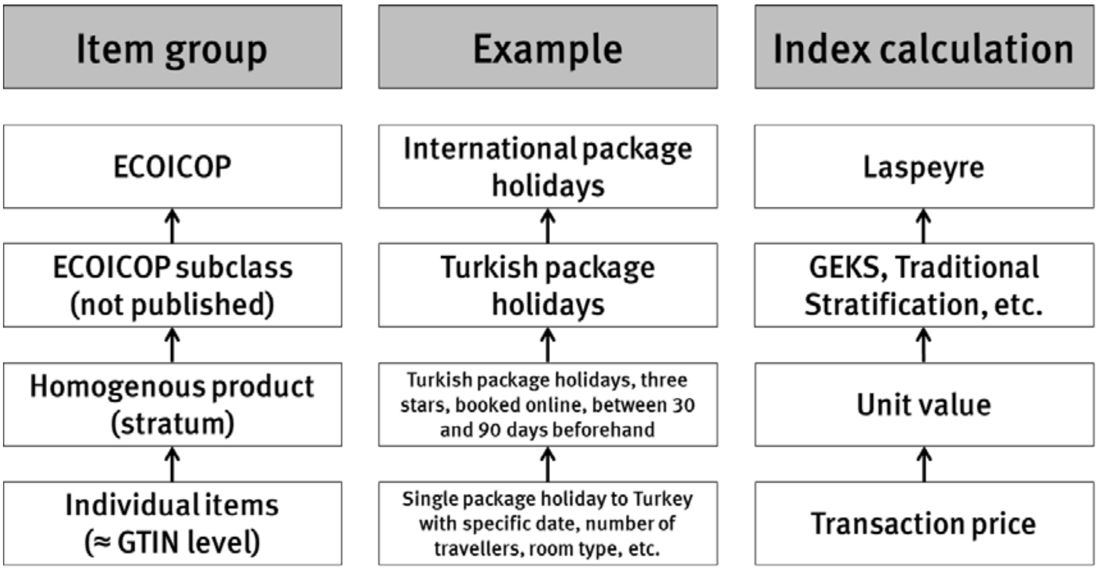
4.3.1 Product definition by a quantitative approach

Setting up the stage: Several items form a product

In price statistics, a proper product definition is key. This is especially true for stratification methods, which group the underlying data according to their price-determining characteristics. Thereby, it is important to distinguish between items and products. More specifically, several items form one product.[28] All items have certain attributes and the question is which attributes are important for product distinction and which ones can be neglected. Obviously, this problem is very much dependent on the product market and especially on the corresponding rate of churn.

Figure 6 illustrates the relationship between items and products in the context of package holidays:

Figure 6: Item hierarchy in the context of package holidays



Source: Own illustration following Chessa (2016)

Product definition in the context of package holidays

The right column of Figure 6 underlines the fact that product definition on a lower level is not connected to the index method since this calculation is performed at a higher aggregation level. In the second column, for illustration purpose, some package holidays would form, for instance, the homogeneous product “Turkish package holidays with three stars and booked online and 30 to 90 days beforehand”. This item group again may form, along with several other ones, an ECOICOP sub-class called “Turkish package holidays”. Note that the Federal Statistical Office currently only publishes at a higher aggregation level (domestic and international package holidays). But if the sample covers sufficient observations, it might be also feasible

[28] This can be made clear in the context of textiles. While a single blue t-shirt of a certain brand with an individual Global Trade Item Number (GTIN, formally known as European Article Number, EAN) is an item, all blue t-shirts of all brands may form the product “blue t-shirt” no matter for example about the fabric or the pattern. This product can be grouped again with t-shirts of other colors and other products to the COICOP group “men’s shirt”.

to publish sub-indices at a more detailed ECOICOP level such as by the category of accommodation or – as tried in this paper – by holiday destinations.

MARS as a balance measure between product homogeneity and continuity

In order to decide on a quantitative level, which price-determining characteristics are important for product definition, Chessa (2018) developed the *Match Adjusted R Squared* (MARS) measure. This measure weighs the two sides of product definition: product homogeneity and product continuity in comparison to a certain base month. Chessa defines homogeneity among a specific product group as the deviation of the average price, whilst assuming that homogeneous items do not vary much in price. Continuity is defined as the share of products that are available in the base period as well as in the current period. These numbers are normalized to one. If for example product definition is only based on the item level (i.e. every single transaction of package holiday), then product homogeneity equals one, but continuity declines as new items appear on the market.[29] Vice versa, if product definition just aggregates all items to one product, the continuity is always one, but homogeneity would equal zero.[30] Chessa suggests multiplying the values for continuity and for homogeneity to obtain a value for MARS, which represents a balance measure between product homogeneity and continuity. This multiplication is similar to a classical loss function since homogeneity increases as continuity decreases and the other way around.

Selection of variables for product definition testing

Note that there are 2^{n} possibilities of combining n different variables for product definition.[31] With regard to reduce this combinatorial problem, not every single variable was analysed but only those variables that had solely a significant influence on $totalPrice$ in the hedonic regression approach (Section 4.2). Seasonal variables like winter and summer season are also excluded.[32] By looking at PPD instead of $totalPrice$, it was possible to omit two variables from the combinatorial problem ($duration$ and $travellerCount$).[33] The variable $bookTime$ was grouped beforehand in order to avoid a too detailed product definition.[34] Moreover, since the shares of one- and two-star accommodations were relatively small in terms of the total revenue in 2015 (less than 1 and 2 %, respectively), these bookings were removed. Moreover, the computation was only performed for the year 2015, which serves also as the base period for the upcoming indices. Finally, the prices of package holidays as given by PPD are stratified according to the variables $topArea$, $star$, $channel$,

---

[29] This assumption is made implicitly for calculating the *double imputation* method (see Section 4.2.1), because no package holiday is grouped with another. Thereby, the lack of continuity is handled by estimating the missing prices.

[30] This assumption is made implicitly for calculating the *unit value price index* (see Section 4.1), because no distinction between items is made.

[31] Moreover, similar to setting up a regression model, one has to consider if grouping a variable is meaningful (e.g. by booking time), making the combinatorial problem even larger.

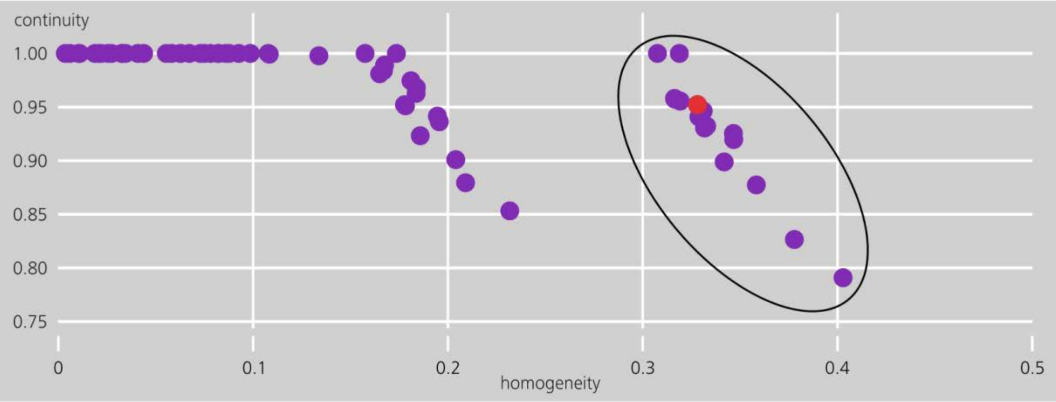
[32] This is due to the fact that seasonal variables create artificial breaks or discontinuities and therefore decrease the value of the continuity dramatically. An alternative could be to stretch the base period to the entire previous year instead of just the previous month. However, this exercise is left for further research.

[33] Alternatively, transactions could be divided also into classes by duration and the number of travellers, which would largely increase the number of groups and might therefore reduce product continuity. It is a fact that for those two variables linearity is not strictly given. Due to simplification and assumable results of the *Unit Value Price Index*, this approach will be used for the stratifications methods anyway.

[34] As shown in Figure 4, the width of possible classes grows by increasing days before departure. So the first group of $\mathrm{bookTime\_Class}$ is from 15 to 30 days, the second from 31 to 90, the third class from 91 to 180, and the fourth class captures all bookings made more than 180 days in advance.

$bookTime\_Class$, $depAirport$ and $weekday$ of departure, yielding $2^6 = 64$ possible combinations to define a product in the context of package holidays.

Figure 7: Continuity and homogeneity of several product definitions following Chessa (2018)



Source: Bundesbank calculations on the basis of booking data from TravelTainment GmbH. The product definition highlighted in red was selected for the subsequent analysis. For illustration purpose, one outlier combination with homogeneity = 1 (i.e., every package holiday represents an own product class) is excluded.

Deutsche Bundesbank S3IN0538.Chart

Picking the best product definition

Figure 7 depicts the average value of continuity and homogeneity for the 64 tested product definitions in 2015. By concept, a combination in the upper right corner, where continuity and homogeneity equal one, would be best. That is why the best definition is to be found in the circled point cloud in Figure 7. Chessa suggests multiplying the monthly values for continuity and for homogeneity in order to obtain MARS.[35] The results for this multiplication are shown in Table A.5 in the Annex. However, in the given application of package holidays, a higher weight would need to be given on the aspect of product homogeneity since the market of package holidays is rather heterogeneous.[36] Additionally, not only the pure values for homogeneity and continuity are to be considered but also the distribution of items per product in order to have enough price representatives per product. For that reason, not the combination with the best value of MARS is chosen but the one that still has high quantitative values of homogeneity and continuity and which is similar to the hedonic regression models above. In Figure 7, this combination is marked in red. Thus, according to the results motivated by Chessa, a product in the context of package holidays is best defined via $topArea$, $star$, $channel$ and $bookTime\_Class$. Moreover, $travellerCount$ and $duration$ are included implicitly by using PPD rather than $totalPrice$ as the dependent variable. Subsequently, these findings define the strata and the data filters that are used in the following two stratification methods.

[35] Note that it is not feasible to multiply the values in Figure 7 in order to calculate MARS, since these represent averages from twelve monthly values in 2015.
[36] In the model from Chessa (2018), this can be thought as a loss function in an additive composition including a parameter λ for manual weighting.

4.3.2 Traditional stratification

Stratification to overcome compositional bias

The *traditional stratification* approach tries to overcome the unit value bias of an average price by grouping transactions into several homogeneous classes before calculating the unit value. In terms of package holidays, transactions that have similar price-determining characteristics are sorted into the same class or stratum. In the following, for each holiday destination as given by $topArea$, the strata are formed by $star$, $channel$ and $bookTime\_Class$, which is consistent with the set of variables approved by the results of the previous section and also the *Hedonic Regression*. The next step is to calculate in each stratum the average PPD in period t (see equation 1) and to normalise the resulting series to $2015 = 100$.[37] In this manner, for each holiday destination, $M = 24$ strata are constructed resulting in 24 elementary price indices, $I_{m,t}^{TS}$.

Aggregation of 24 elementary indices per destination by revenue share in 2015

The aggregation of those elementary price indices to an overall price index for the corresponding destination can be affected by using either a weighted or unweighted mean. In some destinations, certain classes account for only a very small revenue share. For example, there tend to be less package holidays to three-star hotels in Turkey or five-star hotels on the Balearic and Canary Islands, respectively. Thus, an unweighted average price would be biased towards the under-represented classes. For that reason, the weighting is based on the total revenue shares of the individual class in 2015, as given by the transaction data. Finally, for each holiday destination the overall price index according to Traditional Stratification, I^{TS}, in period t is given by:

$$I_t^{TS} = \sum_{m=1}^{M} w_m \, I_{m,t}^{TS}, \qquad (8)$$

where w_m represents the 2015 revenue share of each stratum $m = 1, \dots, M$.

4.3.3 GEKS

Considering also a multilateral method

The origin of the following method goes back to Gini, Eltetö, Köves and Szulc (*GEKS*).[38] Again the sample is stratified as in the *traditional stratification* approach, where unit values for each stratum were calculated and also the price variable is PPD once more. The difference between *GEKS* and the *traditional stratification* lies in the index aggregation. In contrast to the previous section no fixed weights were calculated but the monthly revenue shares where used. Moreover *GEKS* is a multilateral method. *Traditional stratification* is a bilateral method, which compare the current period with a given base period. In contrast, multilateral methods like *GEKS* include also other available periods between, before or even after them into the price index calculation. The main advantage from multilateral methods it that these are transitive and therefore generally free from chain drift.[39] *GEKS* actually combines every considered month to one another in order to get a time series for every one of

[37] Note that this implies a proportional relationship between the total price and both the number of days and the number of travellers. However, a price of a package holiday might be better reflected by a fix-cost (travel-related) component and a non-proportional increase by additional travellers and days. This strict assumption is relaxed in the hedonic regression methods above, which impose a non-linear relationship by using logs.

[38] Eltetö and Köves published an article in 1964 about an index formula that was originally proposed by Gini in 1931. Also in 1964, Szulc published another article about this formula (see OECD and Eurostat 2012, p. 405).

[39] Note that also hedonic regression methods can be constructed in a multilateral way, which is, however, not the case in this paper.

those months. Even though *GEKS* is a rather old method[40], it was revived by Ivancic, Diewert and Fox (2011) in order to face the rising area of scanner data. In particular, in the current month T, *GEKS* compares all months $t = 1, \dots, T$ with the base month 0 using a geometric mean of normalized Fisher indices with changing base periods from 0 to T. Hence, a problem occurs which is characteristic for multilateral methods, which leads to revisions of previous months' index values. However, since price indices should not be revised each month according to the current HICP regulation. That is why Ivancic, Diewert and Fox (2011) propose a chain-link by recalculating the indices for all other months with the help of the new month and applying the growth rate to the new month to the previous index value. Additionally, they propose a rolling window in order to give more recent index values a higher weight in the current index calculation. The *GEKS* index between base period 0 and comparison period t, $I_{0,t}^{GEKS}$, is defined by:

$$I_{0,t}^{GEKS} = \prod_{z=0}^{T} \left(\frac{P_{0,z}^{Fish}}{P_{t,z}^{Fish}} \right)^{\frac{1}{T+1}}, \qquad (9)$$

where $I_{t,z}^{Fish}$ represents the Fisher index between period t and z, whereas T stands for the size of the rolling window. Here, the length of the rolling window was set to 13 months.[41] The Fisher index is given by:

$$I_{t,z}^{Fish} = \sqrt{I_{t,z}^{L} \cdot I_{t,z}^{Pa}} = \sqrt{\frac{\sum_{i=1}^{N_{t,z}} p_z^i q_t^i}{\sum_{i=1}^{N_{t,z}} p_t^i q_t^i} \cdot \frac{\sum_{i=1}^{N_{t,z}} p_z^i q_z^i}{\sum_{i=1}^{N_{t,z}} p_t^i q_z^i}}, \qquad (10)$$

with $I_{t,z}^{L}$ as the Laspeyres index and $I_{t,z}^{Pa}$ as the Paasche index between period t and z. Furthermore, p_t^i and q_t^i denote the price and the number sold of product i in month t. Lastly, $N_{t,z}$ stands for the total number of products that are sold in month t as well as in month z. As for the chaining, the method proposed by Ivancic, Diewert and Fox (2011) was used to obtain non-revised price indices.[42] Note that no dumping-filter was applied, because data cleansing was done beforehand (see Section 3).[43]

---

40 Originally, *GEKS* was used to calculate purchasing power parities.

41 Following De Haan and Krsinich (2014), a window length of 13 months is the smallest that can deal with seasonal products .The window initially starts in January 2014 and ends in January 2015. Greece marks one exception due to the small number of transactions during winter season. Its window initially starts in May 2014 and ends in May 2015.

42 Van Loon and Roels (2018) give an overview about different methods including the one suggested by Ivancic, Diewert and Fox (2011), which is called movement splice. Besides this method the fixed base moving window proposed by Lamboray (2017) was tested. The results were very similar.

43 In German price statistics, the *GEKS* method was already tested on super market scanner data by Bieg (2019).

5 Comparison of results

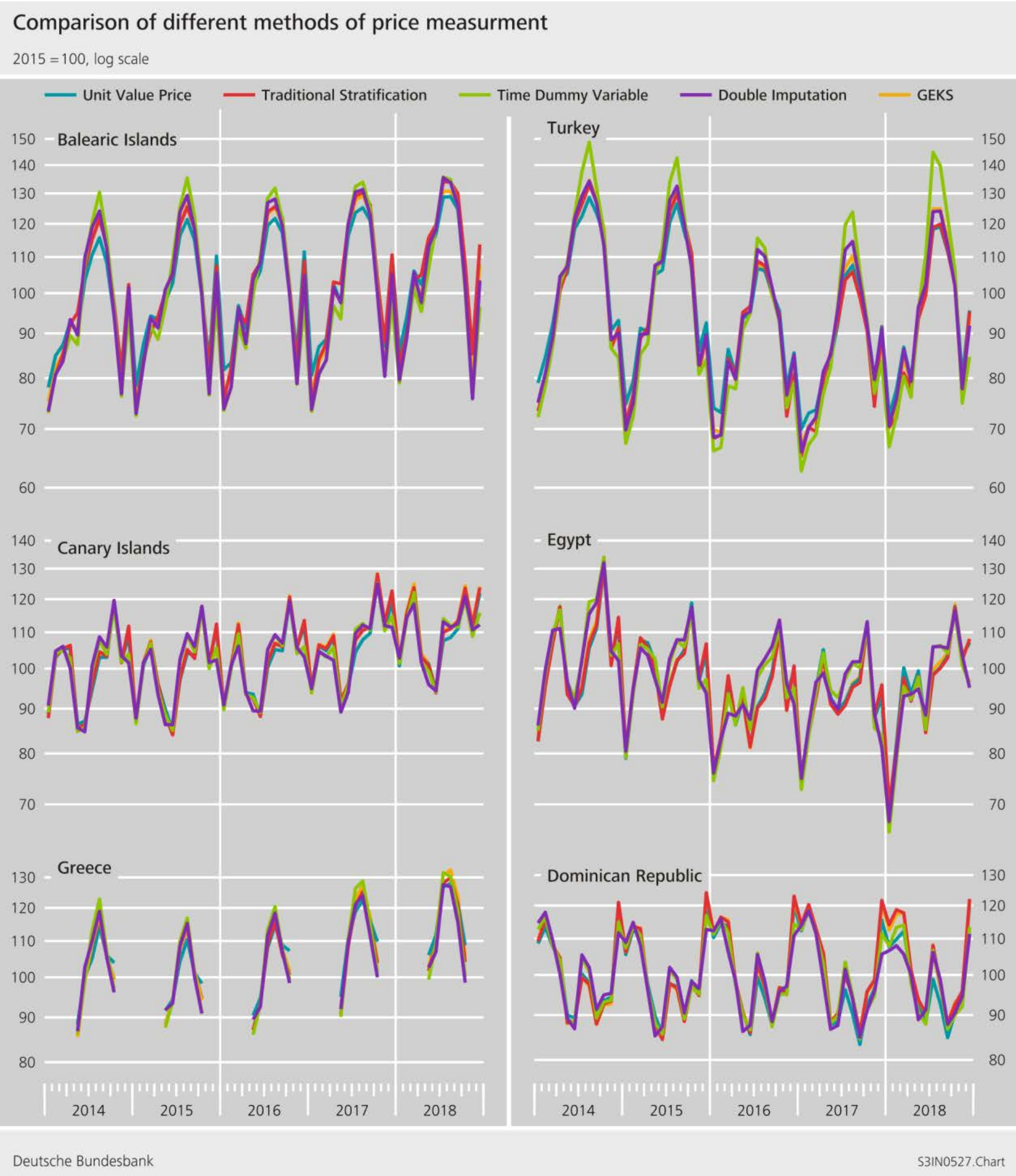
Comparison of price indices based on five different methods

In the following, the price indices based on the five different methods (*unit value price index*, *double imputation*, *time dummy model*, *traditional stratification* and *GEKS*) will be evaluated concerning their seasonal pattern, volatility and robustness. Ideally, the price indices all follow a similar pattern for a given holiday region, so that the selection of the method does not influence the overall movement of the series to a large extent. In this case, the decision on the preferred method could be based on the volatility of the annual rate of change. Moreover, the derived price indices will be compared to the price index according to the current official practice.

Resulting price indices for six major holiday regions

Figure 8 shows the resulting price indices based on the different methods for the six selected holiday regions (Balearic and Canary Islands, Turkey, Greece, Egypt, and Dominican Republic). Overall, the resulting price indices for package holidays in a given region have the same seasonal pattern, with typically higher prices during German summer months and lower prices during winter months. However, there are still some differences across methods within a specific region. For instance, at the end of each calendar year, the price trend for the Canary Islands of *double imputation* differs to the price trends of the other methods. For Egypt, even both methods of *hedonic regression* differ at the end of the year. For Turkey, the *time dummy model* exhibits a higher volatility in comparison to the other methods. For Dominican Republic, the fourth quarter of 2017 and the first quarter of 2018 show differences between almost all methods. Note that although the concept of the *GEKS* as a multilateral index is very different from the bilateral ones, it provides similar results.

Figure 8: Comparison of different methods of price measurement



To have a closer look at the differences in dynamics between methods, the next step is to analyse the annual rates of change. For this purpose, main descriptive statistics are calculated for each method and holiday region. The arithmetic mean (MEAN) indicates whether the price indices have the same trend over time, whereas the standard deviation (SD) as well as the minimum (MIN) and maximum (MAX) indicate the volatility of the annual rates of change. In Table 2, the (absolute) lowest SD, MIN and MAX of a given holiday region are highlighted in green. At a first glance, *traditional stratification* and *double imputation* perform well in terms of these descriptive statistics. The latter exhibits the lowest volatility as indicated by the standard deviation. However, it also appears that the performance of each method seems to depend on the holiday region under consideration. Whereas for the Canary Islands and Egypt, *double imputation* does best, the *traditional stratification* seems to perform well for Balearic Islands and Greece.

Double Imputation and Traditional Stratification exhibit lowest volatility of annual rates of change

Note that the largest variation across methods is found for Dominican Republic, where – in contrast to the other holiday regions – also the sign of the average growth rates (MEAN) differs from each other.

Table 2: Descriptive measures of different index methods by holiday region

Annual growth rates, 2014 - 2018		Unit Value Price	Hedonic Regressions		Stratification	
			Double Imputation	Time Dummy Variable	Traditional Stratification	GEKS
Canary Islands	Mean	2.4	2.1	2.3	2.6	2.8
	SD	4.7	**3.8**	4.8	5.0	**5.0**
	Min	-8.7	**-6.7**	-8.0	-9.6	**-9.3**
	Max	17.1	**14.7**	18.4	**17.5**	18.1
Balearic Islands	Mean	3.2	2.7	2.3	3.3	2.5
	SD	4.7	6.4	**5.3**	**4.3**	4.6
	Min	-8.4	-12.8	**-8.5**	**-7.8**	-8.4
	Max	19.5	25.8	**20.5**	**18.1**	19.0
Turkey	Mean	-2.1	-2.1	-1.8	-2.2	-1.9
	SD	8.2	**8.3**	9.9	**8.8**	9.1
	Min	-16.3	**-16.8**	-21.0	**-17.4**	-18.5
	Max	17.9	**19.6**	21.0	**16.6**	18.2
Greece	Mean	3.4	2.4	2.6	3.2	3.6
	SD	5.5	**5.8**	6.0	**5.3**	5.7
	Min	-5.9	-9.1	**-6.1**	**-5.5**	-5.7
	Max	16.9	**18.2**	18.6	**16.0**	18.2
Egypt	Mean	-1.5	-2.3	-2.6	-1.7	-1.6
	SD	7.9	**7.0**	8.0	**7.3**	7.9
	Min	-19.4	**-15.9**	-18.1	**-17.6**	-18.5
	Max	21.8	**16.9**	20.8	**18.2**	20.9
Dominican Republic	Mean	-0.3	-0.5	-0.3	0.7	0.6
	SD	3.4	3.3	**3.0**	3.4	**3.2**
	Min	-7.7	-8.8	**-7.1**	**-7.9**	-8.1
	Max	8.4	**5.5**	5.8	7.9	**7.8**

Note: Based on the monthly annual growth rates from 2014M1 to 2018M12.

Robustness of data filters and regression specifications

In addition, we performed several robustness tests related to the data itself as well as to the specification of the hedonic regression model. Using all transaction data including last minute bookings as well as non-German departure airports did not affect the index values in a noticeable way. Moreover, excluding bookings with an accompanying child (less than 16 years), which – depending on the tour operator – might lead to a discount on the price of the package holiday, did not affect our results.[44] Finally, estimating the price indices only with the more detailed online data led to very similar results, as shown for the two methods *traditional stratification* and *double imputation* in Figure A.2 and A.3, respectively. In particular, estimating an extended regression model as given in equation (5) did not affect the resulting index of the *double imputation* method. Overall, using the additional information on the meal type (e.g. “all inclusive”) or room category, which plays an important role in setting up a price representative on package holidays, does not seem to change the resulting price index.

[44] For a detailed description of the robustness exercise concerning different data sets, see Table A.6.

Comparison of transaction-based indices with current national practice

For a first approximate comparison of the transaction-based indices with the official price index of international package holidays (ECOICOP 09.6.0.2), the latter have to be aggregated to an overall index, since official results at the level of holiday regions are currently not published. As described in Section 2, the official price index consists of six holiday destinations for international flight package holiday, city trips and cruises. City trips and cruises are not calculated with Amadeus transaction data, instead the official (confidential) index values are used.[45] Similarly, the transaction-based indices for Greece and cruises during winter season are imputed by using all available indices (all-seasonal estimation). For the Dominican Republic, the official index imputes the summer months whereas the methods based on transaction data do not need any imputation for that holiday region.[46] For all investigated methods, a corresponding total index for international package holidays is calculated by summing up the sub-indices using the official (confidential) weights.

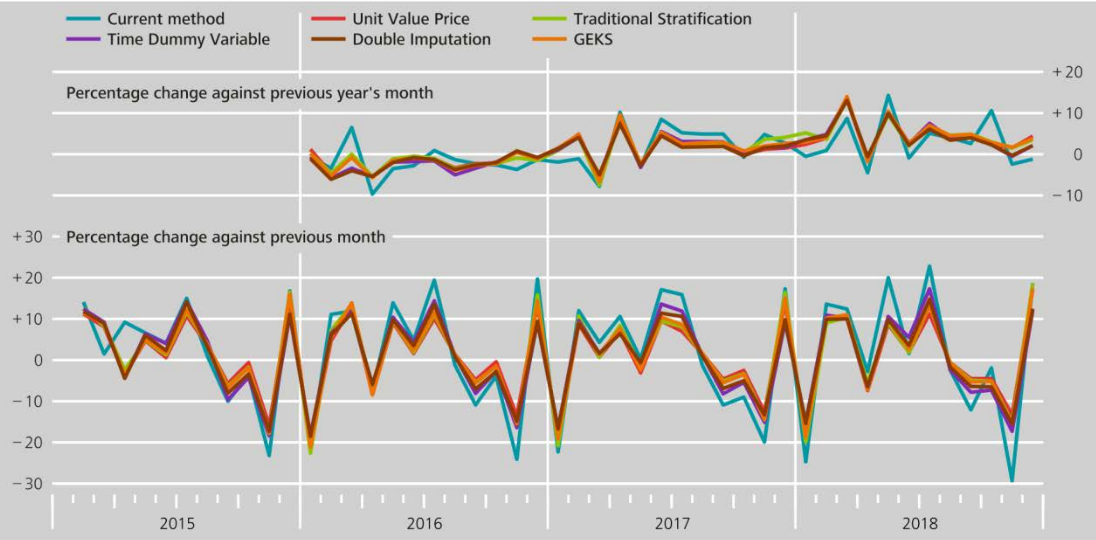
Comparison of the rates of change

Figure 9 depicts the growth rates of the resulting five indices together with the current official index. Note that a comparison of the latter can be only made from 2016M1 onwards, since a new official index computation method was recently introduced back to 2015M1. Concerning the annual rate of change as shown in the upper bottom of Figure 9, there are only four periods (out of a total of 36 periods), when the dynamics of the five methods diverge from each other. In contrast, the official method deviates in eleven out of 36 periods from the sign of the rate of growth indicated by the majority of transaction methods. When concerning the monthly rate of change from 2015M2 onwards, the five methods do not differ in any of the 47 period in terms of the sign of the rate of change, reflecting the dominance of the seasonal pattern in the series. The official method deviates only in four out of 47 periods. Finally, the descriptive statistics on the annual growth rates are shown in Table 3. Evidently, all methods have a smaller standard deviation than the official method. Concerning the different indices, *double imputation* has the lowest standard deviation; however, the difference across methods falls within a rather small range.

[45] Concerning cruises, in the transaction data, there is only information on the destination airport, but not on the room category (e.g., inside or outside cabin), which is obviously an important price determinant when booking a cruise. City trips might be calculated with the underlying transaction data, but is left for further research.

[46] This does not only affect the price movement of the Dominican Republic but also indices for Greek and cruises, because out-of-season months are imputed by the all-seasonal estimation.

Figure 9: Comparison of transaction-based pseudo indices with current sub-index "International Package Holidays" (ECOICOP 09.6.0.2)



Source: Destatis; Bundesbank calculations on the basis of booking data from TravelTainment GmbH. For each transaction-based method, the elementary indices on six holiday regions were aggregated together with the official (confidential) elementary indices "city trips" and "cruises" by using the official fixed-weighting scheme.

Deutsche Bundesbank S3IN0535A.Chart

Table 3: Comparison of transaction-based methods with current national practice (percentage change against previous year's month)

	HICP Int. package holidays (09.6.0.2)	Unit Value	Double Imputation	Time Dummy	Traditional Stratification	GEKS
MEAN	4.3	3.1	3.2	3.5	3.3	3.4
SD	5.3	4.3	**4.1**	4.5	**4.3**	4.4
MIN	-9.7	-6.4	-6.1	**-5.8**	-7.5	**-6.7**
MAX	14.3	13.7	**13.1**	13.6	**13.4**	14.0
Q 0,25	-2.5	-1.4	-1.8	**-1.8**	**-1.2**	-1.5
Q 0,75	4.9	4.2	**3.5**	3.8	4.2	**4.0**

Note: Based on the monthly annual growth rates from 2016M1 to 2018M12, since the method of the official price index changed in 2015. For each transaction-based method, the elementary indices on six holiday regions were aggregated together with the official (confidential) elementary indices "city trips" and "cruises" by using the official weighting scheme.

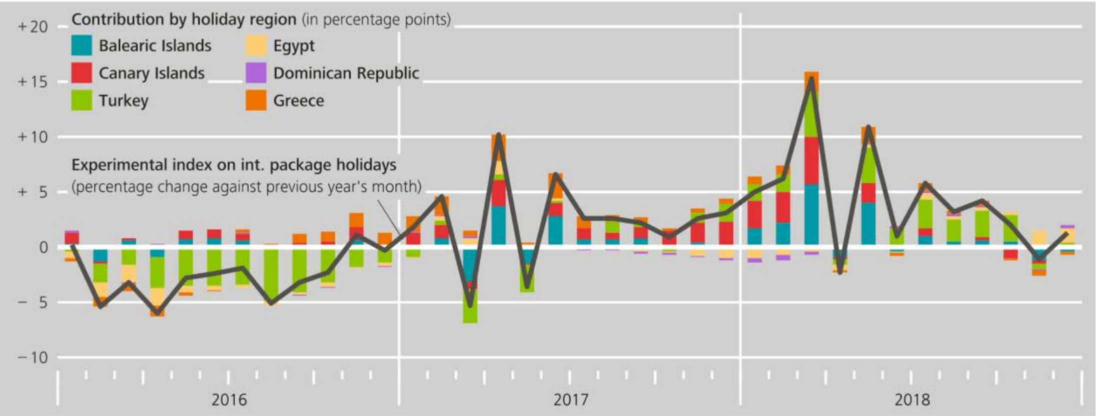
Pure price comparison in the current official method

Concerning the comparison between transaction and offer prices, some remarks have to be made: The current official method follows a pure price comparison of identical price offers over time by tracking the same booking code in each month. That means quality changes should not influence price development. The methods based on transaction data also try to compare like with like but define an identical product in a broader way. For example, in the baseline regression model in Section 4.2.1, the variable $mealType$ is not included whereas the current official practice controls for this information as well. Concerning the data as itself, the transaction data of Amadeus may bring along structural shifts and substitution effects, such as changing weights of booking times, hotels and room categories over the year.

Creating an experimental index for further analysis

Finally, the compilation of regional price indicators allows for a more detailed economic interpretation in the aggregate index of international package holidays. Figure 10 plots an experimental index based on the *double imputation* method, by aggregating the six regional price indicators with their average revenue share from 2015-2016. The resulting series differ from the previously shown indices in Figure 10, because they do not contain the official sub-indices on cruises and city trips. It becomes clear that the negative price trend in 2016 as well as the recent peak in summer 2018 in international package holidays was primarily driven by the developments in Turkey. During the beginning of the sample, the latter experienced a decline in bookings as a response to several terroristic attacks and increasing political uncertainty, with bookings recovering in the summer season 2018. Obviously, this was accompanied by a similar movement in prices for Turkey. Due to the resulting shift in German travellers' preferences, the Balearic and Canary Islands and, to a lesser extent, Greece, could at the same time increase their prices for package holidays during 2017 and 2018.[47]

Figure 10: Experimental index for international package holidays and contributions from holiday regions



Source: Bundesbank calculations on the basis of booking data from TravelTainment GmbH. The experimental index is based on the "double imputation" method and consists of six holiday regions, which are weighted together by their respective revenue share (2015-2016 average).
Deutsche Bundesbank S3IN0536.Chart

[47] See also Section 3 on the revenue shares per holiday destinations over time. Note that, in calculating the contributions to growth, the weight of a given holiday region was held constant (average 2015-2016 revenue share).

6 Summary

A transaction-based price indicator is feasible, but requires further research at the micro level

This paper showed that, by means of transaction data, it is possible to calculate efficiently several experimental price indices that can be disaggregated by holiday regions and therefore allow interpreting movements in the overall index of international package holidays. All five methods under consideration follow a similar pattern, from which the official price index based on offer prices deviates at some points in time. Before valuing those deviations, an important issue has to be noted. Offer data allow for a pure price comparison by tracking the same price offer over time. It is not yet clear to what extent transaction data perform sufficiently well in terms of varying sample and quality adjustment, notably regarding room type. If, for example, travellers tend to book different room types during summer months than in the winter months, the demand changes over time. Therefore, for the underlying transaction data, further analysis at the micro level on booking characteristics is needed.

Way forward

Note that using transaction data might not be the only way to accomplish a regional disaggregation. The Federal Statistical Office is currently pursuing the approach of collecting offer prices automatically from the computer reservation system Amadeus. Thus, a future disaggregation by holiday destination could also be possible by offer data. Nonetheless, in the case of offer data, collected prices still need to be aggregated by external weight information, e.g. on booking time. In this sense, transaction data, which contain weight information on a very detailed level, are more convenient. The question is whether in the case of transaction data, the impact of structural effects could be minimised by acquiring more information, for instance on the room type for offline bookings. Additionally, transaction data from other global distribution systems or even from the touristic operators themselves could make the performed analysis all the more robust. Consequently, a deeper disaggregation could be possible, e.g. price indices by different accomodation categories.

References

Bieg, M. (2019), „Nutzung von Scannerdaten in der Preisstatistik – Eine Untersuchung anhand von Marktforschungsdaten“, WISTA – Wirtschaft und Statistik 2/2019, pp. 25-38.

Chessa, A.G. (2018), “Product definition and index calculation with MARS-QU: Applications to consumer electronics”, mimeo.

Chessa, A.G. (2016), “Processing scanner data in the Dutch CPI: A new methodology and first experiences”, EURONA No. 1/2016, pp. 49-69, online at: https://ec.europa.eu/eurostat/cros/system/files/euronaissue1-2016-art2.pdf.

De Haan, J. and F. Krsinich (2014), “Time Dummy Hedonic and Quality-Adjusted Unit Value Indexes: Do They Really Differ?”, Conference paper at Society for economic measurement in Chicago, USA, August 2014.

Destatis (2009): *Handbook on the application of quality adjustment methods in the Harmonised Index of Consumer Prices*, Statistics and Science Volume 13, Federal Statistical Office of Germany.

Deutsche Bundesbank (2017), “The volatility of the traditional core inflation rate in Germany”, in: Monthly Report, November 2017, pp. 49-51.

European Commission (2009), COMMISSION REGULATION (EC) No 330/2009 of 22 April 2009 laying down detailed rules for the implementation of Council Regulation (EC) No 2494/95 as regards minimum standards for the treatment of seasonal products in the Harmonised Indices of Consumer Prices (HICP), online at: https://publications.europa.eu/en/publication-detail/-/publication/479d86f4-5e76-44b6-9e43-f33f9c2b1888/language-en.

Eurostat (2018), *HICP Methodological Manual*, Luxembourg: Publications Office of the European Union, online at: https://ec.europa.eu/eurostat/documents/3859598/9479325/KS-GQ-17-015-EN-N.pdf.

Eurostat (2017), *Technical manual on Owner-Occupied Housing and House Price Indices*, online at: https://ec.europa.eu/eurostat/documents/7590317/0/Technical-Manual-OOH-HPI-2017/

Eurostat (2013), *Handbook on Residential Property Price Indices (RPPIs),* Luxembourg: Publications Office of the European Union, online at:

https://ec.europa.eu/eurostat/documents/3859598/5925925/KS-RA-12-022-EN.PDF.

Hill, R. (2011), "Hedonic Price Indexes for Housing", OECD Statistics Working Paper No. 2011/01, online at: http://dx.doi.org/10.1787/5kghzxpt6g6f-en.

ILO, IMF, OECD, UNECE, Eurostat, and World Bank (2004), *Consumer Price Index Manual: Theory and Practice*, ed. by Hill, P., Geneva, Switzerland: International Labour Office, online at: https://www.ilo.org/public/english/bureau/stat/download/cpi/cpi_manual_en.pdf.

Ivancic, L., W.E. Diewert and K.J. Fox (2011), "Scanner Data, Time Aggregation and the Construction of Price Indexes", in: *Journal of Econometrics* 161, 24-35.

IMF (2009), *Export and Import Price Index Manual: Theory and Practice*, online at: https://www.imf.org/external/np/sta/xipim/pdf/xipim.pdf.

Lamboray, C. (2017), "The Geary Khamis index and the Lehr index: how much do they differ?", Paper presented at the 15th Ottawa Group meeting, Eltville am Rhein, Germany, 10-12 May 2017, online at: http://www.ottawagroup.org/Ottawa/ottawagroup.nsf/4a256353001af3ed4b2562bb00121564/1ab31c25da944ff5ca25822c00757f87/$FILE/The%20Geary%20Khamis%20index%20and%20the%20Lehr%20index%20how%20much%20do%20they%20differ%20-%20Claude%20Lamboray%20-Paper.pdf.

Linz, S., T. Behrmann and U. Becker (2004): „Hedonische Preismessung bei EDV-Investitionsgütern", WISTA – Wirtschaft und Statistik 6/2004**,** pp**.** 682-689.

OECD and Eurostat (2012), *Eurostat-OECD Methodological Manual on Purchasing Power Parities*, Luxembourg, online at: https://www.oecd.org/sdd/prices-ppp/PPP%20manual%20revised%202012.pdf.

Van Loon, K. and D. Roels (2018), "Integrating big data in the Belgian CPI", Meeting of the Group of Experts on Consumer Price Indices, Geneva, Switzerland, 7-9 May 2018, online at: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/Belgium.pdf.

Annex

Table A.1: Description of variables in Amadeus dataset

Variable	Description	Online	Offline	Type
Information on accommodation				
iffCode	Numeric identifier of the accommodation booked	Y	Y	numeric
accomCategory	Classification of the standard of the accommodation (star rating)	Y	Y	numeric
accomName	Name of accommodation (e.g. "Sea View Hotel")	Y	Y	text string
isCruise	Accomodation represents a cruise ("Y" or "N")	Y	Y	categorical
Information on holiday destination				
accomLocation	Location (lowest level of geography) of the accommodation (e.g. Playa de Palma)	Y	Y	text string
accomProvince	Area of the accommodation (e.g. Balearic Islands)	Y	Y	text string
accomCountry	Country of the accommodation area (e.g. Spain)	Y	Y	text string
Information on flight				
travelDate	Date on which travel is booked to start	Y	Y	date
depAirport	3 letter IATA code of departure airport	Y	Y	alphanum.
destAirport	3 letter IATA code of destination airport	Y	Y	text string
Information on booking process				
tourOperatorId	Numeric identifier of tour operator	Y	Y	numeric
channel	Source of the booking ("Online" or "Offline")	Y	Y	categorical
status	Status of the booking ("Booked" or "Cancelled")	Y	Y	categorical
transactionDate	Date on which the booking is made	Y	Y	date
postcode_travelAgency	Post code of traditional high street travel agency	N	Y	numeric
Information on travellers				
travellerCount	Number of travellers on the booking	Y	Y	numeric
childrenCount	Number of children and infants on the booking	N	Y	numeric
travellerAges	List of ages of each of the travellers	Y	N	alphanum.
Information on transaction price				
totalPrice	The selling price of the booking expressed in EUR	Y	Y	numeric
duration	Length of the travel expressed as a number of days	Y	Y	numeric

mealType	A classification of the level of service provided at the accommodation (e.g. "all inclusive")	Y	N	alphanum.
roomCategory	Description of the accommodation booked (e.g. "with sea view")	Y	N	alphanum.
hasTravelInsurance	Total price includes travel insurance ("Y" or "N")	Y	N	categorical
hasHireCar	Total price includes car hire ("Y" or "N")	Y	N	categorical

Table A.2: Description of new defined variables

Variable	Description	Type
travelMonth	Month of travelDate	numeric
bookingMonth	Month of transactionDate	numeric
bookTime	Difference between travelDate and transactionDate in number of days	numeric
bookTime_Class	bookTime divided into four classes (up to 30, between 31 and 90, between 91 and 180, higher than 180)	numeric
PPD	Price per person per day	numeric
children	Number of children (offline) and travellers less than 16 years of age (online)	numeric
star	accomCategory divided into five classes (one to five stars)	numeric
D(star_one) to D(star_five)	Dummy variables for a given star category (1 or 0)	categorical
D(online)	Online booking only (1 or 0)	categorical
D(GermanAirport)	destAirport is located in Germany (1 or 0)	categorical
topArea	Balearic Islands, Canary Islands, Turkey, Greece, Egypt or Dominican Republic	alphanumeric
D(doubleRoom)	Indicator variable (see Table A.3)	categorical
D(seaView)	Indicator variable (see Table A.3)	categorical
D(highStandard)	Indicator variable (see Table A.3)	categorical
D(lowStandard)	Indicator variable (see Table A.3)	categorical
D(allInclusive)	Indicator variable on whether $mealType$ is “all inclusive” or “Vollpension” (1 in both cases) or not (0)	categorical
D(breakfastOnly)	Indicator variable on whether $mealType$ includes breakfast only or not (1 or 0)	categorical
D(isHoliday)	Easter, Pentecost or Christmas in the course of travel (1 or 0)	categorical
weekday	Weekday of departure date (Monday, …, Saturday, Sunday)	categorical

Table A.3: Categorisation of variable "roomCategory"

Indicator variable	double room	high standard	low standard	sea view
text string	2-zimmer	deluxe	spar	meers
	2 zimmer	superior	eco	mb
	dz	penth		meerb
	2 raum	villa		sea view
	2 räume			seaview
	doppel			meer-u
	zweizimmer			
	zweibett			
	double room			
	doubleroom			
	2er			
	2 be			

Note: The indicator variable equals 1 if the variable roomCategory (converted into lowercase letters) contains one of the pre-defined text strings, and 0 else. The text strings are defined according to the most frequent entries (top-100 values).

A.3: Stability of regression coefficients

Negative coefficient for online and positive coefficients for other variables are plausible

As a necessity to the regression model in Section 4.2, it must be ensured, that the resulting coefficients are stable and have a plausible economic interpretation. Coefficients of the *double imputation* model for each of the six holiday destinations are shown in Figure A.1. On the left side, there are the coefficients for the variables $travellerCount$, $duration$, $bookTime$, D_online and $isHoliday$. As expected, all coefficients are positive, i.e. the price of a package holidays increases with the number of travellers, the duration, the more days the package has been booked in advance and during public holidays. One exception is for online bookings, signalling that a package holiday booked online is on average 8.4 to 11.9 % cheaper (depending on the holiday region) than booked via a stationary travel agency. Concerning volatility over time, it has to be kept in mind that package holidays have a seasonal pattern, which will be reflected in volatile coefficients and partly also a seasonal pattern.

Coefficients for accomodation category have the correct order

The right side of Figure A.1 shows the coefficients on the accomodation category of the underlying hotel, as indicated by one up to five stars. The benchmark in the regression model (3) is a four-star hotel, so five-star hotels are on average expected to be more expensive, whereas one- to three-star hotels are expected to be cheaper. This condition is fulfilled for nearly all holiday destinations. Besides this, the coefficient of a three-star hotel should be on average higher than a two-star hotel, and a coefficient of a two-star hotel higher than a one-star hotel. For most holiday destinations, this is true, but one- and two-star hotels are not common for all holiday destinations and therefore have only a small number of observations. This is reflected in the coefficients of one-star hotels, which are not stable for the Canary Islands and

Turkey: for some month they are higher than for two-star hotels or even positive, and they have also missing values. The same problems occur for two-star hotels in Egypt and Dominican Republic. For example the standard deviation of the coefficient on a two-star hotel in Egypt is higher ($\sigma = 0.09$) than for a three-star hotel ($\sigma = 0.02$) or for five-star hotel ($\sigma = 0.05$). The volatility of some regression coefficients (e.g. two-stars-hotel in Egypt) has only a minor effect on the index, because its implicit weight is very small. In statistical production, the regression model could be adapted and optimized for each holiday region. To sum up, most of the coefficients are stable and show a similar seasonal pattern.

Figure A.1: Stability of regression coefficients over time (*Double Imputation* method)

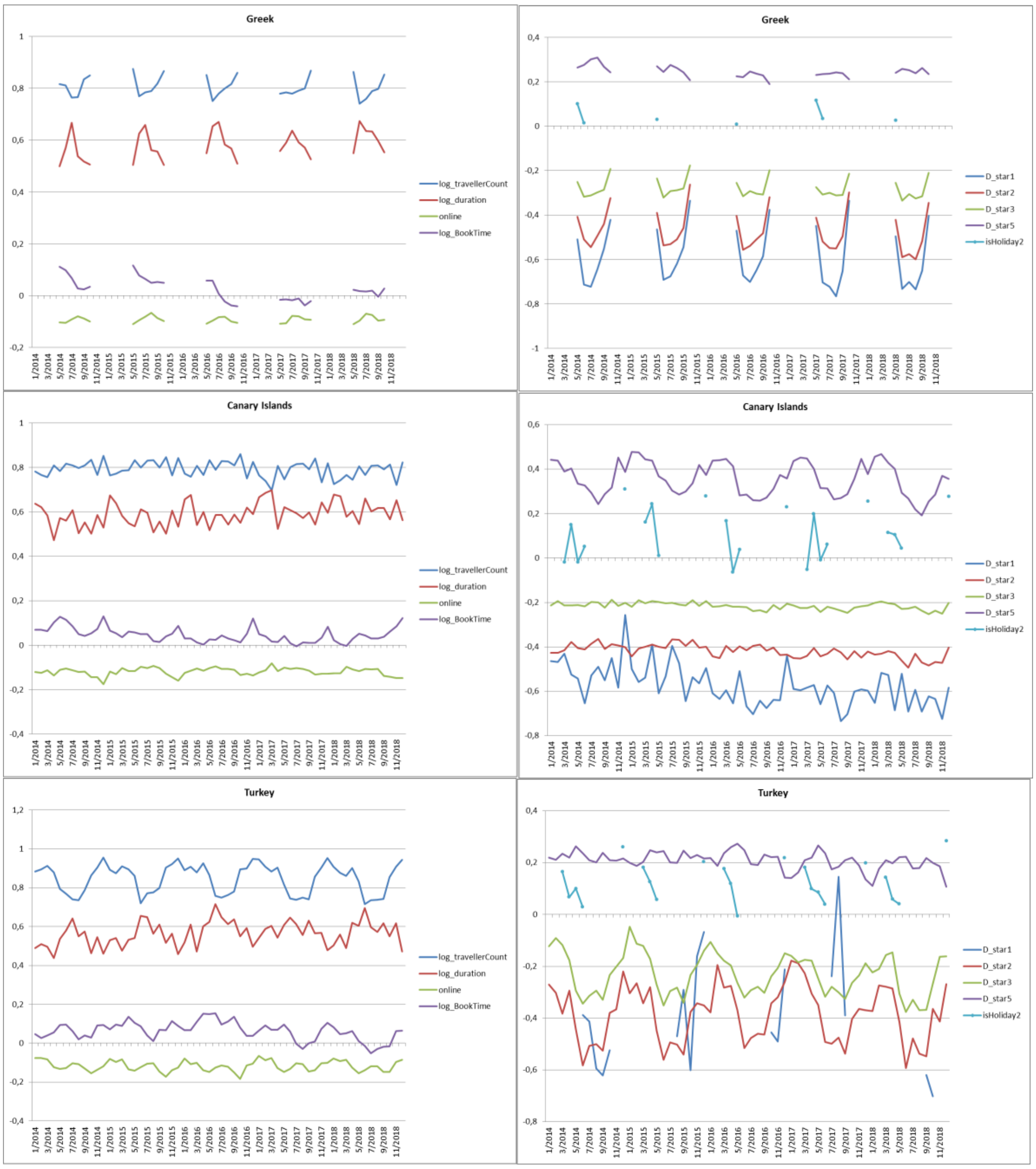


Table A.4: Adjusted R^2 by holiday region

Region / Method	Double Imputation		Time Dummy Model	
	Mean	Max-Min Range	Mean	Max-Min Range
Balearic Islands	0.769	0.161	0.730	0.206
Canary Islands	0.721	0.113	0.677	0.118
Turkey	0.772	0.147	0.794	0.092
Greece	0.753	0.118	0.695	0.156
Egypt	0.704	0.205	0.717	0.175
Dom. Republic	0.785	0.121	0.720	0.099

Table A.5: Top-ten results of MARS for product definition of package holidays

No. of Combination	topArea	star	channel	bookTime _Class	depAirport	weekday
1	1	1	1	1	1	1
2	1	1	1	1		1
3	1	1	1	1	1	
4	1	1		1	1	1
5	1	1	1		1	1
6	1	1	1			
7	1	1	1			1
8	1	1	1	1		
9	1	1		1		1
10	1	1			1	1

No. of Combination	Number of products	Mean of items per product	MARS	Homogeneity	Continuity
1	10681	193.2	0.33	0.40	0.79
2	1008	2047.5	0.33	0.35	0.93
3	1582	1304.6	0.32	0.35	0.92
4	5423	380.6	0.32	0.38	0.83
5	2752	750.0	0.32	0.36	0.88
6	36	57330.1	0.32	0.32	1.00
7	252	8190.0	0.32	0.33	0.95
8	144	14332.5	0.32	0.33	0.95
9	504	4095.0	0.32	0.33	0.93
10	1379	1496.7	0.31	0.34	0.90

Note: This table shows the top-ten results from MARS following Chessa (2018). The values of MARS are calculated as the average of twelve monthly MARS values in 2015. Combination 8 (highlighted in green) was taken for the main analysis in this paper which inherits a high mean of items per product, suggesting having enough price representatives for a bias-free index calculation.

A.4: Robustness of data filters and hedonic regression specification

Table A.6: Robustness of underlying data set

Data filters used / Data set	R1	R2*	R3	R4
excluding outliers as defined by the price per person per day and *duration*	X	X	X	X
German departure airports only		X	X	X
travellers > 16 years				X
excluding last minute bookings (within 14 days before departure)		X	X	X
Online transactions only			X	

Note: R2 denotes the baseline data set used in the main analysis of the paper. R3 (online data only) also includes a more detailed regression equation for *Double Imputation*, as shown in Equation (5).

Figure A.2: Comparison of the monthly annual growth rates for *Traditional Stratification* using different datasets

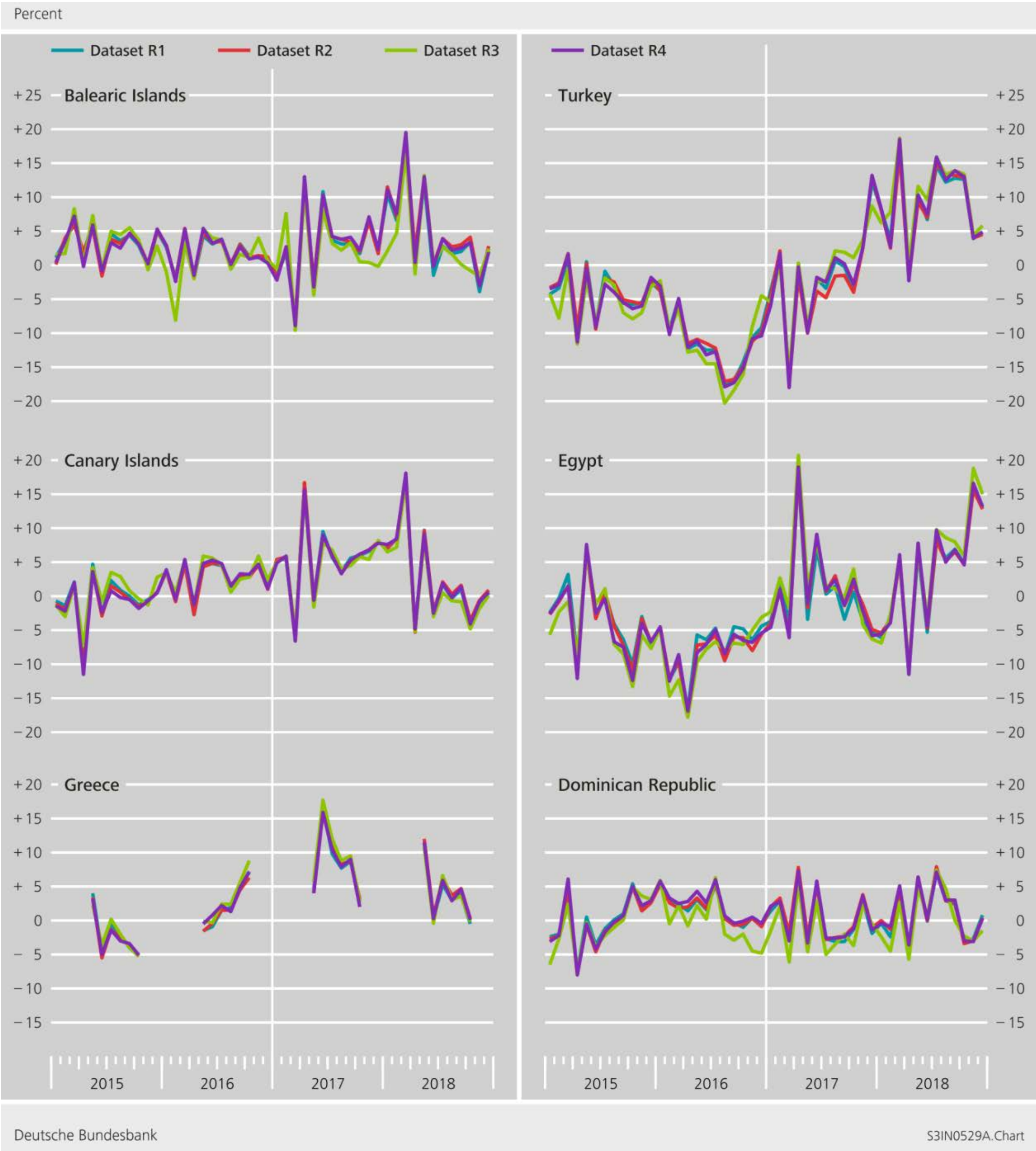


Figure A.3: Comparison of the monthly annual growth rates for *Double Imputation* using different datasets