Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

Working Paper Series

07/22

FORECASTING DOMESTIC TOURISM ACROSS REGIONAL DESTINATIONS THROUGH MIDAS REGRESSIONS

CONSUELO R. NAVA, LINDA OSTI and MARIA GRAZIA ZOIA

The Department of Economics and Statistics "Cognetti de Martiis" publishes research papers authored by members and guests of the Department and of its research centers. ISSN: 2039-4004

Forecasting domestic tourism across regional destinations through MIDAS regressions

Consuelo R. Nava[1*], Linda Osti[2] and Maria Grazia Zoia[3]

[*] *Corresponding author*

[1] *Department of Economics and Statistics "Cognetti de Martiis", University of Turin, Lungo Dora Siena 110A, Turin, Italy. Corresponding Author. Tel:+390116704406. Fax:+390116703895* *consuelorubina.nava@unito.it*

[2] *Faculty of Economics and Management, Free University of Bozen-Bolzano, Brunico, Italy,* *linda.osti@unibz.it*

[3] *Department of Economic Policy, Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20123, Milano, Italy.* *maria.zoia@unicatt.it*

Abstract

Over the years, benefits of domestic tourism have been shadowed by the exponential growth of international tourism, despite the former representing a crucial resource, especially at times of geopolitical instability and pandemics. Therefore, forecasting domestic tourism across different regions and sub-regions becomes fundamental to determine its viability as a substitution of international tourism during the COVID-19 pandemic and to evaluate the effectiveness of governmental incentive policies introduced for its promotion. To this aim, and given the availability of data sampled at different frequencies, mixed data-sampling (MIDAS) models have been employed to estimate and predict domestic tourism expenditures, arrivals, and overnight stays. To this aim, we consider the specific case of Italy for illustrative purposes.

Keywords: Domestic tourism; tourism indicators; MIDAS; U-MIDAS; forecast; COVID-19.
JEL CODE: C01; C53; L83

1 Introduction

Tourism represents for both advanced and emerging economics a channel to gain economic development, helping to create jobs and promote entrepreneurship for an extended range of business activities. Till the worldwide lockdowns due to COVID-19 pandemic, international tourism was acknowledged as one of the main economic activities representing a key component in export diversification capable of reducing trade deficits and compensating weaker exports of other goods and services (UNWTO, 2021a). In this perspective, over the past few decades, Italian tourism policies have focused on international tourism and, in 2019, the last year before the pandemic, ranked 5th in the world for international tourism arrivals and 6th for international tourism earnings, recording 49,596 million USD in international tourism receipts. Consequently, similarly to most countries in the advanced economies, Italy has posed minimal focus on domestic tourism (DT).

Nevertheless, worldwide domestic tourism in 2017 represented 73% of the total global tourism spending, with China, USA, and Germany being the first three countries in the world for travel and tourism domestic spending, and Italy ranking 7th (WTTC, 2018). In September 2020 the UNWTO underlined the role of domestic tourism to drive recovery, stating that in 2018 around 9 billion domestic tourism trips were made worldwide, exceeding 6 times the number of international trips.

The economic super-shock represented by COVID-19 pandemic unprecedentedly impacted travel, hospitality, and tourism worldwide, with international movements for leisure purposes coming to a stop for almost 18 months with the exceptions of some corridors (or bubbles) that were bilaterally created between countries (as, for example, Australia and New Zealand) or some temporary lifting of travel restrictions during the summer months.

1.1 Research objectives

Throughout most of 2020 and 2021, when total lock-downs were not in place, domestic travelling has been the only form of tourism to be allowed. Due to its importance in a time of instability, the purpose of this paper is to forecast domestic tourism indicators (DTIs), such as domestic expenditures (DTE), arrivals (DA) and overnights stays (DOS), across different Italian regions and sub-regions to determine its viability in a time of health crisis, and in specific when governmental incentives are introduced. This objective is pursued by predicting the mentioned DTIs with MIDAS (mixed data-sampling) models which allow to combine effectively data sampled at mixed frequencies, such as financial and (macro)economic variables (see, among other, Ghysels et al., 2007). The R code used for this study is disseminated with the paper, to allow scholars and practitioners to forecast tourism flows of other destinations with a methodological approach based on MIDAS regressions.

In this paper, Italy is taken as a case, with the aim of analysing the first European country that has been severely affected by COVID-19 and has implemented some of the most stringent lockdown measures in the world. In Italy the lockdown started on March 8th, 2020 and it was gradually released in June 15th, 2020, when intra-European borders reopened, and holidaying resumed. With the arrival of the autumn months, Italy entered again into lock-downs. Regional COVID-19 data were used to evaluate the severity of the situation in each region and establish restrictions and lock-downs, till May 24th 2021 when Italy opened again to domestic unrestricted travelling. In light of this context, travelling and tourism can act as important forms of recovery for both the demand (the recovery from stress and anxiety due to the perception of risk and the home confinement) and the supply (economic restart).

More specifically, understanding the main trends of Italian DT offers the possibility to Italian operators and tourist destinations of organizing the reopening, gaining a clearer overview of the COVID-19 impact as well as identifying a road map for the recovery of a sector dramatically affected by the pandemic. The role of DT in this framework has become even more crucial.

In specific, this study first analyzes all the Italian regions and then it moves to specific municipalities (in South Tyrol) to focus on three distinct holiday destinations with distinguishing characteristics.

The results of these analysis also allow an evaluation of the role of recent incentives proposed by the Italian government to support the tourism sector, such as holiday vouchers.

2 Domestic Tourism

International tourism, due to its vital role in the balance of payments of many countries, has shaded the importance of domestic tourism flows (Alvarez-Diaz et al., 2020; Gálvez et al., 2014; Huybers, 2003) from both scholars' and practitioners' perspectives. The exponential growth of international tourism in the last three decades has been favoured by the increased accessibility of tourism destinations due to the expansion of the aviation industry and the recent growth of low-cost airlines (Kim et al., 2019), along with the ease of booking systems directly available to consumers and the higher available incomes. Indeed, with 61% of the population in OECD countries belonging to the middle class (OECD, 2019), and the forecast of a growing percentage of the Chinese and Indian population moving into this class in the next ten years (WTTC, 2018), tourism has become more affordable to many. According to the WTTC (2018), the demand for DT takes off at an annual income level of about US$ 35,000, while international tourism at a level of about US$50,000. Domestic tourism is, therefore, sustained by a growing percentage of middle class individuals in emerging economies and a reduction of the purchasing power of the higher end of the middle class in developed economies, where the costs of some good and services have risen at a faster pace than earnings.

Data before COVID-19 pandemic show the importance of DT also at a time of a global expansion of international tourism. Before the pandemic, domestic tourism represented 73% of the global tourism spending, with variations among countries and high percentages in both emerging and developed economies (WTTC, 2018). In 2017, the top spenders in DT was China, followed by the USA, Germany, and India, Italy ranked 7th (WTTC, 2018). The positive effects of domestic tourism are many and countries like China have supported this type of tourism through the development of an internal rail network, suitable infrastructure for low-cost airlines reaching second- and third-tiered airports. Domestic tourism offers a significant opportunity for economic growth and development and its positive effects include: maintaining and improving tourism infrastructure, especially in regional destinations; retaining economic resources within the country; transferring consumption from the richer to the poorer regions and redistributing spending power; reducing unemployment (Kim et al., 2019; Massidda and Etzo, 2012; Athanasopoulos and Hyndman, 2008).

Domestic tourism has also been identified as a substitution of international tourism (Kim et al., 2019), particularly at times of geopolitical instability and events (Huybers, 2003), as for example during SARS and soon after 09/11. Cafiso et al. (2018) have also shown that proximity tourism (domestic tourism often involves travelling to closer destinations) is preferred by tourists in a time of economic crisis. Kim et al. (2019) stressed that domestic travel can be a substitute of international travel if governmental policies make the former form of tourism viable. When analysing the case of South Korea, they underline that countries with a significant tourism deficit should focus on transforming international outbound tourism into domestic one. Indeed, not only DT can help the balance of payment of countries by limiting the amount of imports represented by outbound tourism, but it also offers business opportunities to holiday destinations.

In 2020, and for most of 2021, the tourism industry has faced a global crisis, with COVID-19 representing an economic super-shock (Dolnicar and Zare, 2020) which forced tourism destinations and operators to undertake strategic analyses to overcome the financial losses tackled for most of 2020 and for more than half of 2021, with still uncertain times ahead.

With the international tourism registering a decline of over 70% in 2020, sitting back to the levels of 30 years ago (UNWTO, 2020), domestic and, to a small extent, intra-European tourism have become the only sources of revenue for most tourism operators. Indeed, international tourist arrivals dropped by 74% in 2020 with one billion fewer tourism arrivals worldwide (UNWTO, 2021a). A slight improvement occurred during the summer seasons. Referring to the 2020 summer months, the WTO talked about "positive signs of gradual but still cautious change in trend" for the Northern Hemisphere (UNWTO, 2020). The winter months of 2020-21 have seen the reintroduction of travel restrictions in most European countries, with again some cautious openings in summer 2021. According to the (UNWTO, 2021b), the month of June and July 2021 "saw a moderate rebound in international arrivals compared to 2020. Nevertheless, 2021 continues to be a challenging year for global tourism, with international arrivals down 80% in June and July compared to 2019". Summer 2021 has seen consumers' confidence lifting due to the relaxation of travel restrictions for vaccinated and tested travellers and the roll-out of COVID-19 vaccines (UNWTO, 2021b). Still, the outlook for winter 2021 remains uncertain, with prospects of a decrease (UNWTO, 2021b). Long term, most of the experts see a recovery to 2019 levels after 2023 in Europe and in 2024 in Asia and the Pacific (UNWTO, 2021b).

With such scenarios, DT represents a viable economic activity as confirmed also by Eurostat data. Indeed, in 2020 (the peak of the pandemic), European DT started to recover earlier than international tourism, reaching in August 2020 almost similar numbers (only 5% less) to August 2019 (EUROSTAT, 2021).

2.1 Tourism in Italy

Before 2019, approximately half of the world's international arrivals and almost 40% of the international tourist revenues were generated from Europe. In this regard, Italy is located in a geographical area with a strong tourist vocation (5th place in the world for tourist arrivals and 6th for tourist revenue, with a 7% increase in 2018 with respect to 2017). At a national level, tourism activities represent more than 5% of GDP and more than 6% of employment. These data are comparable to those of Spain (second country in the world in terms of arrivals and tourist revenues) and are higher than those of France (first country in the world for arrivals) and Germany. The contribution induced by tourism in Italy on the total GDP and employment activated by the expenditure of people employed directly or indirectly in the tourism sector is relevant. Indeed, the overall impact on GDP in 2017 reached 13%, a value higher than the world and European average, while the overall impact on employment also reached levels above the world average (10%), accounting for almost 3.4 million jobs, equal to 15% of the total national employees.

Over the years, DT in Italy has represented an important economic activity, covering approximately half of the overall arrivals and overnight stays. In the period 2012 to 2017 the share of domestic tourism has slightly decreased, shifting from 53% in 2012 to 50% in 2019. Nonetheless, this decrease in share, is not given by a contraction of the demand, rather by unprecedented records in international tourism. Indeed, domestic arrivals increased from 54,994,582 in 2012 to 66,371,433 in 2019 and overnight stays increased from 200,116,495 in 2012 to 216,076,587 in 2019.

From an economic perspective, the contribution of domestic tourism spending on the national GDP surpasses the one of international spending. In the years 2017-2018, about 76% of travel and tourism spending had been generated by domestic tourism (ISTAT, 2021a).

Despite the WHO officially declared a pandemic on 11 March 2020, Italy went into a full lock down on 9 March 2020 (Bourdin et al., 2021). Italy has been the first European nation to be severely affected by COVID-19 and the first one to adopt stringent measures. As a consequence, in 2020, Italy recorded a drop of over 63 billion Euros in tourism consumption. Total tourist arrivals dropped to 55.7 million (131 million in 2019), of which around 16.5 from international tourists (65 million in 2019). Overnight stays recorded similar losses: a 54% loss from 2019 was recorded in international overnight stays, resulting in a loss of about 35 billion Euros compared to the previous year (ISTAT, 2021a,b). The share of international tourism on the total contribution of travel and tourism to the national GDP dropped to 19.4% (Statista, 2021).

This contraction in international tourists visiting Italy, was largely offset by an equally drastic reduction in Italian tourist flows abroad (outbound tourism) which, in terms of nights, fell by 54.1%, with an expenditure level of 13.7 billion euros (-65.7% compared to the previous year). The limitations of movements for tourism purposes have contributed to transforming part of the outbound tourist flows into domestic ones, for which the decrease in overnight stays was at a lower rate (-32.2%). Consequently, part of the tourism expenditure normally spent abroad in previous years remained in Italy, helping to curb the overall decrease in tourism consumption.

The substitute role of domestic tourism in Italy during the first year of pandemic (2020) is further confirmed by the analysis of the travelling behaviour of the Italian population (based on ISTAT, 2021b). In 2020, travel abroad by Italian citizens collapsed recording an 80% loss in number of trips and a 78.2% loss in overnight stays compared to 2019. DT was affected to a lesser extent, and registered a negative variation of 37% in the number of trips and -27.2% in terms of nights (about 76 million less) between 2019 and 2020. In June, thanks to a resumption of inter-regional travel, trips partially recovered but amounted to about half of those in the same month in 2019. The only positive sign of the year was recorded in July and August and concerned short holidays (+19.5% compared to 2019, +31% in terms of overnight stays), which mitigated the overall decrease in summer travels (-19.8%).

After a stable increase in domestic tourism spending in Italy from 93.9 billion Euros in 2010 to 142.8 billion Euros in 2019, 2020 recorded a drop to 71.9 billion Euros, representing a 50% loss compared to the international loss of 62% (from 45.6 billion Euros in 2019 to 17.3 billion Euros in 2020).

2.2 Holiday vouchers

In summer 2020, nations around the world implemented different initiatives to promote domestic tourism, including financial incentives addressing the demand and the supply side. Within this framework, Italian DT was stimulated by holiday vouchers, which are financial incentive aimed at supporting the demand. Starting the first of July, Italian families with an overall economic level of up to €40,000 had the possibility to access to a voucher of up to €500 to be used for the payment of a registered accommodation service within the country. The voucher amount was based on the number of family members (from €150 for one person, €200 for two persons, €500 for three or more persons). The voucher was available only in digital form and granted an immediate discount of 80% at the payment of the accommodation service; the remaining 20% could be discharged as a tax reduction when filing the tax declaration at the end of the financial year. The accommodation services that were accepting vouchers were refunded in the form of tax credits or could transfer the vouchers to third parties, such as credit institutions and financial intermediaries. Initially, the vouchers had to be used within December 2020. However, in 2021, the validity of vouchers requested in 2020 and not used before the end of the year was extended till December 2021. The present study on DTIs in Italy will therefore also discuss the implications of the holiday vouchers as an incentive to boost DT.

3 Data

The analysis of the DTIs has been carried out according to different perspectives with the aim of investigating and comparing domestic tourism dynamics (yearly arrivals, overnight stays and travel expenditures: DA, DOS and DTE, respectively) at different geographic levels and more precisely at national, regional and sub-regional destinations. Thus, first we have investigated the prospects of DT in Italy overall and across regions and, then, we have considered specific tourism destinations in the South Tyrolean Province to highlight DTIs differences across destinations (at municipality level) with distinguishing features. Accordingly, the analysis undertaken is based on the following three datasets (see Figure 1).

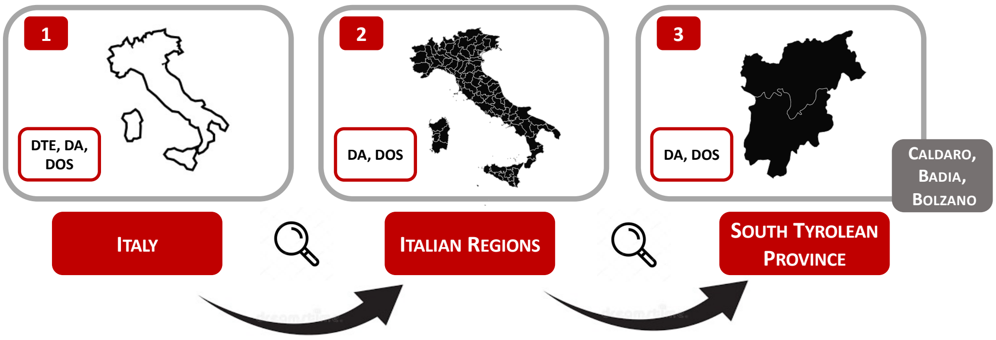


Figure 1: Dataset selection. *Source*: own elaboration

Dataset 1 is national specific and collects: yearly DTE (y_t^{DTE}) to a one-night stay or over from 2009 to 2019, expressed in thousand of Euro (data source: Eurostat); yearly, quarterly and monthly DA (y_t^{DA}) from 1996 to December 2020; yearly, quarterly and monthly DOS (y_t^{DP}) from 1996 to December 2020; monthly and weakly FTSE-MIB ($x_{1,t}$) from 1998 to March 2021; quarterly GDP ($x_{2,t}$) from 1996 to the end of 2020, expressed in current prices (million of Euro); quarterly wage and salaries (WS - $x_{3,t}$) from 1996 to the end of 2020, expressed in current prices (million of Euro). Quarterly GDP and WS are not adjusted for seasonally or calendar effects.

At the moment, DTE is available only in the time span 2009-2019. Unfortunately, measuring the COVID-19 impact on DTE requires data not available yet. Accordingly, one of the aim of this work is to fill this gap by providing estimates of these data.

For the scope of the following analysis, we have also created two further variables based on the logarithm of the DA and DOS ratios, that is:

$$\log\left(\frac{y_t^{DA}}{y_{t-1}^{DA}}\right) \quad \text{and} \quad \log\left(\frac{y_t^{DOS}}{y_{t-1}^{DOS}}\right). \tag{1}$$

These variables are shown in Figure 18 together with the original DTIs in Figure 17. Figure 19 proposes the plot of the yearly DT length of stay that, in 2020, reach the maximum level of 4.5 days. Figure 16, which shows the quarterly WS, the quarterly GDP and the monthly FTSE-MIB series together with their seasonally adjusted series, highlights the impact of COVID-19 in the final part of 2020.

Dataset 2 is regional specific and collects, at least from 1998, time series for the 20 regions of Italy. It includes: yearly y_t^{DA} and y_t^{DP} (at regional level) from 1996 to December 2020; monthly FTSE-MIB from 1998 to March 2021; quarterly GDP and WS (both at national level) since 1996 up to the end of 2020. Logarithms of the DA and DOS ratios are also computed.

Dataset 3 is municipal specific (South Tyrolean Province) and collects yearly DA and DOS (1990-2020) for Badia, Bolzano, and Caldaro. Also in this case, we include: the logarithm of the DA and DOS ratios; monthly FTSE-MIB; quarterly GDP and WS (at national level).

It is worth noting that the South Tyrol is an interesting region collecting tourism destinations with different characteristics, such as mountain, lake, and urban destinations. Therefore, we have considered three different destinations in South Tyrol: Badia, Bolzano and Caldaro. All of them are governed by the same provincial tourism body, which coordinates development strategies and marketing efforts. However, they have substantially different characteristics: a mountain destination with very strong winter (skiing) and summer (hiking) touristic seasons (Badia); a lake destination with one season lasting approximately seven months stretching from spring to early autumn (Caldaro); a urban destination attracting business and cultural tourists all year around (Bolzano).

4 Estimating and forecasting domestic tourism indicators via MIDAS models

This section is devoted to explain how Italian DTIs (yearly DA, DOS and DTE), can be forecast using macroeconomic and financial determinants, typically observed at higher frequencies than DTIs. To avoid the ex-post aggregation of high-frequency data and potentially discarding useful information, we have employed MIDAS models (see, among others, Ghysels et al., 2005, 2006, 2007; Andreou et al., 2010). The latter allow policy makers to predict low-frequency variables as functions of higher-frequency ones and assess in real time, or nowcast, low-frequency indicators by using some extra observations on higher frequency variables which became available after the most recent value of the low-frequency dependent variable (see Figure 2). Compared to other approaches for mixed frequency data, such as State Space and mixed frequency VAR models, MIDAS prove to be more parsimonious and less sensitive to specification errors (Ghysels et al., 2006).

The interest in using data sampled at different frequencies to explain and forecast variables that are emblematic in the tourism sector has increased in recent years. One of the first contributions has employed Google trends to forecast tourism indicators with MIDAS models (Havranek and Zeynalov, 2021). However, when studying DTIs, Google trends appear not to be so much informative as they do not allow to distinguish domestic from foreign tourists. Neglecting this issue may introduce a bias in our forecast.

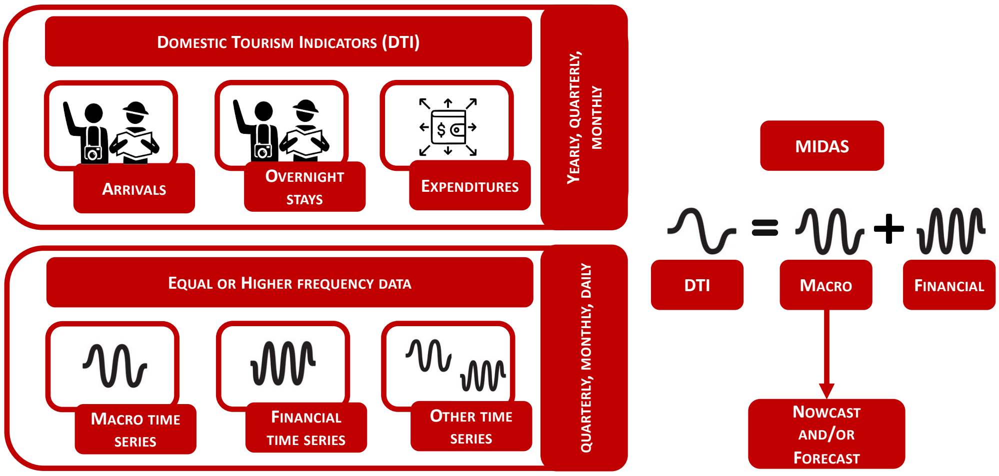


Figure 2: Conceptual framework. *Source*: own elaboration

The approach here developed consists of the following steps (Figure 3):i) select the determinants of DTIs, which are sampled at different frequencies; ii) estimate different MIDAS models for DTIs and choose the best ones on the basis of their out-of-sample performance (Figure 4); iii) forecast DTIs determinants (via, for instance, AR, ARIMA, MA or MIDAS models); iv) nowcast and/or forecast DTIs. The following section explains in detail the econometric framework of MIDAS models.

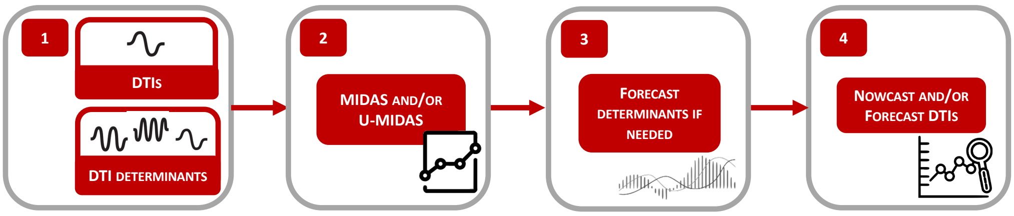


Figure 3: The methodological proposal. *Source*: own elaboration

4.1 MIDAS and U-MIDAS models

MIDAS regressions directly relate variables sampled at different frequencies (see, e.g. Ghysels et al., 2005, 2007, 2006; Andreou et al., 2010). These models are strictly related to distributed lag models as the dependent variable is regressed on distributed lags of the higher frequency ones. The frequency alignment is achieved by using a flexible parsimonious parametric functional constraint expressing the impact of the high-frequency variable on the lower-frequency one. A general univariate MIDAS regression model explaining a dependent low-frequency variable y_t as function of r regressors, $x_\tau^{(i)}, i = 1, 2, \dots, r$, observed at higher-frequencies m_i, that is observed m_i times at time t, meaning $\tau = m_i t$

for all t, can be specified as follows

$$y_t = \alpha + \sum_{i=0}^{r} \sum_{k=0}^{K_i} \beta_k^{(i)} x_{tm_i-k}^{(i)} + \epsilon_t, \quad t = 1, 2, \ldots, T \tag{2}$$

where $\sum_{k=0}^{K_i} \beta_k^{(i)} = 1$ for any i. To reduce the number $\sum_{i=0}^{r} K_i$ of the unknown parameters $\beta_0^0, \ldots, \beta_{K_0}^0, \ldots, \beta_0^r, \ldots, \beta_{K_r}^r$ a weighting scheme of the type

$$\beta_k^{(i)} = \frac{\psi(k, \boldsymbol{\theta}_i)}{\sum_{k=0}^{K_i} \psi(k, \boldsymbol{\theta}_i)} L^k \tag{3}$$

is assumed. In Eq. (3) the function $\psi(k, \boldsymbol{\theta}_i)$, depending on a low dimensional parameter vector $\boldsymbol{\theta}_i$, plays a fundamental role in achieving a flexible parsimonious parametrization. Once the latter is properly chosen, the number of lags K_i is purely based on information criteria. The most widely used function for $\psi(k, \boldsymbol{\theta}_i)$ are the exponential Almon polynomial and the beta polynomial. The former is specified as follows

$$\psi(k, \boldsymbol{\theta}_i) = e^{\sum_{l=1}^{H} \theta_l k^l}, \tag{4}$$

while the latter as

$$\psi(k, \boldsymbol{\theta}_i) = \frac{\left(\frac{k}{K_i}\right)^{\theta_1 - 1} \left(1 - \frac{k}{K_i}\right)^{\theta_2 - 1} \Gamma(\theta_1 + \theta_2)}{\Gamma(k)\Gamma(K_i)} \text{ with } \Gamma(\cdot) = \int_0^\infty e^x x^{\cdot - 1} dx. \tag{5}$$

It is worth noting that MIDAS regressions are linear in the variables, but not in the parameters and as such they are estimated by non-linear least squares (NLS). An autoregressive augmentation can be introduced in Eq. (2):

$$y_t = \alpha + \sum_{j=1}^{n} \beta_j L^j y_t + \sum_{i=0}^{r} \sum_{k=0}^{K_i} \beta_k^{(i)} x_{tm_i-k}^{(i)} + \varepsilon_t. \tag{6}$$

When parameters in Eq. (2) are not constrained to be generated by a given functional distributed lag polynomial, the model is labelled unrestricted MIDAS (U-MIDAS) and can be estimated directly by OLS (Foroni et al., 2015). U-MIDAS are particularly useful in macroeconomic applications where the sampling frequency gap between variables included in the model is not so high: for example when quarterly (monthly) data are used to predict a yearly (quarterly) macro variable. In these cases, the number of quarterly (monthly) lags necessary to estimate the lag polynomials may not be too large, and the curse of dimensionality turns out to be not relevant (Koenig et al., 2003; Clements and Galvão, 2009; Foroni et al., 2015).

Both MIDAS and U-MIDAS regressions result to be suitable econometric tools to nowcast and/or forecast DTIs. Nowcasting refers to the prediction of the present, or the very recent past, based on the information provided by the available data that are sampled at higher frequencies (Rufino, 2019). In this regard, forecasts based on MIDAS models can be obtained through a multi-step forecasting instead of an iterative procedure based on a set of ahead one step forecasts. This is in contrast with many other forecasting approaches, such as autoregressive and state space models, where multi-step forecasts are necessarily done iteratively. Upon noting that model in Eq. (2) can be written in compact form as follows

$$y_t = \alpha + \boldsymbol{B}(L)' \boldsymbol{x}_{t,0} + \epsilon_t \tag{7}$$

where

$$\boldsymbol{x}'_{t,0} = \left[x^{(0)}_{tm_0}, .., x^{(K)}_{tm_K}\right]' \tag{8}$$

$$\boldsymbol{B}(L)' = \sum_{j=0}^{K} \boldsymbol{\beta}'_j L^j, \qquad \boldsymbol{\beta}'_j = [\beta^{(0)}_j, ...\beta^{(i)}_j, ..., \beta^{(K)}_j] \tag{9}$$

with $K = max(K_i)$, an h-step direct ahead forecast can be simply obtained from the model

$$y_{t+h} = \alpha + \sum_{i=0}^{r} \sum_{k=0}^{K_i+m_i h} \beta^{(i)}_k x^{(i)}_{tm_i - k} + \epsilon_{t+h} = \alpha + \boldsymbol{B}_h(L)'\boldsymbol{x}_{t,0} + \epsilon_{t+h} \tag{10}$$

where $\boldsymbol{B}_h(L)'$ are horizon h-specific parameters. In nowcast cases, no new data are required for forecasting while the nowcasts $\hat{y}_{t+h}|T = E(y_{t+h}|I_T)$, with I_T denoting the information available up to the most recent lower frequency T, can be obtained by using the explanatory variable data up to time T.

In nowcasting and forecasting domestic yearly DTIs, quarterly wage and salaries (WS), quarterly GDP, monthly or weekly values of the FTSE-MIB and monthly domestic overnight stays (DOS) have been used as covariates of the following MIDAS models. It is worth noting that even if the considered DTIs are not a very long, all the following specified models are adequate for performing the desired nowcast/forecast:

M1 A parsimonious U-MIDAS model with a trend variable and a unique covariate: either WS (model M1a) or GDP (model M1b).

$$y_t = \alpha_0 + \alpha_1 * trend + \sum_{k=0}^{3} \gamma_k x_{tm-k} + \epsilon_t. \tag{11}$$

The model proves suitable insofar as, being DTIs observed yearly, the difference between its sample frequencies and that of the independent variable is not high and the curse of dimensionality is avoided.

M2 An U-MIDAS model specified as in Eq. (11) with the inclusion of an autoregressive component $\beta_1 y_{t-1}$. The model is denoted M2a or M2b, depending on it uses either WS or GDP.

M3 A MIDAS model with either exponential Almon or beta polynomial as weighting schemes for the coefficients of the independent variable over two years, that is

$$y_t = \alpha_0 + \alpha_1 * trend + \sum_{k=0}^{7} \beta_k x_{tm-k} + \varepsilon_t. \tag{12}$$

where $\beta_k = \frac{\psi(\boldsymbol{\theta},k}{\sum_{k=0}^{7}(\psi(\boldsymbol{\theta},k)}$, $\sum_{k=0}^{7} \beta_k = 1$, with $\psi(.)$ defined as in Eq. (4) or as in Eq. (5). The model is denoted M3.1a or M3.1b if the covariate is WS and the weighting scheme is either an Almon or a beta polynomial. It is denoted M3.2a or M3.2b if the covariate is the GDP and the weighting scheme is either an Almon or a beta polynomial.

M4 A MIDAS model including two high-frequency determinants of DTIs, that is

$$y_t = \alpha_0 + \alpha_1 * trend + \sum_{i=0}^{1} \sum_{k=0}^{K_i} \beta_k^{(i)} x_{tm_i - k}^{(i)} + \varepsilon_t. \tag{13}$$

The model is denoted M4.1a, M4.1b, M4.2a , M4.2b, M4.3a and M4.3b depending on which couple of covariates – WS and GDP, WS and DOS (only for the expenditures case), WS and FTSE-MIB – and of weighting scheme – Almon (a) or a beta (b) polynomial – are employed.

M5 A MIDAS model specified as in M4, enriched by the introduction of an autoregressive component, $\beta_1 y_{t-1}$. The model is denoted M5.1a or M5.1b if the covariates are WS and GDP and the weighting scheme is either an Almon or a beta polynomial, respectively. The model is denoted M5.2a or M5.2b if the covariates are FTSE-MIB and GDP and the weighting scheme is either an Almon or a beta polynomial, respectively.

5 Results

The results discussed in this section have been obtained from MIDAS/U-MIDAS models. The undertaken analysis consists of five steps which are shown in Figure 4.

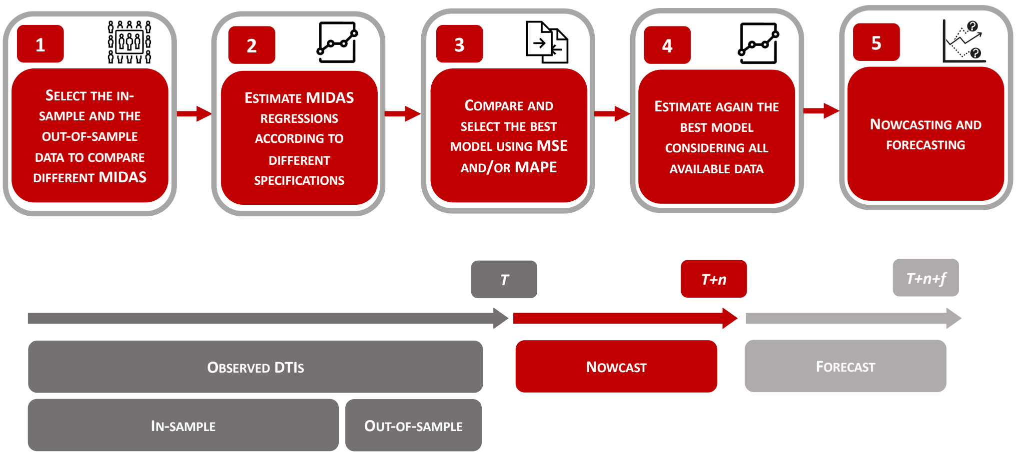


Figure 4: Applied methodology for nowcasting and forecasting. *Source*: own elaboration

Dataset 1

First, the yearly DTE in Italy, which is available from 2009 to 2019, has been extended up to 2021 by using U-MIDAS/MIDAS regressions. The best model to nowcast and forecast this series has been selected among those of Section 4 which employ a single covariate.

The models have been estimated over the period 2009-2018 and employed to forecast DTE in 2019 as shown in Figure 4. Looking at Table 2, which reports MSEs (mean squared errors of the forecasts) and MAPEs (mean absolute percentage error) for the forecasts, we can conclude that M3.2a is the best model for DTE. Accordingly, the values of the latter have been nowcasted up to 2020 and the value for 2021 has been forecast using model M3.2a, which employs the national GDP as independent variable. Figure 5 shows the sample values (in black) and the forecast (in red) values of DTE. Looking at this figure, we see that DTE exhibits a severe decrease in 2020, when compared to the value registered in 2019 (-62% given an expenditure of 59,243.39 million Euros), which reaches the minimum level in this year. Moreover, the forecast values for this variable show a slight increase in 2021 when it should reach 81,190.38 million Euros. However, this slight increase reflects a 47% reduction compared to DTE in 2019.

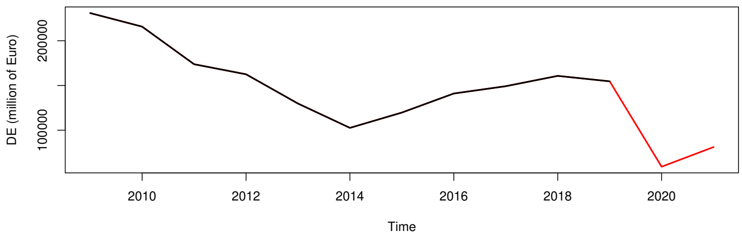


Figure 5: Nowcast/forecast of the Italian time series of DTE (data source EUROSTAT) for 2020 and 2021 using model M3.1b. *Source*: own elaboration

Then, the yearly DA and DOS have been investigated. Given that the FTSE-MIB series, which is a potential explanatory variable for the aforementioned series, is available only from 1998, in the following analysis the sample range has been reduced to meet this constraint. Thus, the subsequent results refer to the period 1998 - 2020. As data on DA and DOS are available up to the end of 2020, we have forecast these series up to 2021. Models introduced in Section 4 have been still used again for forecasting purposes.

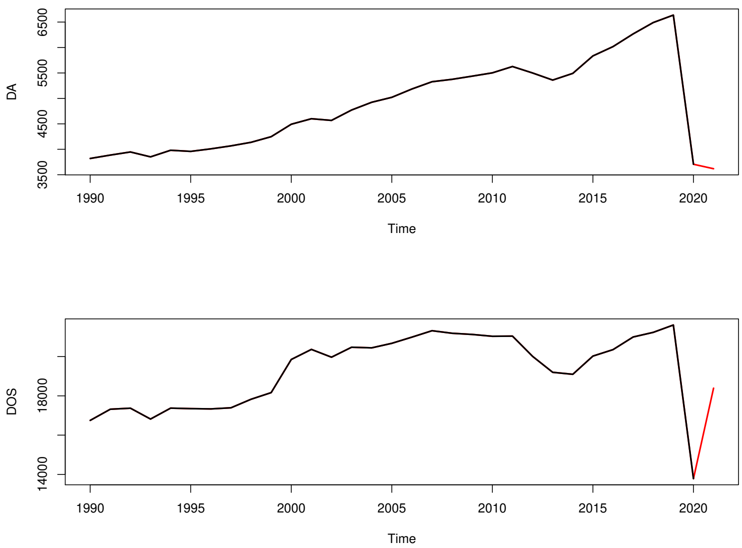


Figure 6: Italian series of DA (top panel) and DOS (bottom panel) together with their forecast for 2021 using model M2b and M4.2a, respectively. *Source*: own elaboration

To avoid negative forecasts of DTIs, DA and DOS for Italy have been forecast by considering the log of the yearly ratios (i.e., $\log(y_t/y_{t-1})$) in Eq. (1). The MIDAS/U-

MIDAS models introduced in the previous section have been estimated and the out-of-sample MSEs and MAPEs are reported in Table 3. The best out-of-sample predictions have been obtained by using models M1b and M4.1a for DA and DOS, respectively. Forecast values engendered by these models are shown in Figure 7. As for DA, the forecast value -0.921 in 2021 means that, taking into account 37,058,635 observed arrivals in 2020, a further important reduction is expected in 2021 when this series reaches its minimum level of 14,759,715. Similarly, the forecast of -0.0490 for DOS in 2021 means that, taking into account 137,841,845 observed overnight stays registered in 2020, a further reduction is expected in 2021 when this series should reach its minimum level of 131,244,542. The forecasts, shown in Figure 18, have been computed as follows

$$forecast_{2021} = \exp(lnfy_{2021}) * observed_{2020}, \tag{14}$$

where $lnfy_{2021}$ is the forecast of $\log(y_{2021}/y_{2020})$.

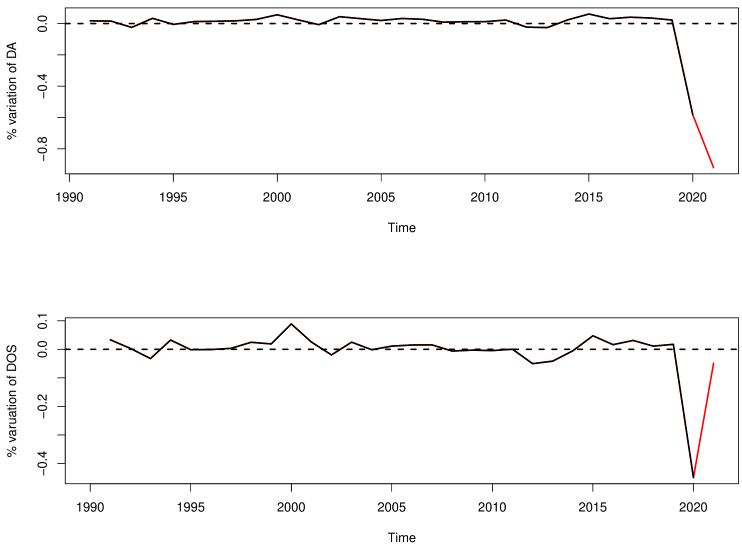


Figure 7: Italian log ratio (variation) of DA (top panel) and DOS (bottom panel) together with their 2021 forecasts using model M1b and M3.2a, respectively. *Source*: own elaboration

Dataset 2

We have then focused on regional DTIs. First, we have selected the optimal models to explain Italian regional DA and DOS among the ones previously considered. The results for each region are reported in Figure 20. These models have then been employed to forecast the log ratios and the related absolute values of both DA (Figure 22) and DOS (Figure 24) in 2021.

Figure 21 shows the DA trend from 2016 to 2018 across different Italian regions. We can see that DA is stable in most regions, with variations around zero and ranging mainly from +0.2 to –0.4. The regions lying in the central part of northern Italy were the ones which attracted most domestic tourists. Particularly attractive was Emilia Romagna with its short, but well-known Adriatic coast, followed by Lombardy (with the city of Milan symbolising the Italian capital of business tourism), and Veneto, renowned for culture (Venice and Padova) and its Adriatic coast.

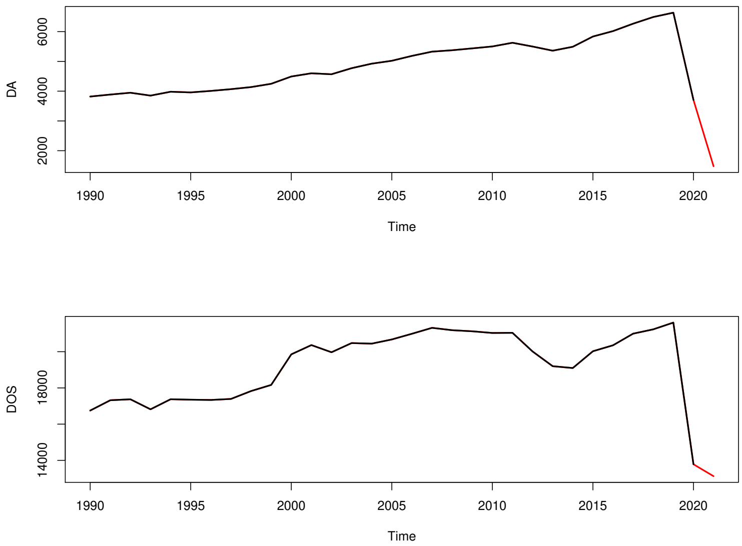


Figure 8: Italian DA (top panel) and DOS (bottom panel) together with their 2021 forecasts obtained from model M1b and M3.2a, respectively. *Source*: own elaboration

Figure 22 shows DA in 2019 and 2020 (also based on the variation from 2018 and 2019, respectively) along with its forecast for 2021. Also in 2019 DA was stable in the northern regions, despite a slight decrease in variation from 2018 in most Italian regions, including Emilia Romagna and Veneto, which performed particularly well in 2018. Emilia Romagna, Lombardy, and Veneto remained the most attractive in terms of absolute numbers of DA, despite a slight contraction in the variation of arrivals.

As expected, and also as indicated in the official data of the Italian Institute of Statistics (ISTAT, 2021a, and Section 2.2), DA in 2020 dropped in all regions. Lombardy and Piedmont in the north (with the two industrial cities Milan and Turin) and Lazio (with the capital regions of Italy Rome) are the regions which suffered the most in terms of DA variations. Also Emilia Romagna and Veneto experienced a contraction, however less severe than the three regions mentioned above. Interesting is the situation of the central regions of Marche and Abruzzo, and of the two island Sardinia and Sicily. Up to 2019, Sardinia and Sicily were stable in terms of both variations and numbers of DA. Over the years both regions have been able to internationalize and attract high spending international visitors, reducing, consequently, their shares of domestic visitors. Marche, which experienced a growth in DT in 2019, performed similarly to Sardinia and Sicily. The islands and Marche remained low in absolute numbers, however managed to maintain a percentage of domestic arrivals higher than other regions in Italy, showing their capabilities of substituting part of their international visitors with domestic ones.

Also Trentino South Tyrol in northern Italy performed better (in terms of variance) than other regions. The region relies on its mountains (among which the UNESCO listed Dolomites), its untouched nature, and high-quality accommodation. Furthermore, being located at the border with Austria and being part of Italy only since 1919, it exerts to many Italian tourists a touch of exotic. The forecasts for 2021 show minor signs of recovery for the most of Italy. Positive variations in terms of DA are seen in Campania, Tuscany, Piedmont, and Lombardy. The latter industrial region is performing well also in terms of DOS. The forecasts signal a recovery of domestic business tourism for 2021. On the other side, DA and DOS remain stable for the two islands if confronted with 2020.

It can be concluded that regions which feature typical summer destinations in 2019 attracted around half of its visitors form international markets. Among these, there are Trentino South Tyrol, Sardinia, and Sicily, which were able to cover some of the losses caused by the pandemic through domestic tourism in 2020. Other regions, heavily relying on DT before 2019, as for example Emilia Romagna, underwent harder pressures. Furthermore, the forecast for 2021 shows a slow improvement for business tourism in industrial regions and a decrease for the two islands. The models indicate the capability of DT to partially substitute international tourism in a time of health crisis, even if international outbound tourism by Italian citizens is resumed when positive economic and pandemic signals appear. Therefore, Italian tourists who were choosing international destinations for their holidays before the pandemic, turned their choices into domestic destinations in 2020, although they continued to show interest towards international destinations when a slight easing of the pandemic occurred. Figure 9 shows the forecast for 2021 in more details, by comparing the variations in DA and DOS directly.

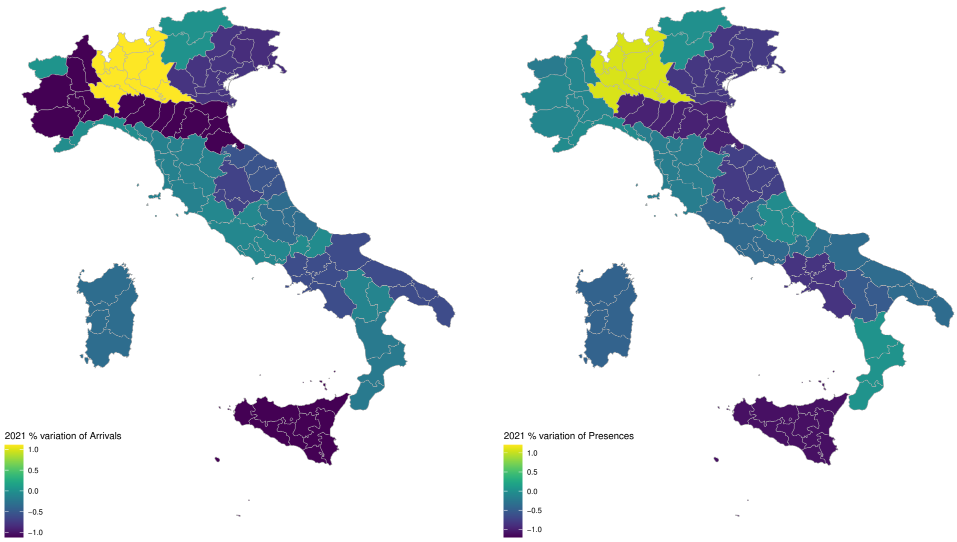


Figure 9: Regional log ratio (variation) of DA (left panel) and DOS (right panel) - forecast 2021

Dataset 3

Finally, the same analysis has been conducted for three Italian municipalities: Badia, Bolzano and Caldaro. In this case, the evaluation of the out-of-sample forecast performances of the MIDAS models, previously introduced, has lead to the results for DA and DOS highlighted in Figures 10-15.

DA and DOS forecast have been then obtained for these three tourist destinations. As for Badia, the mountain destination with strong winter skiing and the summer hiking seasons, DA is predicted to decrease in 2021 accompanied by a slight increase in DOS. Indeed, the length of holidays is predicted to slightly increase for domestic tourists.

Forecasts show that Bolzano, the urban destination attracting business and cultural tourists will outperform the other two destinations in terms of overnight stays. In this regard, forecasts at Italian regional level have also shown improvements for business tourism in 2021.

Finally, Caldaro, the lake destination with a longer summer season, is forecast to perform better than the other two destinations in terms of DA, despite a predicted decline in DOS. According to this result, shorter holiday breaks are forecast to increase for this type of destination. Having a longer season spanning from late spring to early autumn, Caldaro can attract higher numbers of visitors during long week-ends and short Easter and Autumn breaks.

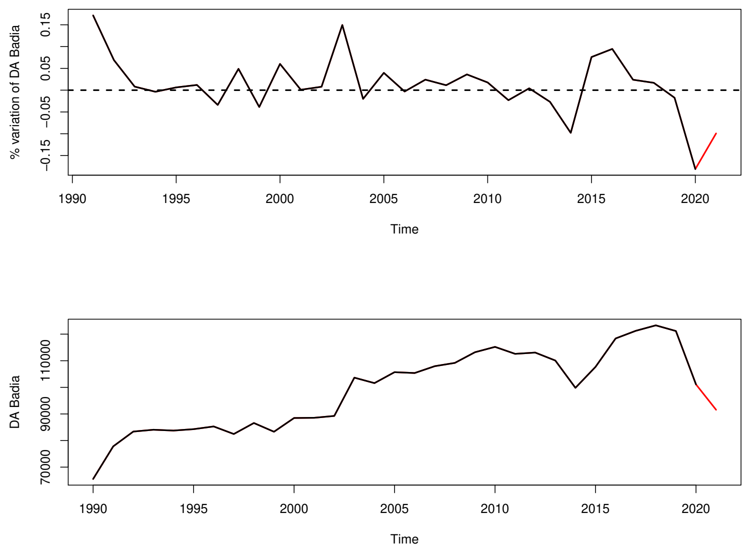


Figure 10: Italian log ratio of DA (top panel) and DA (bottom panel) for Badia together with their 2021 forecasts (optimal model M4.1a). *Source*: own elaboration

It is worth noting that the analyses here proposed are easily reproducible and not computation intensive. They have been conducted using the free software R and taking advantage of functions built in the MIDASR package of Ghysels et al. (2016). Part of the code used to obtained the results (we report only the code for domestic expenditures) here presented can be founded in Appendix B.

6 Discussion and conclusions

The objective of this work was to forecast DTIs across different Italian regions and sub-regions, and to determine its viability in a time of crises, in specific when governmental incentives are introduced. A further objective was to propose MIDAS regressions as suitable forecasting tools for the specificity of our empirical issue. Accordingly, in this section we first discuss the forecasts in relation also to the governmental incentives introduced; and secondly, we argue the MIDAS effectiveness in combining information sampled at different frequencies.

Forecasts of DTIs across Italian regions highlight the role played by DT to be a valuable resource during a period of crisis. However, the substitute role of DT is volatile and with positive economic and pandemic signs, international tourism tend to take off again. For this reason, incentive for the DT promotion should be in place also beyond the crises.

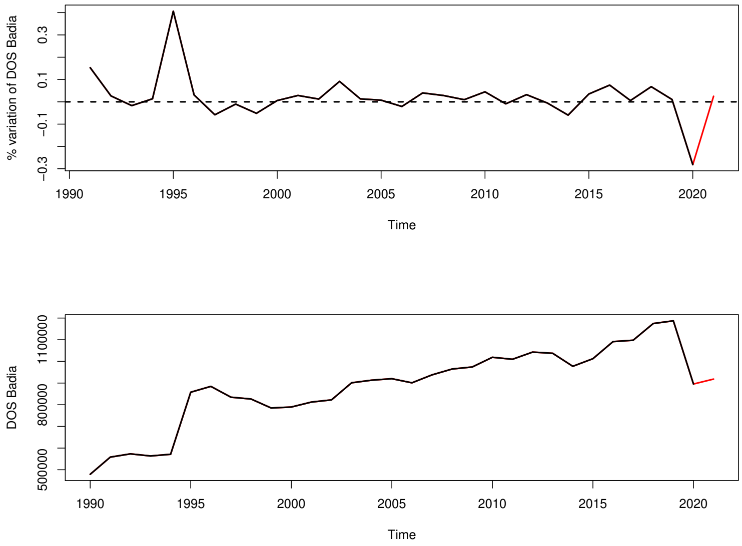


Figure 11: Italian log ratio of DOS (top panel) and DOS (bottom panel) for Badia together with their 2021 forecasts (optimal model M2b). *Source*: own elaboration

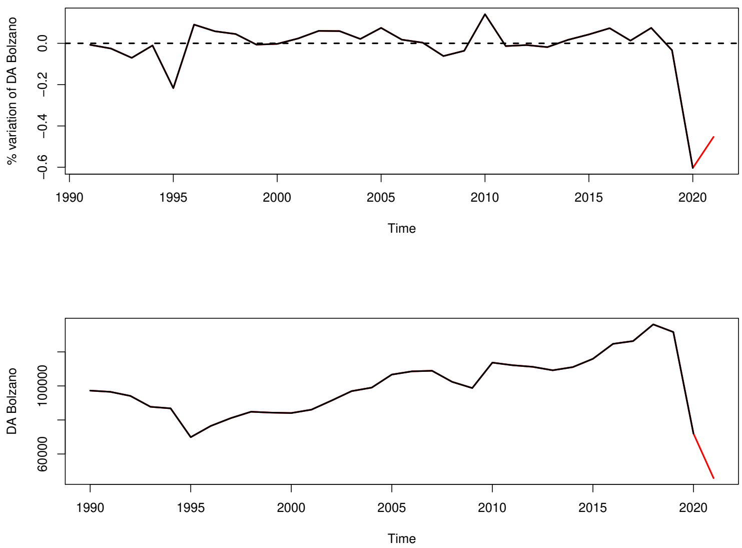


Figure 12: Italian log ratio of DA (top panel) and DA (bottom panel) for Bolzano together with their 2021 forecasts (optimal model M2b). *Source*: own elaboration

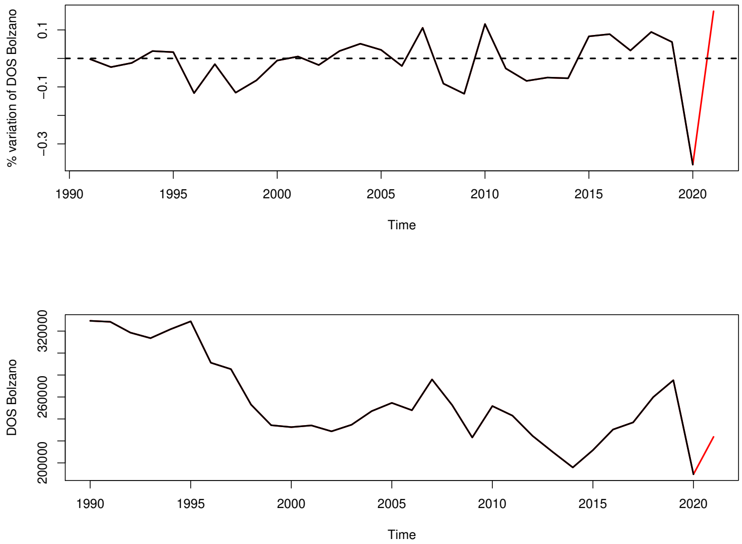


Figure 13: Italian log ratio of DOS (top panel) and DOS (bottom panel) for Bolzano together with their 2021 forecasts (optimal model M2a). *Source*: own elaboration

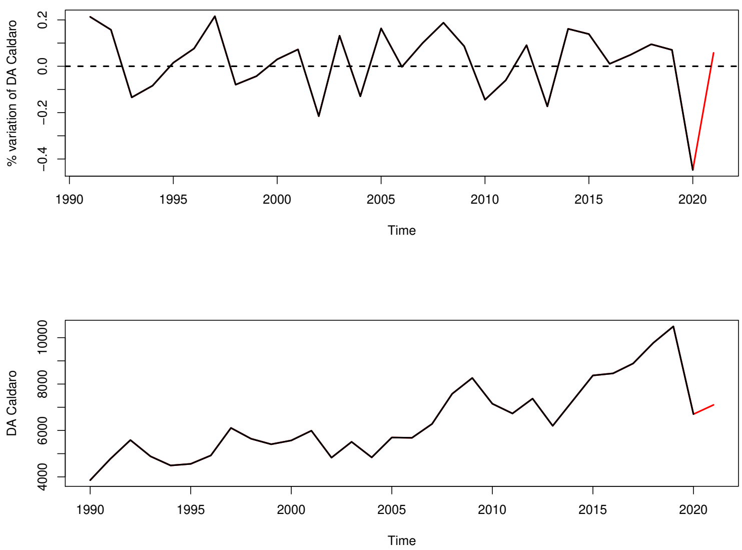


Figure 14: Italian log ratio of DA (top panel) and DA (bottom panel) for Caldaro together with their 2021 forecasts (optimal model M3.2a). *Source*: own elaboration

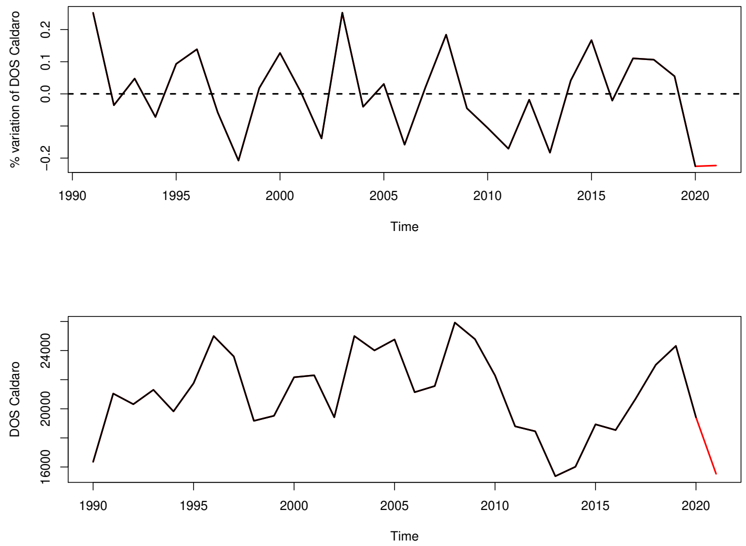


Figure 15: Italian log ratio of DOS (top panel) and DOS (bottom panel) for Caldaro together with their 2021 forecasts (optimal model M4.1a). *Source*: own elaboration

Incentives can take different forms and developed and emerging economies many have used tax incentives aimed, at the supply side, to stimulate investment growth in the tourism sector (Munyanyi and Chiromba, 2015). COVID-19 has had such a disruptive effect on the tourism sector that some countries (including Italy) have decided to target incentives directly to the demand. Other nations, which used similar incentives in summer 2020 are Greece, Island, Slovenia, Denmark, Malaysia, and Thailand. In specific, Italy has offered a voucher of up to €500 to each family belonging to a certain economic level to be used for accommodation services. Official ministry data (https://io.italia.it/dashboard/) show that a total of 1,885,802 holiday vouchers were generated, for a value of 829,431,050 Euros. The ministerial decree, which introduced the holiday vouchers foresaw a budget of 1,667.2 million Euros for 2020 and 733.8 million Euros for 2021. It becomes evident that the incentive has not been fully exploited by Italian families. In our analysis, we forecast a total expenditure of 59,243.39 million Euros for 2020. The sum foreseen by the government for the incentive would therefore have covered roughly almost 3% of the total expenditure. Nonetheless, only 829.43 million Euros were actually requested by the Italian population; this sum covers only 1.4% of the total expenditure. The number of vouchers actually used within 31st December 2020 further underlines the limited effectiveness of the incentive.

By calculating an average value of €440 per voucher, the estimated value of the voucher spent for 2020 is about 339,240,000 Euros. Therefore, only 0.6% of the 2020 expenditure was generated by the vouchers.

Thus, at national level, the incentive had a managerial impact on the total expenditure. However, further research could analyse data of the used vouchers at regional level, not available yet. Our results show the volatility of DT; therefore we suggest to Italian policy makers to further incentivize DT for the next few years of recovery (international tourism is forecast to return to pre-pandemic levels between 2023 and 2024). This could be done through other forms of governmental spending, such as: tax incentives directed to the supply for the development of products responding to domestic needs; budget increments for marketing and promotion; and/or the improvement of existing amenities. This latter factor, for example, has been identified by Bernini et al. (2020) as a factor capable of reinforcing loyalty and increasing attractiveness.

Therefore, some discussion points arise on the limited success of the incentive: how much did the families spend in total (above the voucher) for their holidays? Did the vouchers triggered higher expenses that otherwise would have not occurred? This issue needs the collection of ad-hoc data, and we urge the Italian Ministry of Finance to do so.

Moreover, in Section 2 we have discussed that DT takes off at an annual income level of about US$ 35,000. The Italian incentive targeted families with an economic level (ISEE is the Italian economic level considered, which also includes assets as for example house properties) of up to 40,000 Euros. Is the targeted income level too low? Was the voucher too low for the families within the economic level targeted? Would there have been a higher use of the vouchers if higher values were given to families? Should future incentives (possibly with lower vouchers) include also families with an economic level higher than 40,000 Euros? Table 1 shows the values for this discussion.

Table 1: The numbers of the Italian holiday voucher

Forecast expenditure 2020 (own elaboration)	59,243.39 million Euros
Number of vouchers granted by 31st December 2020	1.9 million vouchers granted
Value of vouchers granted by 31st December 2020	829.43 million Euros
Ministerial budget for the incentive 2020	1,677.2 million Euros

To conclude the first part of this discussion, also single regional and sub-regional destinations should work on the promotion of DT through the creation of loyalty programs. Those visitors, who substitute their international vacation with a domestic one, should be enticed to revisit the destination also at the end of the crisis without returning to international trips. For this reason, destination at regional or sub-regional level should perform specific analysis of the domestic demand to better understand its trends and outlooks. To this end, the results of this work show that despite a regional analysis is capable of offering good insights, a narrower sub-regional approach at destination level presents deeper understandings. By analysing three different destinations belonging to the same province (South Tyrol) in northern Italy, we have discovered that the demand for DT differs depending on the destination being a mountain one with two strong seasons, a urban destination open all-year-around, or a lake destination with a longer summer season.

Concerning MIDAS regression, in this paper we have shown how MIDAS models prove to be a suitable approach for obtaining estimates and forecasts of DTIs. MIDAS regressions allow to combine information included in variables sampled at different frequencies in a very effective way. The use of functional constraints on parameters associated to higher-frequency variables allows to make them compatible with the lower-frequency ones and to contain parameter proliferation (Rufino, 2019). By avoiding time-averaging techniques, based on the aggregation of higher frequency variables, the information included in the latter and the timing of the same is preserved (Ghysels et al., 2006). Accordingly, MIDAS approach exploits a much larger information set with high level of flexibility and efficiency compared to other econometric approaches finalized to deal with mixed sampled-data, such as state space models. In this specific study, MIDAS allow a combined use of economic and financial variables to estimate and predict DTIs. The use of a such rich information set reduces uncertainty in the DTI's nowcast/forecast. In this specific study, MIDAS models allow a combined use of economic and financial variables to estimate and predict DTIs. The use of a such rich information set reduces uncertainty in the DTIs nowcast/forecast.

In this regard, referring to yearly DTIs data (instead of monthly or quarterly) is justified by the need to obtain prudential forecasts capable to reduce errors implied by uncertainty in the tourism sector due to COVID-19.

References

Alvarez-Diaz, M., D'Hombres, B., Ghisetti, C., and Pontarollo, N. (2020). Analysing domestic tourism flows at the provincial level in spain by using spatial gravity models. *International Journal of Tourism Research*, 22(4):403–415.

Andreou, E., Ghysels, E., and Kourtellos, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics*, 158(2):246–261.

Athanasopoulos, G. and Hyndman, R. J. (2008). Modelling and forecasting australian domestic tourism. *Tourism Management*, 29(1):19–31.

Bernini, C., Cerqua, A., and Pellegrini, G. (2020). Endogenous amenities, tourists' happiness and competitiveness. *Regional Studies*, 54(9):1214–1225.

Bourdin, S., Jeanne, L., Nadou, F., and Noiret, G. (2021). Does lockdown work? a spatial analysis of the spread and concentration of covid-19 in italy. *Regional Studies*, pages 1–12.

Cafiso, G., Cellini, R., and Cuccia, T. (2018). Do economic crises lead tourists to closer destinations? italy at the time of the great recession. *Papers in Regional Science*, 97(2):369–386.

Clements, M. P. and Galvão, A. B. (2009). Forecasting us output growth using leading indicators: An appraisal using midas models. *Journal of Applied Econometrics*, 24(7):1187–1206.

Dolnicar, S. and Zare, S. (2020). Covid19 and airbnb–disrupting the disruptor. *Annals of tourism research*.

EUROSTAT (2021). Tourism statistics - nights spent at tourist accommodation establishments.

Foroni, C., Marcellino, M., and Schumacher, C. (2015). Unrestricted mixed data sampling (midas): Midas regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pages 57–82.

Gálvez, T. G., Romero, J. M., and Devesa, M. J. S. (2014). Measuring and analysing domestic tourism: the importance of an origin and destination matrix. *Tourism economics*, 20(3):451–472.

Ghysels, E., Kvedaras, V., and Zemlys, V. (2016). Mixed frequency data sampling regression models: the r package midasr. *Journal of statistical software*, 72(1):1–35.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, 76(3):509–548.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1-2):59–95.

Ghysels, E., Sinko, A., and Valkanov, R. (2007). Midas regressions: Further results and new directions. *Econometric reviews*, 26(1):53–90.

Havranek, T. and Zeynalov, A. (2021). Forecasting tourist arrivals: Google trends meets mixed-frequency data. *Tourism Economics*, 27(1):129–148.

Huybers, T. (2003). Domestic tourism destination choices—a choice modelling analysis. *International Journal of Tourism Research*, 5(6):445–459.

ISTAT (2021a). Conto satellite del turismo in italia: anticipazione anno 2020.

ISTAT (2021b). Viaggi e vacanze in italia e all'estero (anno 2020).

Kim, H.-R., Yi, C., and Jang, Y. (2019). Relationships among overseas travel, domestic travel, and day trips for latent tourists using longitudinal data. *Tourism Management*, 72:159–169.

Koenig, E. F., Dolmas, S., and Piger, J. (2003). The use and abuse of real-time data in economic forecasting. *Review of Economics and Statistics*, 85(3):618–628.

Massidda, C. and Etzo, I. (2012). The determinants of italian domestic tourism: A panel data analysis. *Tourism Management*, 33(3):603–610.

Munyanyi, W. and Chiromba, C. (2015). Tax incentives and investment expansion: evidence from zimbabwe's tourism industry. *AD-minister*, (27):27–51.

OECD (2019). Pressure under: The squeezed middle class. *OECD*, 169.

Rufino, C. C. (2019). Nowcasting philippine economic growth using midas regression. *DLSU Business & Economics Review*, 29(1):14–23.

Statista (2021). Cinternational and domestic tourism contribution to italian gdp 2014-2020.

UNWTO, W. T. O. (2020). *World Tourism Barometer*, volume 18. Madrid: World Tourism Organization Press.

UNWTO, W. T. O. (2021a). *International Tourism Highlights, 2020 Edition*, volume 18. Madrid: World Tourism Organization Press.

UNWTO, W. T. O. (2021b). Vaccines and reopen borders driving tourism's recovery. 4th october 2021.

WTTC (2018). Tourism importance and economic impact.

Appendix

A Supplementary Figures and Tables

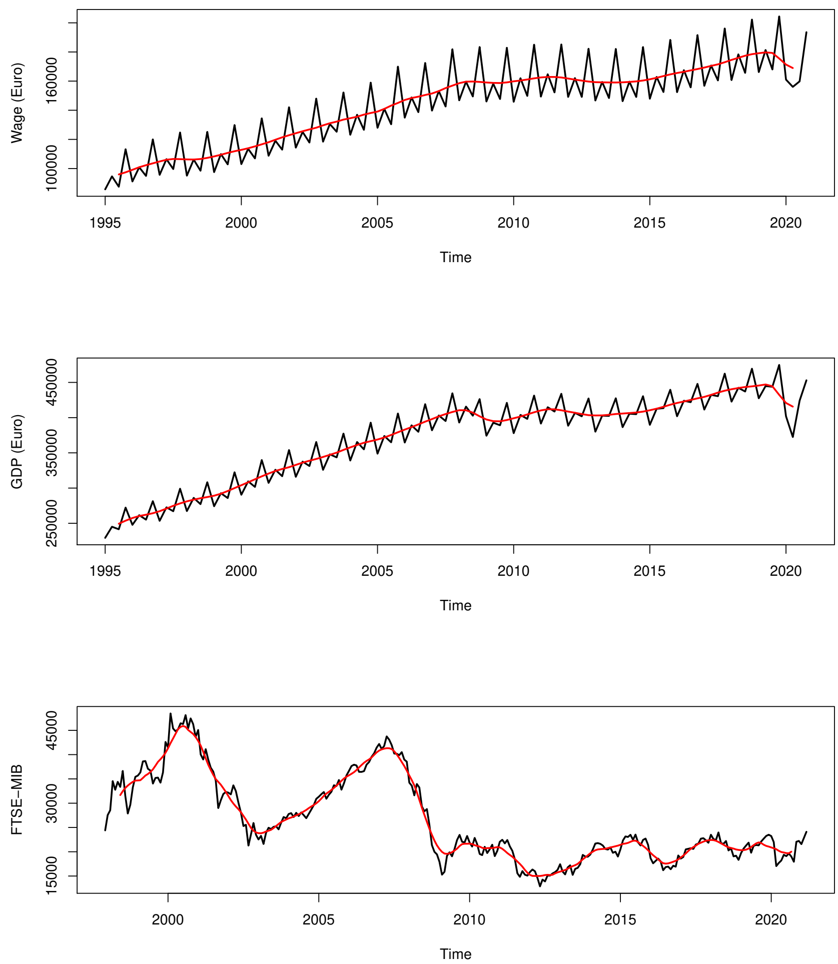


Figure 16: The first panel shows the Italian row and trend (red line) WS together with the series trend. Similarly, the second panel shows the Italian GDP and the third panel the monthly FTSE-MIB. *Data source*: ISTAT and Yahoo finance

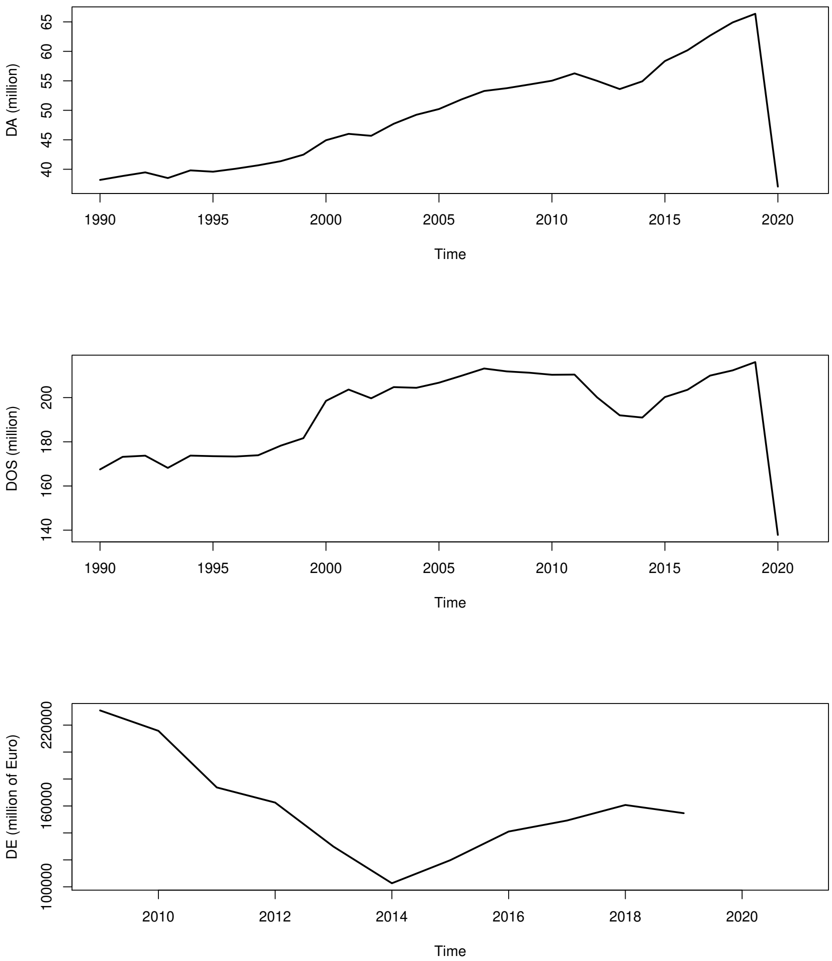


Figure 17: Italian yearly DA (first panel), DOS (second panel) and DTE (third panel). *Data source*: ISTAT and Eurostat

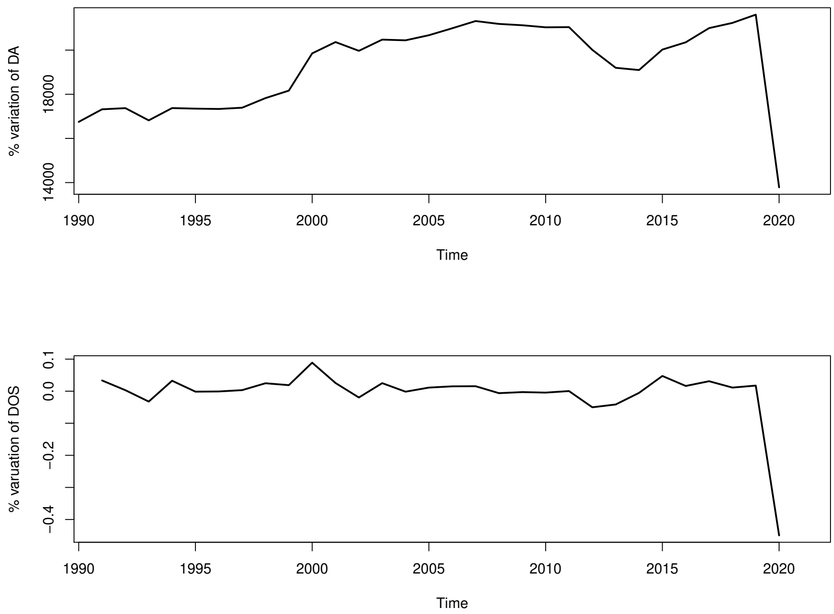


Figure 18: Italian yearly variation of DA (first panel) and DOS (second panel). *Data source*: ISTAT

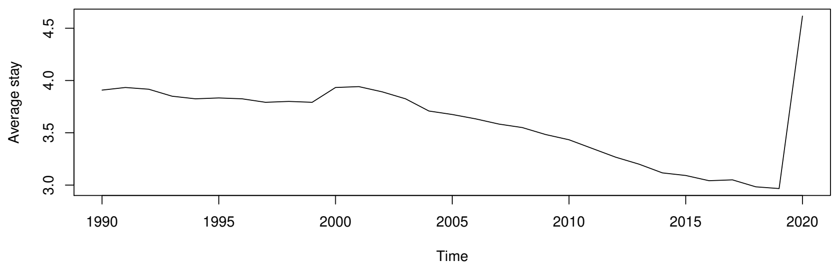


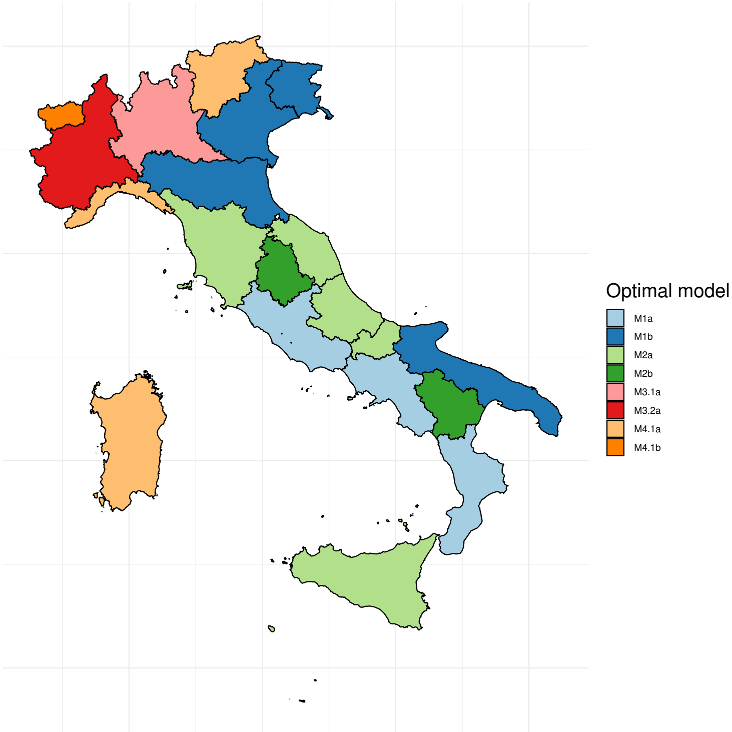
Figure 19: Italian yearly average length of stay of domestic travellers *Data source*: ISTAT

Table 2: MSEs and MAPEs of the out-of-sample forecasts obtained from different MIDAS specifications

Model	DTE		DA		DOS	
	MSE	MAPE	MSE	MAPE	MSE	MAPE
M1a	16480.208	3.264	2058.323	7.453	7256.560	7.256
M1b	7473.341	2.198	2053.945	7.445	7172.076	7.213
M2a	24201.821	3.956	2933.028	8.896	8152.802	7.691
M2b	49734.683	5.671	**2030.364**	**7.098**	7166.278	7.210
M3.1a	4082.840	1.625	2763.905	8.636	7065.586	7.160
M3.1b	**1810.496**	**1.082**	2578.590	8.342	-	-
M3.2a	4083.005	1.625	2588.117	8.357	-	-
M3.2b	1810.958	1.083	-	-	-	-
M4.1a	22480.620	3.813	2717.693	8.564	6854.995	7.052
M4.2a	3812.250	1.570	-	-	**5921.730**	**6.554**
M4.3a	14429.991	3.055	2842.489	8.758	7281.708	7.268
M5.1a	8011.882	2.276	27539.820	27.261	6881.959	7.066
M5.2b	15287.455	3.144	3091.177	9.133	8137.159	7.683

Table 3: MSE and MAPE out-of-sample of MIDAS models employed to explain DTIs

Model	Expenditure		Arrivals		overnight stays	
	MSE	MAPE	MSE	MAPE	MSE	MAPE
M1a	16480.208	3.264	0.577	9.953	0.487	10409
M1b	7473.341	2	**0.557**	**9.775**	0.460	10.118
M2a	24201.821	3.956	0.594	1.0099	0.508	10628
M2b	49734.683	5.671	0.621	1.0324	0.529	10852
M3.1a	4082.840	1.625	0.607	1.0203	0.453	10034
M3.1b	1810.496	1.082	0.606	1.0199	-	-
M3.2a	4083.005	1.625	0.613	1.0252	-	-
M3.2b	1810.958	1.082	-	-	-	-
M4.1a	2248.062	3.813	0.579	9.971	**0.451**	**10018**
M4.2a	**3812.25**	**1.57**	-	-	0.468	10201
M4.3a	14429.991	3.055	0.614	10.261	0.466	10183
M5.1a	8011.882	2.276	50.85	295.415	360158624	8950958
M5.2b	15287.455	3.144	0.618	1.0294	0.467	10195



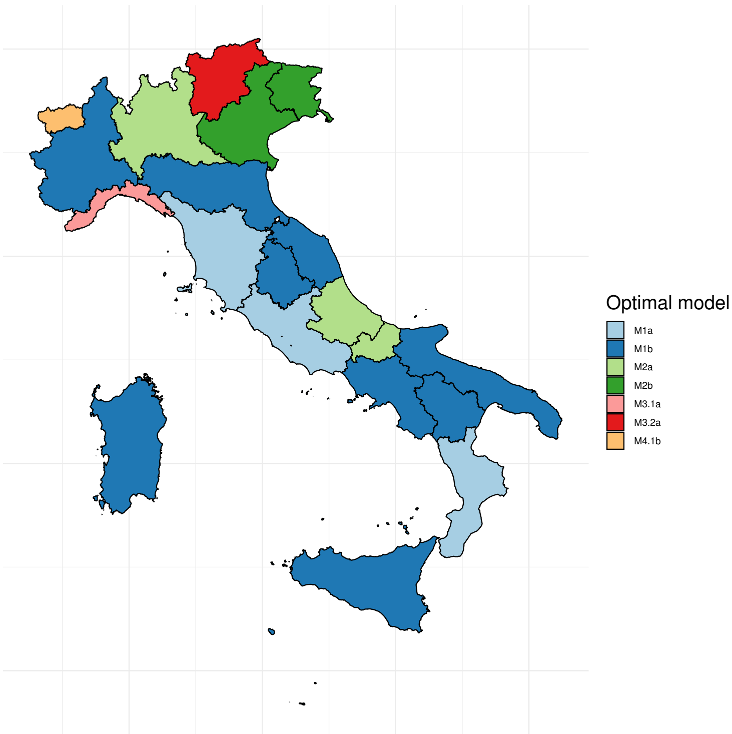


Figure 20: Optimal models for DA (top panel) and DOS (bottom panel)

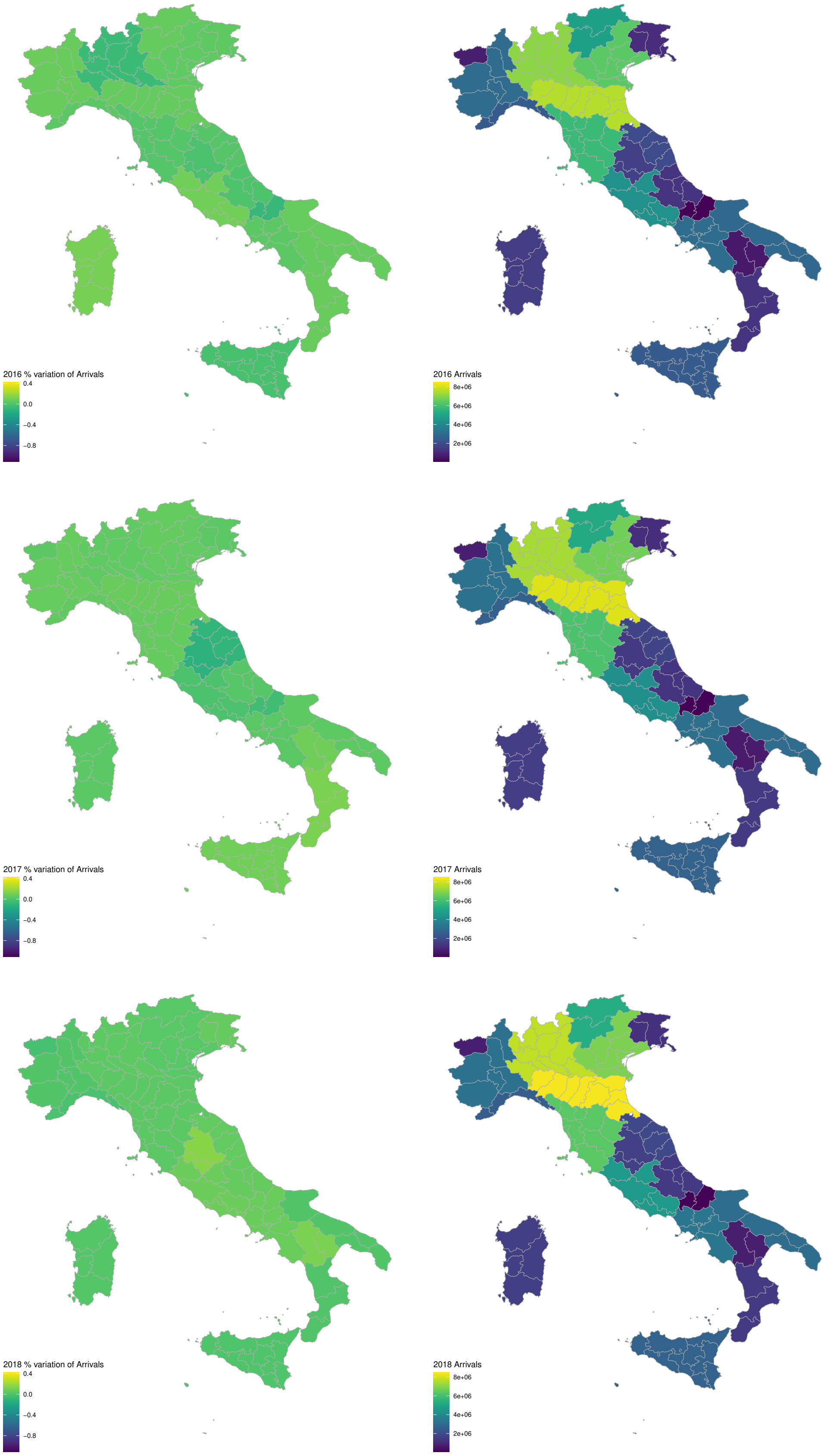


Figure 21: Regional log ratio (variation) of DA (left panel) and DA (right panel) 2016-2018

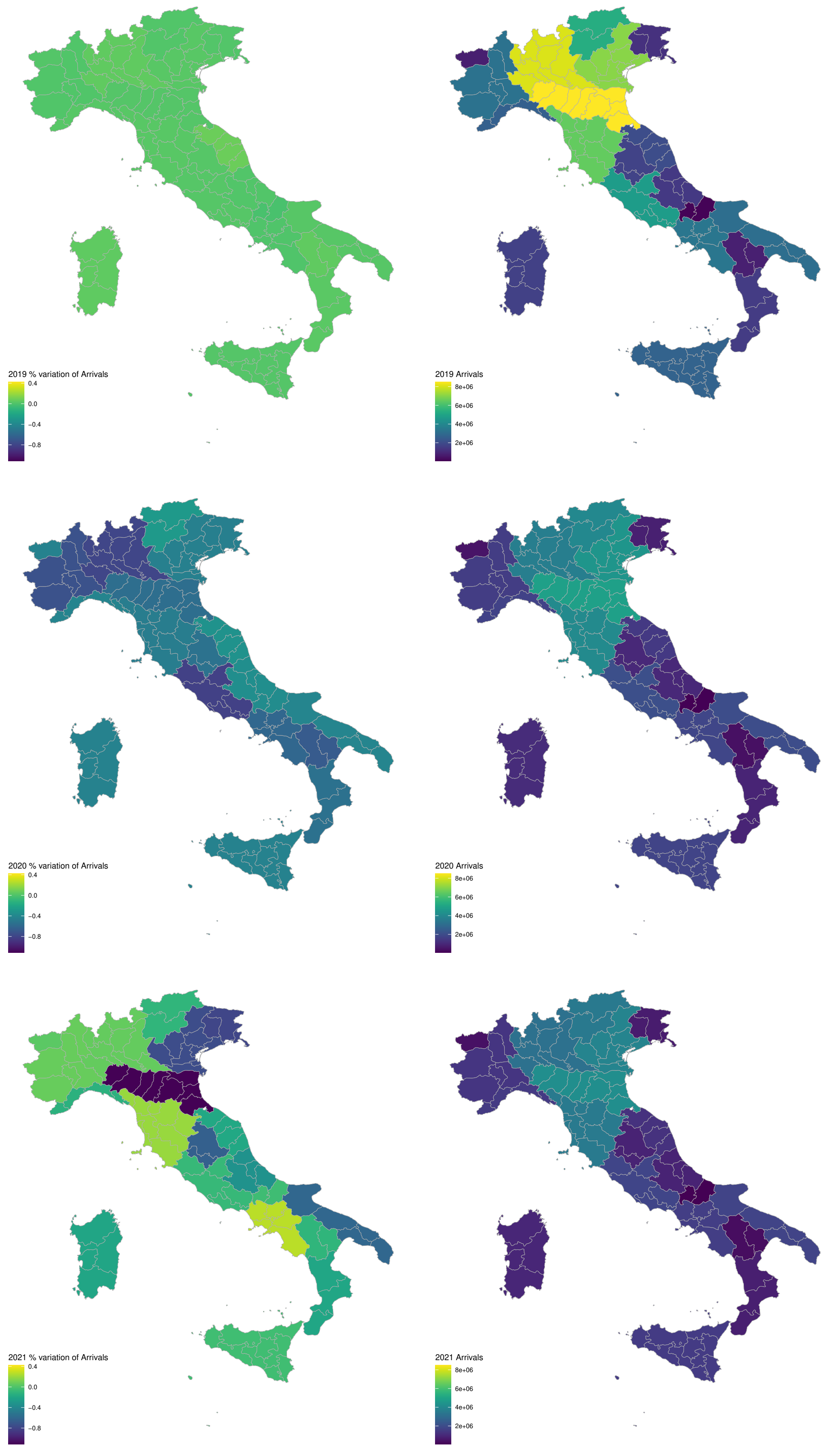


Figure 22: Regional log ratio (variation) of DA (left panel) and DA (right panel) 2019-2021

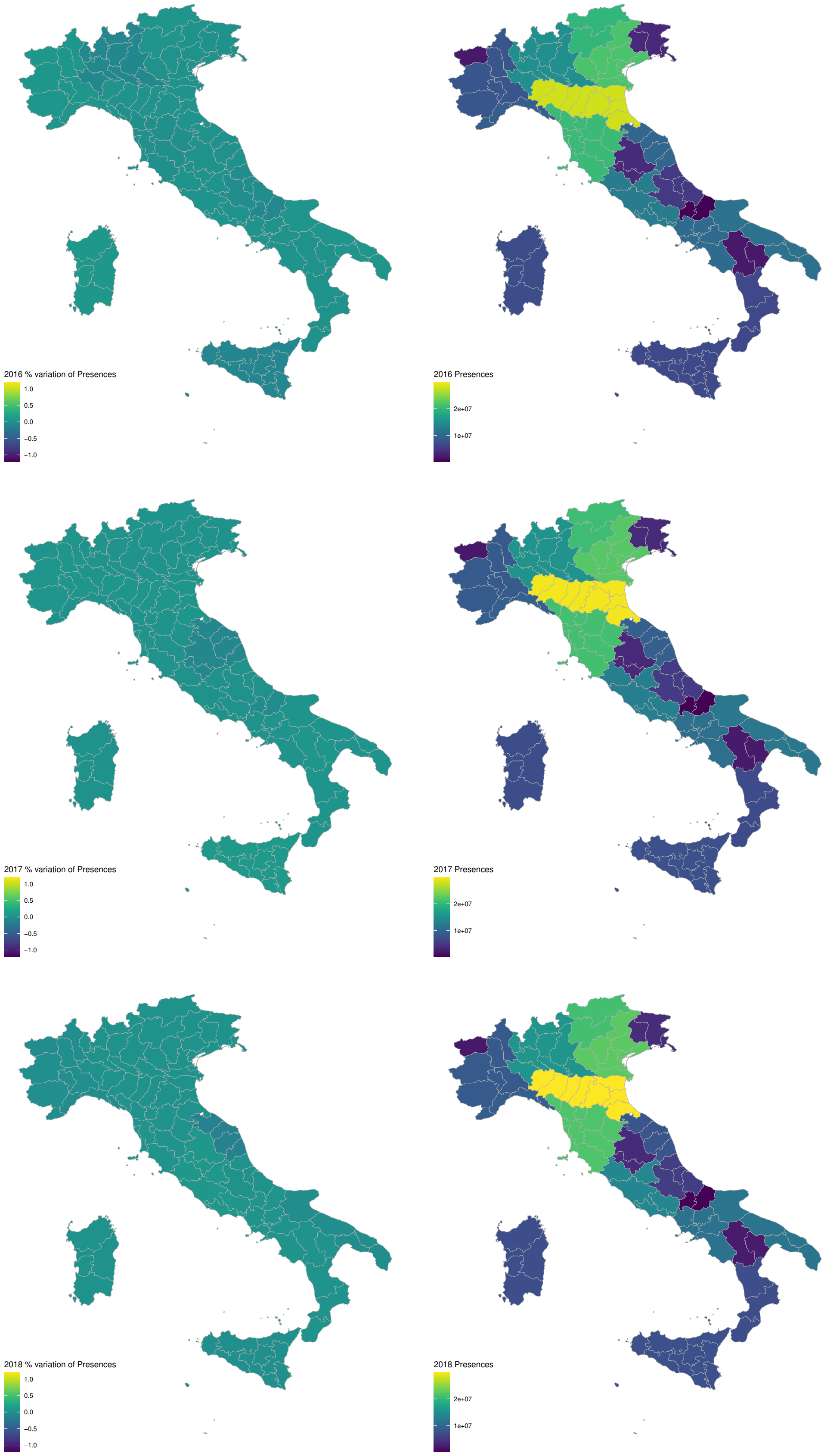


Figure 23: Regional log ratio (variation) of DOS (left panel) and DOS (right panel) 2016-2018

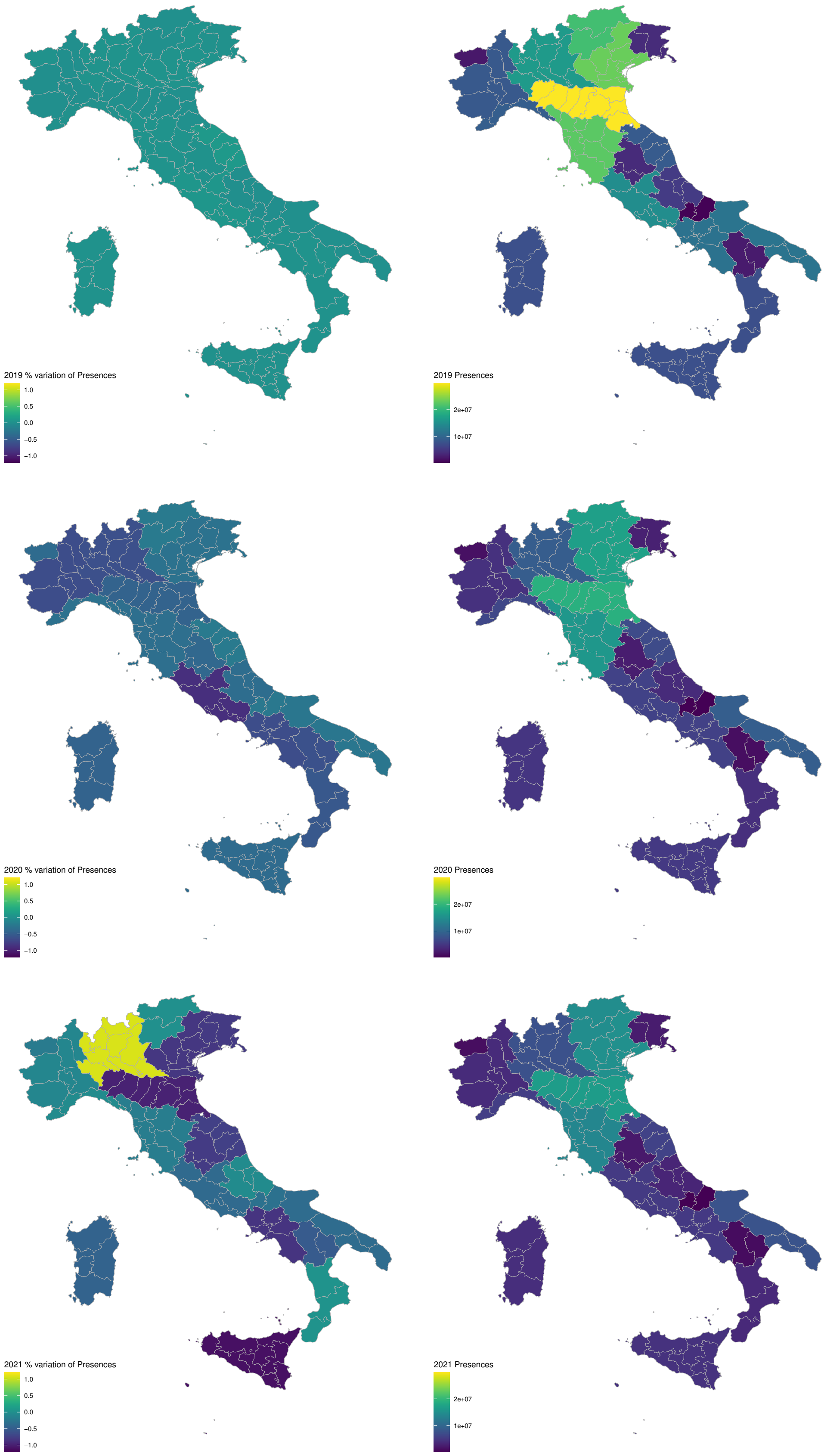


Figure 24: Regional log ratio (variation) of DOS (left panel) and DOS (right panel) 2019-2021

B R Code

Here we propose the code used for nowcasting and forecasting of domestic tourists expenditures. Note that we refer to the steps shown in Figure 4.

```
#########
# STEP 0
#########
# LOAD LIBRARY

library(midasr)
library(openxlsx)
library(TSstudio)
library(ggfortify)
library(tseries)
library(forecast)

#########
# STEP 1
#########
# VARIABLE DEFINITION (given preliminary importation)
# we set the in-sample to be from 2009 up to 2018
# Italian domestic expenditure
yy <- window(I(italy_exp/100), start=1, end=10)
# Italian wage and salaries
xx <- window(italy_wage, start=5, end=44)
# Italian GDP
zz <- window(italy_gdp, start=5, end=44)
# trend variable
trend <- 1:length(yy)
# Italian domestic overnight stays, monthly observed
rr2 <- window(PRESENCE, start=c(2009,1), end=c(2018,12))

#########
# STEP 2
#########
# Different MIDAS model specification
# M1a
M1a <- midas_r(yy~trend+mls(xx,0:3,4), start=NULL)
M1b <- midas_r(yy~trend+mls(zz,0:3,4), start=NULL)
M2a <- midas_r(yy~trend+mls(yy,1,1)+mls(xx,0:3,4), start=NULL)
M2b <- midas_r(yy~trend+mls(yy,1,1)+mls(zz,0:3,4), start=NULL)
M3.1a <- midas_r(yy~trend+mls(xx,0:7,4, nealmon),
                 start=list(xx=c(1,2)))
M3.1b <- midas_r(yy~trend+mls(zz,0:7,4, nealmon),
                 start=list(zz=c(250,1)))
M3.2a <- midas_r(yy~trend+mls(xx,0:7,4, nbeta),
                 start=list(xx=c(5,0.1,4)))
M3.2b <- midas_r(yy~trend+mls(zz,0:7,4, nbeta),
                 start=list(zz=c(10,0.1,6)))
M4.1a <- midas_r(yy~trend+mls(zz,0:7,4, nealmon)+
                 mls(xx,0:7,4, nealmon),
                 start=list(xx=c(0,0),zz=c(1,2)))
M4.1b <- midas_r(yy~trend+mls(rr2,0:11,12, nealmon)+
                 mls(xx,0:7,4, nealmon),
                 start=list(xx=c(0,0),rr2=c(10,20)))
M4.2a <- midas_r(yy~trend+mls(zz,0:7,4, nbeta)+mls(xx,0:7,4, nbeta),
                 start=list(xx=c(1,-0.2,1),zz=c(1,-0.2,0.7)))
```

```
M5.1a <- midas_r(yy~trend+mls(yy,1,1)+mls(zz,0:7,4, nealmon)+
                 mls(xx,0:7,4, nealmon),
                 start=list(xx=c(1,2),zz=c(10,20)),
                 weight_gradients = list(nealmon_gradient))
M5.1b <- midas_r(yy~trend+mls(yy,1,1)+mls(rr2,0:11,12, nealmon)+
                 mls(xx,0:7,4, nealmon),
                 start=list(xx=c(0,0),rr2=c(100,20)))


# check convergence of the different models
# (here just a case is proposed)
out3.1a$convergence

# test model restriction (here just for a case is proposed)
hAh_test(out3.2a)
hAhr_test(out3.1a)

#########
# STEP 3
#########
# Compare and select the best model
full_data <- list(xx=window(italy_wage, start=5, end=52),
                  zz=window(italy_gdp, start=5, end=52),
                  rr2=window(PRESENCE, start=c(2009,1), end=c(2019,12)),
                  yy=window(I(italy_exp/100), start=1, end=11),
                  trend=1:12)

insample <- 1:length(yy)
outsample <- (1:length(full_data$yy))[-insample]
avgf <- average_forecast(list(out1a,out1b,out2a,out2b,out3.1a,out3.1b,
                      out3.2a, out3.2b, out4.1a, out4.1b,out4.2a,
                      out5.1a,out5.1b),
                  data=full_data,insample=insample, outsample=outsample,
                      type="fixed",
                      measures=c("MSE","MAPE","MASE"),
                      fweights=c("EW","BICW","MSFE","DMSFE"))
sqrt(avgf$accuracy$individual$MSE.out.of.sample)
sqrt(avgf$accuracy$individual$MAPE.out.of.sample)

#########
# STEP 4
#########
# define again the variables including also 2019 for expenditure
# and all the other observed quarterly values for GDP and wage
yy <- window(I(italy_exp/100), start=1, end=11)
xx <- window(italy_wage, start=5, end=48)
zz <- window(italy_gdp, start=5, end=48)
trend <- 1:length(yy)

# Estimate again the best model
out3.1b <- c()
out3.1b <- midas_r(yy~trend+mls(zz,0:7,4, nealmon),
          start=list(zz=c(250,1)))
summary(out3.1b)

#########
# STEP 4
#########
```

```r
# nowcasting (2020)
f3.1b <- forecast(out3.1b, newdata=list(trend=c(12),
                     xx=italy_wage[c(49:52)],zz=italy_gdp[c(49:52)]))
f3.1b$mean

# forecasting (2021)
# prepare data including the nowcast in 2020
yy3.1b <- window(c(yy, f3.1b$mean), start=1, end=12)
xx3.1b <- window(italy_wage, start=5, end=52)
zz3.1b <- window(italy_gdp, start=5, end=52)
trend3.1b <- 1:length(yy3.1b)
out3.1b.1 <- midas_r(yy3.1b~trend3.1b+mls(zz3.1b,4+0:7,4, nealmon),
                     start=list(zz3.1b=c(250,1)))
summary(out3.1b.1)

f3.1b.1 <- forecast(out3.1b.1, newdata=list(zz3.1b=rep(NA,4),
                     trend3.1b=c(13)))
f3.1b.1$mean
# put observed and forecast data all together
yy3.1b_f <- window(c(yy, f3.1b$mean, f3.1b.1$mean), start=1, end=13)
```