Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The Daily Economic Indicator: tracking economic activity daily during the lockdown

Nuno Lourenço, António Rua

Banco de Portugal, Nova SBE, Portugal

ARTICLE INFO

JEL classification:
C22
C58
E32

Keywords:
Daily economic index
High-frequency
Measurement of economic activity
Factor model
COVID-19

ABSTRACT

The SARS-CoV-2 outbreak made clear the urgent need to depart from traditional statistics, typically released with a lag and available at a relatively low frequency. This led to unparalleled efforts to put forward high-frequency indicators to track economic developments timely. By resorting to non-traditional data sources, we propose a novel daily economic indicator to track economic activity in Portugal. It corresponds to the latent variable of a set of daily series within a factor model framework. We find a sudden and sharp drop in economic activity in mid-March 2020, when the lockdown of several activities was declared due to the COVID-19 pandemic. Since in this approach we address the complexities of high-frequency data without further smoothing, we are able to identify sudden changes of economic activity in a timely and daily manner in contrast with other approaches.

1. Introduction

Monitoring economic developments in real-time is often grounded in the assessment of a large panel of variables available at a monthly or quarterly frequency. These indicators are also made available with a release lag, thus not proving to be suitable to track sudden changes of economic activity. The use of high-frequency data, i.e., data recorded at infra-monthly intervals, dates back to the 1920s. In fact, statisticians and econometricians at the time relied on weekly data to support policy analysis. Fisher (1923) pioneered this work, by proposing a weekly index number of wholesale prices, based on price quotations of 200 commodities. Early efforts to model the statistical properties of high-frequency data were remarkable. In this regard, Crum (1927) studied the series of weekly bank debits outside New York City from 1919 to 1926 and developed a seasonal adjustment based on the median-link-relative method suggested by Persons (1919).

Nowadays, with the rapid dissemination of statistical data that are being processed weekly, daily or even hourly, the information set available to economic agents has been made progressively larger. Nevertheless, there is still room to improve on data collection while making it available to economic policy agents as most of the data are typically owned by private firms. Furthermore, relying on high-frequency data poses challenges to empirical work (e.g., multiple seasonal patterns and non-integer periodicities, calendar effects, high volatility and a marked irregular component, existence of outliers). Therefore, taking on board such a high-frequency data into the analysis of short-term economic prospects is not straightforward which explains why its use has been rather limited.

The SARS-CoV-2 outbreak has made clear the urgent need to depart from traditional statistics and rely instead on high-frequency non-conventional data sources. One of the reasons behind this shift relates to the rapid change of economic conditions. The unprecedented nature of the shock makes it difficult to assess the impact of the COVID-19 pandemic on the real economy. In this regard, a lot of effort is being devoted to measure economic activity at a high-frequency so as to help designing sound economic policies.

To the best of our knowledge, this is the first work that develops a daily economic indicator (DEI) to assess the impact of the lockdown on economic activity at a high-frequency. This is achieved by combining several tools and exploiting non-conventional data sources. Based on those series, a latent factor that can be regarded as an indicator of economic activity is estimated drawing on a factor model. This approach is well-rooted in the literature since the seminal work by Stock and Watson (1989). Given the high-frequency, a pre-treatment of the series has to be performed to cope with data issues such as seasonality and calendar effects, which poses a key challenge in daily data. For that purpose,
we resort to STL (Seasonal-Trend decomposition based on Loess) and RegARIMA methods. We find marked seasonal intra-weekly patterns in the series that have to be accounted for as well as the presence of significant calendar effects.

Other attempts to obtain high-frequency composite indicators include Lewis et al. (2020) and Bundesbank (2020). Lewis et al. (2020) proposed a weekly economic index (WEI) to track the rapid economic developments associated with the response to the novel coronavirus in the United States. It corresponds to the single factor estimated from ten weekly series. It should be noted, however, that this is a weekly indicator in contrast to what we put forward. Furthermore, corrections for moving holidays are not carried out. Similarly, the Bundesbank (2020) developed a weekly activity index (WAI) for the German economy based on a dataset comprising seven high-frequency indicators. Not only the frequency of the DEI is higher but also the WAI relies on 12-week growth rates of its rolling 12-week averages for each high-frequency indicator, which implies a strong smoothing of the indicator, thus not allowing to identify sudden changes in economic activity in a timelier manner.

The usefulness of several non-conventional data sources to track economic developments has been documented in the literature. This can be particularly relevant in the face of sudden disruptions of economic activity that pose challenges to the more traditional statistics. A very promising data source refers to that of card-based payments, as these provide an encompassing view of transactions in the economy. The applications on the use of payments data to track economic activity are quite extensive and include, inter alia, Carlens and Storgaard (2010) for Denmark, Barnett et al. (2016) for the United States, Duarte et al. (2017) for Portugal, Galbraith and Tkacz (2018) for Canada and Aprigliano et al. (2019) for Italy. In addition, transportation activities have also been used to gauge prospects on economic activity. In fact, as pointed out by Lahiri et al. (2004), transportation activities are very sensitive to the business cycle. However, their use to track business cycle fluctuations has been rather limited. One of the reasons behind this refers to a non-negligible release lag of these statistics, a situation that has been circumvented with the rise of systems with electronic toll collections that allow these statistics to be more promptly available (see also Lahiri and Yao (2012)). Moreover, Askitas and Zimmermann (2018) consider transportation activity performed by heavy transport vehicles across Germany to track the general state of the economy, while Fenz and Schneider (2009) use this type of data to predict Austrian exports. Interesting insights can also be derived from electricity consumption data due to its broad coverage and wide availability. In this respect, Arora and Lieskovsky (2014) show evidence of the usefulness of electricity consumption to trace GDP developments, which is also corroborated by the work of Aprigliano et al. (2019).

After an extensive data collection and a preliminary analysis, we end up with five daily series, namely card-based payments, road traffic of heavy commercial vehicles, cargo and mail landed, electricity consumption and natural gas consumption. These series present the strongest co-movement with GDP developments among the series assessed and allow to cover several important dimensions of economic activity as discussed above.

The proposed composite indicator presents statistically significant information content about current GDP developments. In particular, the evolution of economic activity during the lockdown is analysed. We find that the DEI started to decrease substantially in mid-March, at the time the Portuguese government declared the highest level of alert due to the COVID-19 pandemic followed a few days after by the declaration of the State of Emergency in the entire Portuguese territory. Such analysis reinforces the usefulness of high-frequency indicators such as the DEI. Naturally, the use of high-frequency indicators is not limited to the analysis of particular events such as the lockdown, but can also be used for monitoring real activity in a timely manner and, thus, anticipating the path of low frequency variables which is of utmost importance for policymakers and private economic agents.

The remainder of the paper proceeds as follows. In section 2, we present the econometric framework pursued. Section 3 describes the data. The daily economic indicator is put forward and discussed in section 4. Section 5 concludes.

2. Econometric framework

Despite the potential usefulness of considering high-frequency indicators, the use of such series involves tackling several issues. In fact, daily raw data typically present strong seasonal patterns. However, daily or weekly series are available for much shorter time spans than monthly or quarterly data. Naturally, this means that intra-annual seasonality can be hard or even impossible to estimate using such high-frequency data as it typically encompasses just a few years. To overcome this issue, one can consider year-on-year growth rates (as in, for instance, Lewis et al. (2020) for weekly data). Such data transformation is easily interpretable and allows to accommodate slowly moving seasonality. One should mention that the computation of year-on-year growth rates corresponds to a filtering procedure that tends to smooth out the series by reducing the importance of high frequencies but also entails a phase shift (see, for example, Quenneville and Findley (2012)). Note that a phase delay is also present in standard alternative seasonal adjustment procedures (as discussed by Findley and Martin (2006)).

However, although we are comparing a given date vis-à-vis the same date in the previous year, the weekday might differ. In fact, in many economic series, the behaviour differs across the seven days of the week. Such distinction can be particularly marked between, for example, the weekend and working days. To account for intra-weekly seasonality we resort to STL (see also Bergmeier et al. (2016) and Hyndman and Athanasopoulos (2014)).

Furthermore, there are specific events that occur every year which one should also take into account. In this respect, national holidays are often linked to a date but there are moving holidays such as Easter which do not always fall on the same date. Hence, after coping with intra-weekly seasonality, the impact of those calendar effects is purged resorting to seasonal RegARIMA models.

After the pre-treatment of the data, one can proceed with factor model estimation to obtain the daily economic indicator. In the remainder of the section, we run through the above-mentioned building blocks in detail.

2.1. STL

The STL is a non-parametric method proposed by Cleveland et al. (1990). It allows to decompose a seasonal time series ($Y_t$) into trend ($T_t$), seasonal ($S_t$) and remainder ($R_t$) components, i.e,

$$Y_t = T_t + S_t + R_t, t = \{1, \ldots, N\}. \quad (1)$$

STL is a very versatile and robust method for decomposing time series and it has several advantages over other decomposition methods. It can handle any type of seasonality, meaning that it is not restricted to monthly and quarterly data. The seasonal component is allowed to change over time, with the rate of change controlled by the parameter seasonal width. The smoothness of the trend-cycle is also user-defined. It can be specified to be robust to outliers, so that sporadic unusual

---

1 For an overview of different seasonal adjustment procedures for daily and weekly data see, for example, Ladadji et al. (2018).

2 Naturally, as the series become longer in the future other alternatives may be considered such as the procedure proposed by Ollech (2018).

3 Furthermore, in our case, as well as in Lewis et al. (2020), the use of year-on-year growth rates also allows to cope with the non-stationarity of the series which is essential to proceed with the estimation of a factor model.

4 Note that if one is interested in addressing multiple periodicities, one can use STL sequentially (see Cleveland et al. (1990) and Ollech (2018)).
observations will not affect the estimates of the trend-cycle and seasonal components. Although the above model corresponds to an additive decomposition, a multiplicative decomposition can be obtained by first taking logs of the data.

The STL operates by iterating through smoothing of the seasonal and trend components where all smoothing is done resorting to Loess. For the seasonal smoothing, observations are separated into cycle-subseries. These subseries are separately smoothed and then recomposed. The trend smoothing is done after removing the previously estimated seasonal component. This is iterated until convergence. The STL also has an outer loop which, after each inner loop, computes robustness weights that are passed on to the inner loop, in order to cope with outliers.

2.1.1. Loess

Loess was initially proposed by Cleveland (1979) and developed further by Cleveland and Devlin (1988). Loess combines the simplicity of linear least squares regression with the flexibility of nonlinear regression. This is achieved by fitting, at each point in the data set, a low-degree polynomial to a localized subset of the data. The polynomial is estimated using weighted least squares, giving more weight to the data points nearer the point of estimation and less weight to the data points that are further away. The value of the regression function for a specific data point is obtained by evaluating the local polynomial at that data point. This is done for each of the N data points.

The usual weight function used for Loess is the tri-cube weight function

\[
W(d) = \begin{cases} 
(1 - d^3) & 0 \leq d < 1 \\
0 & d \geq 1 
\end{cases} 
\] (2)

where \( d \) is the distance of a given data point from the point of estimation, scaled to lie in the range from 0 to 1. Therefore, the weight for a specific point \( t_i \) in any localized subset of data is obtained by evaluating the weight function at the distance between that point and the point of estimation \( t \), after scaling the distance so that the maximum absolute distance over all of the points in the subset of data is exactly one. That is,

\[
w(t_i) = W\left(\frac{|t_i - t|}{|q_i - t|}\right). 
\] (3)

This means that observations close to \( t \) have the largest weights while decreasing as the distance increases, becoming zero at the \( q^\text{th} \) farthest point. Note that the parameter \( q \) determines the number of neighbouring observations included in the local regression and therefore controls the smoothing degree. The choice regarding the degree of smoothing can be informed by diagnostic methods (see Cleveland and Terpenning (1982) and Cleveland et al. (1990)).

2.1.2. Inner loop

The inner loop is basically an iteration between seasonal smoothing (smoothing of the cycle-subseries) and trend smoothing. The inner loop iteratively updates the trend and seasonal components. This is done by subtracting the current estimate of the trend from the raw series. The time series is then partitioned into cycle-subseries. The cycle-subseries are Loess smoothed and then passed through a low-pass filter. The seasonal components are the smoothed cycle-subseries minus the result from the low-pass filter. These are subtracted from the raw data. The result is Loess smoothed, which becomes the trend. What is left is the remainder. More formally, the \( k^\text{th} \) iteration of the inner loop is composed by the following steps:

1. **Detrending**: The first step is to detrend the series, calculating \( Y_t - T_t \).

2. **Cycle-subseries smoothing**: The detrended time series is broken into cycle-subseries. For example, daily data with a weekly seasonality would yield seven cycle-subseries: all Sundays will be one time series, all Mondays a second, etc. Each cycle-subseries is smoothed by Loess with a span of \( q \). The smoothed values yield a temporary seasonal time series \( C_{s,k+1} \).

3. **Low-pass filtering of smoothed cycle-subseries**: To prevent low-frequency power from entering the seasonal component, a low-pass filter is applied on \( C_{s,k+1} \) yielding \( L_{s,k+1} \). The filter is composed of moving average and Loess filters.

4. **Detrending of smoothed cycle-subseries**: \( Y_t^{(k+1)} = C_{s,k+1} - L_{s,k+1} \). This is the \((k + 1)^{\text{th}}\) estimate of seasonal component.

5. **Deseasonalizing**: The original series is deseasonalized by subtracting the seasonal component computed in the previous step, that is, \( Y_t - S_t^{(k+1)} \).

6. **Trend smoothing**: Smoothing by Loess the deseasonalized time series results in \( T_t^{(k+1)} \), the \((k + 1)^{\text{th}}\) estimate of the trend component.

2.1.3. Outer loop

If outlying values are apparent in the time series, one may wish to run the outer loop to down-weight these values. In the outer loop, robustness weights are assigned to each data point depending on the size of the remainder. This allows for reducing or eliminating the effects of outliers.

After running the inner loop, given the estimates for the trend and seasonal components, the remainder can be obtained simply by

\[
R_t = Y_t - T_t - S_t. 
\] (4)

Let

\[
h = 6 \times \text{median}(|R_t|). 
\] (5)

Then the robustness weight at time \( t \) is calculated as

\[
\rho_t = B(|R_t|)/h 
\] (6)

where \( B \) is the bisquare weight function

\[
B(u) = \begin{cases} 
(1 - u^2)^2 & 0 \leq u < 1 \\
0 & u \geq 1 
\end{cases} 
\] (7)

For the Loess regressions in the steps 2 and 6 of the inner loop, the neighborhood weight is multiplied by the robustness weight, which allows to dampen the influence of outliers.

2.2. RegARIMA

To determine the impact of moving holidays on the time series, we consider a RegARIMA model (see, for example, Findley et al. (1998)):

\[
q_p(L) \Phi_p(L)Y_t[(1 - L)^q(1 - L)^D] Y_t - \sum_{j=1}^{r} \beta_j Z_{jt} = \theta_q(L) \Phi_q(L) \varepsilon_t 
\] (8)

where \( L \) is the lag operator, \( p, q, d, s, P, D, Q \) are non-negative integers \((s \geq 2)\) and \( Z_{jt} \) is the set of independent variables. The parameters \( p, q, P, Q \) determine the order of the lag polynomials and \( d \) and \( D \) control for the non-seasonal and seasonal differencing, respectively. In the case of the daily series, \( s = 365.6 \). The parameter \( \beta_j \) captures the impact of regressor \( Z_{jt} \) on \( Y_t \). To capture the impact of moving holidays, regressors for these holidays are included in the set of independent variables. As the event may affect beyond the date of the event itself, a number of days preceding and/or following the event can also be considered in the set of regressors.\(^7\)

\(^5\) Loess or Lowess stands for “LOcally WEighted Scatterplot Smoother”.

\(^6\) In this case, the leap day, February 29, has to be dropped.

\(^7\) See, for example, Ladiray (2018) about impact models to estimate the before or after effect on the activity.
2.3. Factor model

After the above pre-treatment of the data, one can proceed into the estimation of the factor model. Let $X_t$ be a $M$-dimensional column vector of time series, observed for $t = \{1, \ldots, N\}$, that is, $X_t = [Y_{1t}, \ldots, Y_{Mt}]$. The variables in $X_t$ are represented as the sum of two orthogonal components: the common component, driven by a small number of unobserved common factors that accounts for most of the co-movement among the variables; and the idiosyncratic component, driven by variable-specific shocks.

The data generating process for $X_t$ admits a static factor representation written as:

$$X_t = \Lambda F_t + \xi_t$$

where $F_t = (f_{1t}, \ldots, f_{rt})'$ is a $(r \times 1)$ vector of latent factors, $\Lambda$ is a $(M \times r)$ matrix of unknown factor loadings and $\xi_t$ denotes a $M$-dimensional vector of idiosyncratic terms. This representation is without loss of generality as it can be shown that the dynamic factor model representation has an equivalent static factor formulation (see, for instance, Stock and Watson (2005)).

The space spanned by the latent factors can be estimated through the principal components estimator which has been shown to be consistent under relatively general assumptions (see Stock and Watson (1998, 2002b), Bai and Ng (2002) and Amengual and Watson (2007)).

The proposed DEI corresponds to the latent factor $f_{1t}$ driving the co-movements of the variables in $X_t$ (see also Lewis et al. (2020)).

3. Data

The lack of priors on which economic time series prove to be more relevant for tracking economic activity at a high-frequency basis motivated an unparalleled data collection effort. A comprehensive dataset comprising around two dozens of daily economic time series with a minimal sample period for in-sample analysis was built at the first place. The series were chosen to encompass several domains of the economy, thus signalling developments across different sectors.

We collected data on traffic from Instituto da Mobilidade e dos Transportes (IMT). In addition to total traffic, we also gathered disaggregated data on traffic of heavy commercial vehicles and light passenger vehicles. These traffic statistics provide a measure of daily traffic intensity as they record the number of vehicles passing through tolls. As many goods across Portugal are transported by road and Spain is the country’s main trading partner, these statistics may signal prospects on both the production and international trade. We considered daily traffic data since the beginning of 2012.

Also on transport activity, we gathered data on the movement of passengers at national airports, the number of aircrafts landed as well as freight movements (cargo and mail shipments) which are compiled by ANA Aeroportos de Portugal and the Portuguese Civil Aviation Authority (ANAC) and kindly provided by Statistics Portugal for the period since 2009.

Energy consumption developments are also of utmost relevance to assess the stage of economic activity in the industrial and domestic sectors. Therefore, we collected data on electricity consumption across different distribution grids (e.g. very high-voltage, high-voltage, medium-voltage) which was provided by EDP Distribuição. As these series are only available since the beginning of 2011, this set of information was complemented with aggregate electricity consumption (adjusted for temperature effects) released by Redes Energéticas Nacionales (REN).

We also collected data on natural gas consumption in the conventional market for the period since the start of 2013, which has also been provided by REN.

An additional source of data aiming at tracking the consumer behaviour and transactions in general refers to card-based payments and ATM cash withdrawals, carried out both with cards issued in Portugal and abroad. These series are compiled and have been provided by Banco de Portugal. Although these series are available for a longer time span, we will restrict henceforth the analysis to the period from January 1, 2009 to April 30, 2020.

4. Results

4.1. Preliminary analysis

Given the composition of the initial dataset and to avoid over-representation of any particular economic activity dimension, we pursue a parsimonious approach regarding the number of variables considered. In this respect, one should mention the acclaimed work by Stock and Watson (1989) who proposed a composite economic indicator for the United States based on a factor model using just four series (in the same vein, Azevedo et al. (2006) developed a composite indicator for the euro area activity also resorting to a small number of series). More recently, focusing on weekly data, Lewis et al. (2020) consider ten high-frequency indicators whereas Bundesbank (2020) take on board seven weekly series.

Having in mind the objective of obtaining a broadly based measure, we consider those variables that display more information content about the evolution of economic activity. We end up with five daily series: road traffic of heavy commercial vehicles; electricity consumption; natural gas consumption; cargo and mail landed; card-based payments. As in Lewis et al. (2020) we disregard financial data as the aim is not to track financial conditions, but the real activity instead.

4.2. Intra-weekly seasonality

As described in section 2, we firstly estimate the weekly seasonality using the robust STL. In particular, we take logs of the series and set the value of the smoothing parameter to 151, as in Ollech (2018). This means that around 3 years of data are used in Loess regressions.

---

8 The typical assumptions allow for some heteroskedasticity and limited dependence of the idiosyncratic components in both the time and cross-section dimensions, as well as for moderate correlation between the latter and the factors.

9 IMT currently manages 15 motorway concessions in Portugal from north to south, corresponding to the universe of existing concessions.

10 For this type of data, there are two discontinuities. Until October 31, 2015 there is no split between operations with cards issued in Portugal and abroad. From November 1, 2015 onwards data are available for cards issued in Portugal whereas data regarding cards issued abroad are available only since January 1, 2018.

11 In the former case, the high-frequency indicators include measures of same-store retail sales, steel production, fuel sales, railroad traffic, electricity consumption, an index of consumer sentiment, initial and continued claims for unemployment insurance, an index of temporary and contract employment and tax collections from paycheck withholdings. In the latter case, the series include electricity consumption, truck toll mileage, Google searches related to unemployment and short-time work, worldwide number of flights, air pollution and cash withdrawals.

12 In particular, the co-movement between the year-on-year quarterly growth rate of each daily series and GDP growth has been assessed to pre-select the variables.

13 In this section, we use the STL function in R, whereas in the remaining sections most of the results are obtained with Matlab.

14 In this respect, Cleveland et al. (1990) argue that it should be an odd number and at least 7, but beyond that there is no unique choice. As a robustness analysis, we considered lower parameter values allowing for a more rapidly evolving weekday effect. We find that the main results remain qualitatively unchanged.
Such a value is high enough to avoid including annual periodic movements into the intra-weekly seasonal factors while leaving some room for variation of the seasonal factors. Following Ollech (2018), we compute the smoothed periodogram for both the differenced original series and the differenced seasonally adjusted series (see Fig. 1). One can see that the spectral peaks associated with intra-weekly seasonality (at frequency $2\pi/7$ and its harmonics, denoted by the dashed lines in Fig. 1) have been filtered out.

In Fig. 2, we display the estimated seasonal factors for all indicators. For each variable, we plot the cycle-subseries of the weekly seasonal component. Each cycle-subseries is plotted separately against time. In particular, first the Sunday values are plotted, then the Monday values are displayed, and so forth. The midmean for each weekday corresponds to the horizontal line. As the series are log-transformed before seasonal adjustment, it implies that the seasonal component in Fig. 2 conveys approximately the percentage by which the data and the seasonally adjusted data differ on the original scale.

The immediate finding is that there is a strong weekly seasonal pattern in all series. As expected, in the case of the road traffic of heavy commercial vehicles, there is a striking difference between working days and the weekend. In particular, traffic on Sundays is 90 per cent lower than on an average day, while on Saturdays it is almost 60 per cent lower. One can also conclude that such magnitudes have fallen somewhat during the period under analysis. Concerning electricity consumption, usage on Sundays is 15 per cent lower while on Saturdays it is near 10 per cent lower than the average day. One should also mention that electricity consumption on Mondays seems to be a bit below the observed for the remaining working days, most probably reflecting the gradual resume of production activity after the weekend. The natural gas consumption exhibits a relatively similar seasonal pattern to the observed for electricity consumption. Both series are reflecting, to a large extent, the intra-weekly seasonality of industrial activity. In particular, gas consumption on Sundays is around 20 per cent lower while on Saturdays it is close to 15 per cent lower than the average day. Regarding cargo and mail landed, Sundays record a value 45 per cent below the average day but this figure has been decreasing over time. On Saturdays, cargo and mail landed are also below average and the gap has been widening. Mondays and Tuesdays correspond basically to an average day, while the remaining weekdays are characterized by cargo and mail landed above average. In what concerns card-based payments, purchases are 25 per cent lower on Sundays and 10 per cent higher on Fridays and Saturdays. The remaining weekdays behave close to an average day.

4.3. Calendar effects

After removing the intra-weekly seasonality, one proceeds with the estimation of the effects of Carnival and Easter (which involves two national holidays, the Good Friday and the Easter Sunday). Besides considering regressors for each of those days, to capture potential effects in the activity during the days before or after those events, we also consider regressors for these periods. In particular, we consider a window width of one week, that is, we consider three days before and after each event. In the case of Carnival, this means considering the period from Saturday to Friday (as Carnival occurs on a Tuesday) whereas for Easter the period ranges from Tuesday to Wednesday next week. Regarding the specification of the RegARIMA model, we pursue the well-known Box-Jenkins approach. As mentioned earlier, we set $s = 365$ given the daily frequency of the data whereas the polynomial lag orders have been determined by coefficient significance tests and minimization of the BIC criterion for the sake of parsimony. In particular, we end up with $p = 0, d = 1, q = 1, P = 0, D = 1, Q = 1$. The estimated effects are reported in Fig. 3 along with the 95 per cent confidence intervals. Regarding the road traffic of heavy commercial vehicles, there is a substantial decline of traffic on Carnival, by around 70 per cent, and a reduction of 10 per cent on the previous day. In the remaining days, the impact is not statistically different from zero. In what concerns Easter, the major effect is on Good Friday with a negative impact of 90 per cent and around 30 per cent in the subsequent three days. A very similar pattern is found for both electricity and natural gas consumption, but with smaller magnitudes. In the case of Carnival day, the negative impact on electricity and natural gas consumption is estimated to be around 15 per cent. For Good Friday, the effect is close to 15 per cent for electricity and near 25 per cent for natural gas consumption. For cargo and mail landed, the Carnival effect is close to 15 per cent and there is also a statistically significant effect on the following day of around 20 per cent. It is also found a significant impact throughout the Easter period, with the major effect being recorded on Good Friday (almost 60 per cent). Regarding card-based payments, there are no significant effects during the Carnival period. For Easter, it is observed a positive impact on the days immediately before Good Friday, a negative effect between Good Friday and Monday with the largest impact recorded on Sunday (close to 50 per cent).

4.4. The Daily Economic Indicator

After the series have been corrected for the above effects, year-on-year growth rates are computed. As usual, prior to factor model estimation, the series are outlier-adjusted following Stock and Watson (2005) and normalized to a common scale. To cope with missing values and unbalanced datasets in the context of a factor model one can draw on the EM algorithm as suggested by Stock and Watson (2002a).

Considering data until the end of 2019, the resulting weights attached to each variable are as follows: road traffic of heavy commercial vehicles (0.25), cargo and mail landed (0.15), natural gas (0.27), electricity consumption (0.18) and card-based payments (0.17). The road traffic of heavy commercial vehicles and natural gas consumption present a slightly higher weight whereas cargo and mail landed record the lowest weight. The variance explained by the latent variable amounts to 53 per cent. Through the recursive computation of the weights, one can also conclude that most of the weights have been relatively stable over the last years. Only the weight attached to electricity consumption seems to have decreased slightly whereas the weight for card-based payments has moved in the opposite direction.

Fig. 4 displays the estimated latent factor from January 1, 2014 to April 30, 2020. Such a daily economic indicator is certainly prone to some irregular movements, which reflect its high-frequency nature. In fact, if one takes the centered weekly average of the DEI, that is, the seven-day centered moving average, one rapidly obtains a much smoother indicator. This is, of course, even truer for the centered quar-

---

15 One should bear in mind that STL, as well as other well-established seasonal adjustment procedures such as X-12-ARIMA, TRAMO-SEATS and unobserved component models, despite their advantages, are not free of caveats when applied to daily data and may not be completely effective on removing intra-weekly seasonality (see, for example, Ladiray et al. (2018)).

16 For example, in the case of the weekly economic index developed by Lewis et al. (2020), the variance explained is around 54 per cent. To reinforce the use of the first factor as an overall economic indicator, we computed two criteria, the Eigenvalue Ratio (ER) and Growth Ratio (GR) estimators, recently developed by Ahn and Horenstein (2013). These determine the number of factors based on the empirical distribution of the eigenvalues and perform better than alternative criteria suggested in previous literature. We find that both ER and GR point to a single factor.

17 Note that this is the period for which all the constituent series are available. For the period which goes back to the beginning of 2010, the results should be read with caution as the information set is more limited.
Fig. 1. Smoothed periodogram.
Fig. 2. Cycle-subseries of the weekly seasonal component.

18 In the case of Lewis et al. (2020), they consider the year-on-year growth rate of weekly frequency data, whereas in Bundesbank (2020) 12-week growth rates are computed from the rolling 12-week averages for each high-frequency indicator.

In Fig. 5, we narrow the analysis to the latest period covering the lockdown of economic activity due to the coronavirus pandemic. In particular, we plot the DEI along with the year-on-year growth of quarterly GDP. One should note that, due to the estimation method, the scale of DEI is not interpretable. Therefore, to compare it more directly with the recent evolution of GDP growth, the DEI has been normalized so that
Fig. 3. Carnival and Easter effects on the daily series.
its quarterly counterpart has the same mean and standard deviation of GDP growth over the last years.

The DEI started to fall on March 12, when the Portuguese government declared the highest level of alert due to the COVID-19 pandemic. It sharpened the decline immediately after the declaration of the State of Emergency in the entire Portuguese territory on March 18, which led to the lockdown of several economic activities. In April, the figures point to an unprecedented decline of economic activity during the first half, with some very mild signs of recovery at the end of the month. Hence, this type of analysis emphasizes the usefulness of high-frequency indicators such as the DEI.

Finally, as in Lewis et al. (2020) we also assess the information content of the DEI to track current GDP developments. We start by regressing GDP growth on the quarterly average of the DEI, i.e.,

\[ y_t = \beta_0 + \beta_1 x_t + \sum_{s=1}^{S} \gamma_s y_{t-s} + u_t \]  

with the number of autoregressive lags selected according to BIC criterion following a general-to-specific approach. The estimation results for the period since the beginning of 2010 until the end of 2019 are presented in Table 1.\(^{19}\)

| Variables          | Quarterly data regression | Mixed frequency regression |
|--------------------|----------------------------|----------------------------|
| Low-frequency regressors |                             |                            |
| Constant           | 0.356***                   | 0.413***                   | 0.397***                   |
| \( y_{t-1} \)      | 0.550***                   | 0.400***                   | 0.410                       |
| Quarterly DEI      | 10.684***                  | 1.845                      |
| Parameters associated with DEI |
| \( \beta_1 \)     | 15.192***                  | (2.479)                    |
| \( \theta_1 \)    | 2.375*                     | (1.222)                    |
| \( \theta_2 \)    | 1.450*                     | (0.600)                    |
| R-squared          | 0.938                      | 0.940                      |
| Adjusted R-squared | 0.935                      | 0.940                      |

Notes: *, **, *** denote statistical significance at the 10, 5 and 1 per cent significance levels, respectively. The standard errors are reported within brackets.

The results reveal that the quarterly average of the DEI conveys significant information about current GDP growth, even after controlling for its past behaviour.\(^{20}\)

We also investigate the usefulness of the high-frequency availability of the DEI in a regression context. For that purpose, we consider a Mixed-Data Sampling (MIDAS) approach. The MIDAS approach, which has been put forward by Ghysels et al. (2004), allows series \( x_{t}^{m} \) sampled at lower-frequencies (such as quarterly GDP), to be regressed on a variable \( x_{t}^{1/m} \) sampled at frequency \( m \), say daily (see, for example, Andreou et al. (2013), for a comprehensive overview). Consider the following MIDAS model:

\[ y_t = \beta_0 + \beta_1 B(L^{1/m}; \theta) x_t^{1/m} + u_t \]  

where \( B(L^{1/m}; \theta) = \sum_{k=0}^{K} B(k; \theta) L^k \) and \( L^{1/m} \) is a lag operator such that \( L^{1/m} x_t^{(m)} = x_t^{(1/m)} \) and the lag coefficient in \( B(k; \theta) \) of the corresponding lag operator \( L^k \) is parameterised as a function of a small-dimensional vector of parameters \( \theta \). Thus, \( B(k; \theta) \) is a weighting scheme used for aggregation. As the number of lags of \( x_t^{(m)} \) can increase quite substantially, in particular when the frequency of the variables is very different, say quarterly and daily, the use of a convenient parametric function of \( B(L^{1/m}; \theta) \) allows to cope with parameter proliferation. In particular, we consider the Beta polynomial which is characterized by two parameters, \( \theta_1 \) and \( \theta_2 \), and the traditional Almon polynomial (of degree two). The model has also been augmented with lags of the dependent variable. From Table 1, one can conclude that both mixed-frequency models present a quite high explanatory power and slightly higher than in the previous case.

5. Concluding remarks

During rapidly evolving economic conditions, traditional statistics fail to provide timely signals on the state of the economy. The outbreak presented...
of the novel coronavirus at the end of 2019 and its lightning-fast spread across continents led economists to depart from traditional statistics and rely on high-frequency data instead. In fact, alternative data sources have become an increasingly important tool to track economic activity in the aftermath of the COVID-19 pandemic, due to the lack of relevant up-to-date data. This effort in collecting timely data has been particularly notable in central banks and other major institutions. Such a line of research can greatly benefit from the engagement of private firms, which typically own this type of data.

In this paper, we carried out an extensive data collection of non-traditional statistics from several sources at a daily frequency to assess which time series prove to be more relevant for tracking economic activity. However, when dealing with high-frequency data, there are relevant issues that have to be addressed, including marked seasonal patterns and calendar effects, which are expected to influence the dynamics of the time series. In this respect, we estimated the intra-weekly seasonal pattern by means of the STL method, while calendar effects have been addressed through RegARIMA models.

After a preliminary analysis and bearing in mind a widespread coverage of economic activity, we selected five daily series. These series cover electronic payments, transportation activity and energy consumption. The latent factor underlying the evolution of those series constitutes the proposed daily economic indicator. The analysis of the recent behaviour of DEI during the lockdown period in Portugal reveals a sudden and sharp drop of economic activity from mid-March 2020 onwards. In particular, when the Portuguese government declared the highest level of alert due to the COVID-19 pandemic on March 12, the DEI started its downward path. It declined further after the declaration of the State of Emergency in the entire Portuguese territory on March 18, which led to the lockdown of several economic activities. Moreover, the DEI points to an unprecedented decline of economic activity in April during the first half, with some very timid signs of recovery at the end of the month.

Lastly, we also assess the information content of the DEI to track current GDP developments. The results reveal that the quarterly average of the DEI conveys significant information about current GDP growth, even after controlling for its past behaviour. These findings are corroborated in a Mixed-Data Sampling regression approach, hence highlighting the usefulness of the DEI in tracking economic activity.

Acknowledgements

The opinions expressed in this article are those of the authors and do not necessarily coincide with those of Banco de Portugal or the Eurosystem. Any errors and omissions are the sole responsibility of the authors. The authors would like to thank all data providers for their prompt responses and outstanding efforts in collecting and sharing the daily data. The authors also thank the Editor and two anonymous reviewers for comments and suggestions.

References

Ahn, S.C., Horenstein, A.R., 2013. Eigenvalue ratio test for the number of factors. Econometrica 81 (3), 1203–1227.
Amengual, D., Watson, M., 2007. Consistent estimation of the number of dynamic factors in a large N and T panel. J. Bus. Econ. Stat. 25 (1), 91–96.
Andreou, E., Ghysels, E., Kontos, A., 2013. Should macroeconomic forecasters use daily financial data and how? J. Bus. Econ. Stat. 31 (2), 240–251.
Aprigliano, V., Ardizzi, G., Monteforte, L., 2019. Using payment system data to forecast economic activity. International Journal of Central Banking 15 (4), 55–80.
Arora, V., Lieskovsky, J., 2014. Electricity Use as an Indicator of U.S. Economic Activity. Working paper series. U.S. Energy Information Administration.
AISKIN, N., ZIMMERMANN, K., 2018. Nowcasting business cycles using toll data. J. Forecast. 34, 299–306.
Azevedo, J., Koopman, S., Rua, A., 2006. Tracking the business cycle of the Euro area: a multivariate model-based band-pass filter. J. Bus. Econ. Stat. 24 (3), 278–290.
Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica 70 (1), 191–221.
Barnett, W., Chauvet, M., Leiva-Leon, D., Su, L., 2016. Nowcasting Nominal GDP with the Credit-Card Augmented Divisia Monetary Aggregates. Working papers series in theoretical and applied economics no. 2016/05. University of Kansas.
Bergmeir, C., Hyndman, R.J., Benítez, J.M., 2016. Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation. Int. J. Forecast. 32 (2), 303–312.
Bundesbank, 2020. A weekly activity index for the German economy. Monthly report may 2020 2 (5), 68–72.
Carlsen, M., Storgaard, P.E., 2010. Dankort Payments as a TIMely Indicator of Retail Sales in Denmark. Technical report.
Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, L.J., 1990. STL: a seasonal-trend decomposition procedure based on loess. J. Off. Stat. 6 (1), 3–33.
Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. J. Am. Stat. Assoc. 74, 829–836.
Cleveland, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. J. Am. Stat. Assoc. 83, 596–610.
Cleveland, W.S., Terpenning, L.J., 1982. Graphical methods for seasonal adjustment. J. Am. Stat. Assoc. 77, 52–62.
Crum, W.L., 1927. Weekly fluctuations in outside bank debits. Rev. Econ. Stat. 9 (1), 30–36.
Duarte, C., Rodrigues, P.M., Rua, A., 2017. A mixed frequency approach to the forecasting of private consumption with ATM/POS data. Int. J. Forecast. 33 (1), 61–75.
Feng, Z., Schneider, M., 2009. A leading indicator of Austrian exports based on truck mileage. Monetary policy & the economy 44–52 q1/09.
Findley, D.F., Martin, D.E.K., 2006. Frequency domain analyses of SEATS and X-11-ARIMA seasonal adjustment filters for short and moderate-length time series. J. Off. Stat. 22 (1), 1–34.
Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C., Chen, B., 1998. New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. J. Bus. Econ. Stat. 16 (2), 127–152.
Fisher, I., 1923. A weekly index number of wholesale prices. J. Am. Stat. Assoc. 18 (143), 835–840.
Galbraith, J.W., Tkacq, G., 2018. Nowcasting with payment systems data. Int. J. Forecast. 34 (2), 366–376.
Ghysels, E., Santa-Clara, P., Valkanov, R., 2004. The MIDAS Touch: Mixed Data Sampling Regressions. Discussion paper, UNC and UCLA.
Hyndman, R.J., Athanasopoulos, G., 2014. Forecasting: Principles and Practice. O’Texts, Melbourne, Australia.
Ladiray, D., 2018. Calendar effects. (Chapter 5). In: Mazzi, G.L., Ladiray, D., Rieser, D. (Eds.), Handbook on Seasonal Adjustment. European Commission.
Ladiray, D., Mazzi, G.L., Palate, J., Proietti, T., 2018. Seasonal adjustment of daily and weekly data. (Chapter 29). In: Mazzi, G.L., Ladiray, D., Rieser, D. (Eds.). Handbook on Seasonal Adjustment. European Commission.
Lahiri, K., Yao, W., 2012. Should transportation output be included as part of the coincident indicators system? OECD Journal: J. Bus. Cycle Meas. Anal. (1), 1–24 2012.
Lahiri, K., Yao, W., Young, P., 2004. Transportation and the economy: linkages at business-cycle frequencies. 1864–Transport. Res. Rec. 103–111.
Lewis, D., Mertens, K., Stock, J., 2020. U.S. Economic Activity during the Early Weeks of the SARS-CoV-2 Outbreak. Staff report no. 920, april 2020, Federal Reserve Bank of New York.
Ollee, D., 2018. Seasonal Adjustment of Daily Time Series. Discussion paper no 41/2018. Deutsche Bundesbank.
Persons, W.M., 1919. Indices of business conditions. Rev. Econ. Stat. 1, 147–162.
Quenneville, B., Findley, D.F., 2012. The timing and magnitude relationships between month-to-month changes and year-to-year changes that make comparing them difficult. Taiwan Economic Forecast and Policy 44–52 q1/09.
Stock, J., Watson, M., 1989. New indexes of coincident and leading economic indicators. In: NBER Macroeconomics Annual 1989. vol. 4. National Bureau of Economic Research, pp. 351–409.
Stock, J., Watson, M., 1998. Diffusion Indexes. Working paper no. 6702. National Bureau of Economic Research.
Stock, J., Watson, M., 2002a. Macroeconomic forecasting using diffusion indices. J. Bus. Econ. Stat. 20, 147–162.
Stock, J., Watson, M., 2002b. Forecasting using principal components from a large number of predictors. J. Am. Stat. Assoc. 97, 1167–1179.
Stock, J., Watson, M., 2005. Implications of Dynamic Factor Models for VAR Analysis. Working paper no. 11467. National Bureau of Economic Research.