Postal Traffic in Portugal: Applying Time Series Modeling

Carla Machado and Filipa Silva

1 Introduction

The postal sector is adapting to the Internet technology (IT) revolution. Short message service (SMS) or email competes with correspondences (e-substitution). On the other hand, the purchase of physical products over the Internet (e-commerce) affects positively the volume of parcels delivered, a phenomenon expected to grow, spurred in part by digital single market goals.[1] Because parcel growth is an e-commerce phenomenon, fixed broadband penetration appears more correlated with postal traffic than gross domestic product (GDP) (EGIDE et al. 2015).

This paper aims to present models that fit and forecast postal traffic and parcels postal traffic. The two are estimated and forecasted separately because of their different tendencies over the period of the analysis. The estimation methods used are ARIMA, ARIMAX, decomposition, and multiple linear regressions. The dataset used for estimations is quarterly traffic of postal mail in Portugal (domestic and outgoing international), beginning in the first quarter of 2005 (1Q2005) and ending in the second quarter of 2018 (2Q2018), from postal operators, collected by ANACOM.

After this introduction, Sect. 2 provides a literature review. Section 3 presents the Portuguese data considered and its sources, as well as the analysis of the postal traffic in Portugal, for both correspondences and parcels, and analysis of their drivers. Section 4 explores the methodological framework and compares the main results of

[1] https://ec.europa.eu/digital-single-market/en/new-eu-rules-e-commerce

ANACOM. The opinions expressed are those of the author and do not necessarily reflect the views of ANACOM.

C. Machado (✉) · F. Silva
ANACOM, Lisbon, Portugal
e-mail: carla.machado@anacom.pt

P. L. Parcu et al. (eds.), *The Changing Postal Environment*, Topics in Regulatory Economics and Policy, https://doi.org/10.1007/978-3-030-34532-7_15

the estimated models under different approaches. The best model is then applied to forecast the short-term evolution of correspondence and parcels traffic in Portugal. Conclusions are discussed in Sect. 5.

2 Literature Review

The postal sector is a pillar of socioeconomic development, and historically, according to UPU (2014) until the year 2000, letter-post volume per capita grew at a similar pace as GDP per capita in industrialized countries. However, after that year, there was a decoupling of the two trends: GDP per capita continued to grow, while the number of letters per capita declined noticeably.

In Portugal, EGIDE et al. (2015) mention that GDP lost its main power to explain postal traffic since 2005, although ANACOM (2011) refers that the main senders of postal traffic (big enterprises) remain very sensitive to the price of the postal services. With the development of IT, there was a decrease of the traditional letter mail volumes (because of e-substitution and changes in interpersonal communication) and an increase of parcels volumes, especially cross-borders (due to the adoption of e-commerce) – UPU (2019), Wik (2019), Copenhagen Economics (2018), and EGIDE et al. (2015), among others, all document this result.

EGIDE et al. (2015) explained that postal traffic in European countries was decreasing because of five main forces: economic growth,[2] population, digitalization, costs for the clients,[3] and legal provisions,[4] although digitalization, in general, is the main driver of the decline.

On the other hand, the authors mention that parcels traffic is increasing in every European country. UPU (2018) mentions economic growth, trade, and connectivity as the factors that mostly affect postal traffic – GDP and exports create demand for postal services, while connectivity (measured as the rate of Internet penetration) and digitization encourage the increase of e-commerce. However, Wik (2019) warns that available data on cross-border parcels are limited and underestimate the actual volumes, namely, because they do not consider small packets delivered through the letter-post stream.

Wik (2019) and ERGP (2019) also alert that there might be a problem in the definition of what is a postal service or a postal operator (for instance, what are the boundaries between postal and transport services), in addition to the fact that different countries have different definitions of postal services and operators.

[2] Namely, GDP.

[3] Increases of the costs have impact on the traffic volumes because these are the drivers for digitalization.

[4] This last, especially where e-government policies have been adopted, is the case in Portugal.

ERGP (2019) includes the environment impact of postal delivery (ecological footprint) and labor conditions in the drivers of the postal traffic (e.g., temporary employment, subcontracted workers, and the need for more flexible deliveries, at night, extra work during peak periods). In addition, they refer that the postal market is becoming more receiver-oriented, than sender-oriented, due to the increase of the power of postal receivers to influence the quality characteristics of the deliveries (for instance, requiring efficient track and trace).

3 Postal Traffic in Portugal

3.1 Data Source and Samples

In this article, the demand for postal services is described as outgoing international volumes of correspondences, and parcels, from the quarterly statistical indicators of ANACOM, the Portuguese postal regulator, collected from the postal service providers.[5]

The time span of the data is from 1Q2005 to 2Q2018. In 2012, the full competition regime was adopted on national territory, as well as of international services to or from national territory, due to the legal regime (Law no. 17/2012, of 26 April[6]) that governs the provision of postal services. That regime also changed some of the postal definitions, namely, new types of postal items[7] related to postal services (prior to that, information about editorial mail and direct mail was not clearly separated), and the maximum weight limit for postal parcels (previously, 20 kg) was removed from the definition, affecting the data collected by ANACOM from postal operators. As a result, from 4Q2012 onward, ANACOM adopted the definitions of postal services indicated in the new law and started collecting information with these new criteria.

With the new law, certain providers considered that (at least) some of their services were not anymore within the definition of postal services; therefore, they stopped reporting statistical information since 2014, according to ANACOM

[5] Statistical indicators include data on traffic, revenues, postal network, and human and material resources. Data is subject to changes with revised or updated new quarterly information. More information is available at https://www.anacom.pt/render.jsp?categoryId=337756.

[6] Transposes the national legal system Directive 2008/6/EC of 20 February 2008. More information is available at https://www.anacom.pt/render.jsp?contentId=1127982.

[7] According to the definition, a postal item is an item addressed in the final form, which observes the physical and technical specifications that allows it to be sorted by a postal network, as well as delivered at the address indicated on the object itself or on its wrapping, namely, (a) correspondence, which consists of a communication in written form on any kind of physical medium, including direct mail; (b) editorial mail as books, catalogues, newspapers, and other periodicals; and (c) postal parcel, which is a package containing merchandise or objects with or without commercial value.

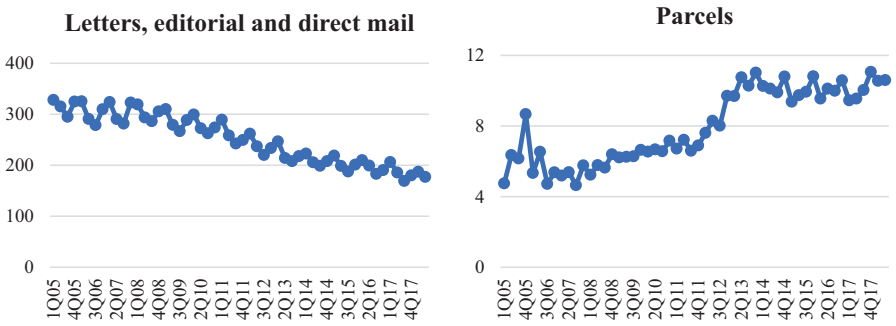

Fig. 1 Evolution of the postal traffic by item (domestic and outgoing international), Portugal. Unit: Millions of objects. (Source: Authors, with data from ANACOM)

(2018). Wik (2019) and ERGP (2019) also note a similar development for other countries.

3.2 The Evolution of Postal Traffic

Portugal observed a drop in postal traffic over the last few years, with a decrease of the total traffic of 4.4% per year, in the last 10 years, mainly due to the decrease of the letters sent. By type of item, considering data from the 1Q2005 to the 2Q2018, there is a negative trend in letters, editorial mail, and direct mail and a positive trend in parcels.

Figure 1 presents the postal evolution by type of item, gathered according to these common trends.

The volume of correspondences (letters, editorial mail, and direct mail) has a negative trend at least since the 1Q2005, with a higher decrease between the end of 2007 and the end of 2011 and an eased negative trend from then on.[8] The volume of parcels had a positive but slow increasing trend from 2008 to the end of 2011, in 2012 it took a leap, and from 2013 on, it stabilized around the 10 million objects.[9]

In the 2Q2018, the volume of parcels represented only 6% of the total postal traffic (domestic and outgoing international,), while around 80% were letters, 7% editorial mail, and 8% direct mail, a distribution that is slowly changing since 2012 – in the 2Q2012 parcels represented 3% of the total postal traffic in Portugal.

[8] The structural break analysis identified two breaks (4Q2007 and 4Q2011) in the time series of correspondence traffic.

[9] The structural break analysis identified one break in the time series of correspondence traffic (4Q2012).

3.3 Postal Traffic Determinants

This report considers data to capture the effect of digitalization (fixed broadband penetration[10]) and e-commerce (exports) on postal traffic, together with data to capture the economic evolution in Portugal (GDP, consumer price index, exports, unemployment rate, economic sentiment indicator, and consumer confidence indicator).[11]

3.3.1 Correspondence (Letters, Editorial Mail, and Direct Mail) Determinants

The SIMPLEX program, adopted by the Portuguese Government in 2006 to simplify administrative and legislative procedures using, among others, electronic channels (e-government), explains partially the postal traffic decline between 2007 and 2011. Among other government initiatives were the promotion of the access to Internet and to computers at a small price to young and adult people (e-initiatives program, launched in 2007)[12] and the availability of electronic forms to citizens and companies and the dematerialization of some acts of registration, making it possible to request them without physical relocation to conservatoires. Between 2007 and 2018, according to Eurostat data, the Internet usage in Portugal increased notoriously (from 40% to 75%), especially in people with ages between 25 and 44.

After the initial phase of the SIMPLEX program, the decrease of postal traffic started to ease.

The financial crisis between 2010 and 2014 in Portugal may have helped to delay the process of dematerialization of the invoices of the Portuguese companies and, as a consequence, to slow down the correspondence decrease.

Considering the explanatory variables initially mentioned, from the 1Q2005 to the 2Q2018, the variable with the highest correlation with the volume of correspondences is fixed broadband penetration (−94%[13]). This correlation is strongly negative, acting as a proxy for digitalization (e-substitution, e-government, and e-invoice). The second main variable related with the volume of correspondences is exports (−92%), also with a negative relation. If one considers the evolution in time of the correlations, instead of just the correlation in one specific period, Fig. 2 shows that the correlation between postal correspondence and fixed broadband penetration growth in time, during the

[10] Fixed broadband accesses per 100 inhabitants. Data from ANACOM, Portugal.

[11] Data from Instituto Nacional de Estatística, Portugal (www.ine.pt).

[12] E-initiatives (e-opportunities, e-school, and e-teachers) was launched by the Portuguese Government, in September 15, 2007, in order to promote the information and communication technologies, which permitted to students (young and adult) and teachers to receive a laptop and access to the Internet at a symbolic price, with training classes to explain how to work with the equipment. In 2 years, 852 thousands individuals took part in this program, 1/3 were less than 16 years old, 1/3 between 17 and 35 and the other 1/3 more than 35 years (15% with more than 45 years).

[13] The Pearson correlation coefficient was considered in correlation analysis.

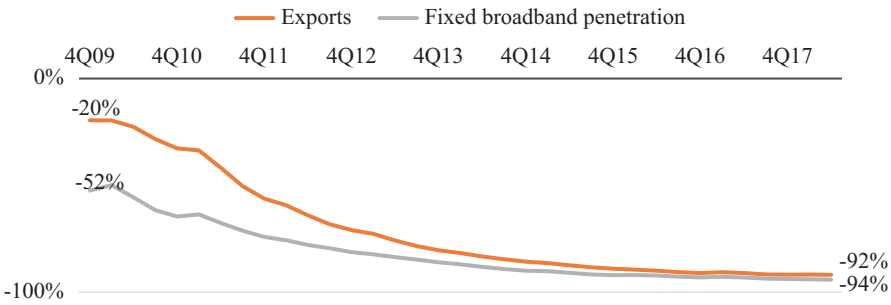

Fig. 2 Correlation between correspondence traffic (letters, editorial mail, and direct mail) exports and fixed broadband penetration, between the 1Q2005 and the date mentioned in horizontal line, Portugal. (Source: Authors, with data from ANACOM)

1Q2005 and until the 4Q2009, was −52%, but if we extend the data until the 2Q2018, the negative correlation was −94%. A similar trend holds for the correlation between postal correspondence and exports.

3.3.2 Parcels Determinants

The growth of parcels traffic in 2012, compared with previous years, might be explained by the new legal regime that governs the provision of postal services under a full competition regime, on national territory, as well as of international services to or from national territory,[14] layed down in 2012.

However, according to ANACOM (2018), from 2014 on, certain providers started considering that (at least) some of their services were not falling anymore in the definition of postal service; therefore, they stopped reporting statistical information or providing statistical information to ANACOM using a new reclassification of object type. ANACOM (2018) mentions that this might be a move to escape the legal regime, and its fees, applicable to all postal service providers. As for the parcels traffic evolution, this lack of data reporting may help explain why from 2014 parcels traffic has stabilized around 10 million objects.

Using the explanatory variables already mentioned for correspondences, fixed broadband penetration results again as the variable with the higher relation with parcels' volumes, exhibiting a strong positive correlation (+86%) between 1Q2005 and 2Q2018, probably due to the effect of e-commerce. Exports is the second important variable, with a high correlation with parcels (+83%), but slightly lower than with correspondences.

[14] Law no. 17/2012, of 26 April, that transposes the national legal system Directive 2008/6/EC of 20 February 2008. More information is available at https://www.anacom.pt/render.jsp?contentId=1127982.

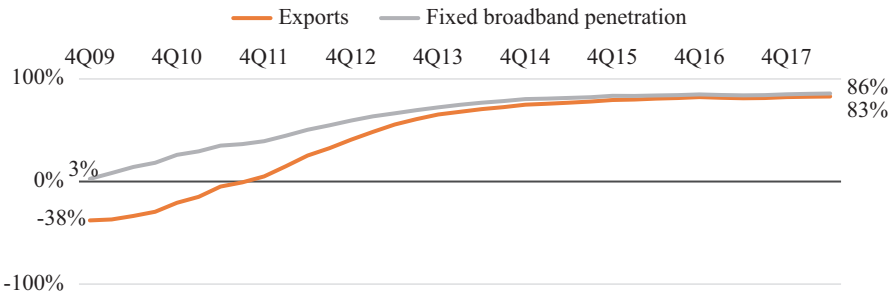

Fig. 3 Correlation between parcels traffic and Internet penetration, between the 1Q05 and the date mentioned in horizontal line, Portugal. (Source: Authors, with data from ANACOM)

The evolution of the correlations shows that while until about the end of 2011, exports were negatively related with the volume of parcels, recently they are highly positively correlated (+83%), potentially because, on the one hand, it creates demand for postal traffic but also associates with the growth of the e-commerce conditions in the last years (Fig. 3).

4 Time Series Modeling

4.1 Methodological Framework

To model the time series of postal traffic, we consider three methodological approaches: (1) ARIMA (autoregressive integrated moving average) and ARIMAX (extend ARIMA models through the inclusion of exogenous variables) models, (2) decomposition models, and (3) multiple linear regressions.

ARIMA models are based on the Box-Jenkins approach (Box and Jenkins 1970). They include two nonseasonal components in time: the autoregressive (AR) component, related with the effect of the time series in itself, and the moving average (MA) component, related with the effect of the error in the time series. If the time series has seasonality, this becomes a SARIMA model, and the seasonal AR and MA components are considered.

The extended ARIMA models through the inclusion of exogenous variables (*ARIMAX models*) were also analyzed and discussed in Sect. 3.3.

Decomposition models are based on the splitting of the time series in three different components: (a) tendency, (b) seasonality, and (c) other oscillating movements. These models include exponential smoothing models, in particular, Holt (1957) developed the approach with the tendency and Winters (1960) developed the approach with seasonality. These are now known as Holt-Winters (HW) models. They can be additive or multiplicative, depending on the relation between the three previous components of the model. Forecasts derived from exponential smoothing are weighted averages of

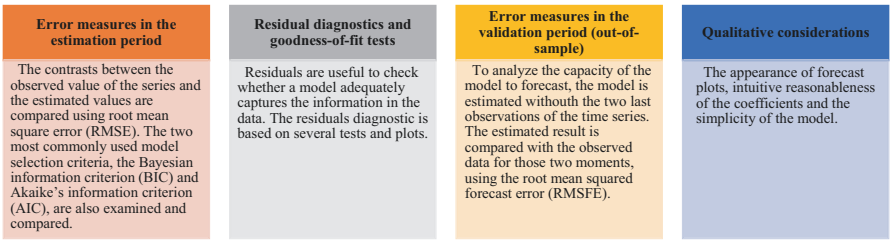

Fig. 4 Criteria to compare the results after estimation. (Source: Authors)

past observations, with the weights decaying exponentially as the observations get older, so the more recent observations have a higher weight in the predictions.

The *multiple linear regression (MLR) method* also expresses the long-term trend in which a common linear regression model is established between the incidence studied and time t. These models may also include dummies for seasonality and for structural breaks (to identify when a time series abruptly changes at a point in time).

The MLR models enable to include exogenous variables in the estimation, in our case those discussed in Sect. 3.3.

The objective of the estimations is to determine which of the candidate models best explain the postal traffic, considering the available data, which involves minimizing the loss of information. To compare the results after the estimation, different criteria are used (Fig. 4).

4.2 Estimation, Goodness-of-Fit and Forecast

4.2.1 Correspondence (Letters, Editorial, and Direct Mail)

The traffic of the postal correspondence, available quarterly for the 1Q2005 to the 2Q2018 (54 observations), is a nonstationary[15] time series, with negative trend and seasonal effects. The log transformation stabilizes its variance.

According to the ACF and PACF functions and to the residuals diagnostic, it was selected an ARIMA model, with seasonality (seasonal order is 4, due to quarterly data) denoted as $SARIMA(p, d, q)(P, D, Q)_s$.[16] This is a seasonal exponential smoothing

[15] Verified by the Augmented Dickey-Fuller test and Phillips-Perron test for unit root.

[16] An ARIMA with seasonality is denoted as SARIMA (p, d, q) $(P, D, Q)_s$ and is given by $\prod(B^s)\Phi(B)\Delta^D{}_s\Delta^d Y_t = \Theta(B^s)\theta(B)\varepsilon_t$. The nonseasonal AR and MA components are represented by polynomials $\Phi(B)$ and $\theta(B)$ of orders p and q, respectively, and the seasonal AR and MA components by $\prod(B^s)$ and $\Theta(B^s)$ of orders P and Q. Nonseasonal and seasonal difference components by $\Delta^d=(1-B)^d$ and $\Delta^D{}_s=(1-B^S)^D$, where, p, d, and q are the orders of nonseasonal AR, differencing, and MA, respectively; P, D, and Q are the orders of seasonal AR, differencing, and MA, respectively, and s represents seasonal order ($s = 4$ for quarterly data).

Table 1 Estimated results Box-Jenkins approach (compared SARIMA and ARIMAX model) for postal correspondence traffic

Variable	(1.1) SARIMA
ln_Correspondence	
_cons	−0,001
ARMA ar	
L1.	–0,408∗
Sigma	
_cons	0,025∗∗∗
ARMA4 ma	
L1.	−0,565∗∗∗
Statistics	
RMSE	25,5
AIC	−213,7
BIC	−206,2
Skewness	−0,182
Pr(skewness)	0,564
Kurtosis	3,178
Pr(kurtosis)	0,485

Source: The authors
Legend: ∗ $p < 0.05$; ∗∗ $p < 0.01$; ∗∗∗ $p < 0.001$

model, conceptually similar to the Winters model: (1.1) *SARIMA*(p, d, q)(P, D, Q)$_s$ = *SARIMA*(0, 1, 1)(0, 1, 1)$_4$.

ARIMAX models[17] were developed using exogenous variables that were highly correlated with the time series for correspondence traffic, as presented in Sect. 3.3. If one considered the fixed broadband penetration, the residuals were not normal distributed, while in the case of exports, the model had certain misspecification problems (Table 1).

The results of the exponential smoothing models with seasonal effects (additive and multiplicative) show that, although the additive and multiplicative methods have similar goodness-of-fit, the multiplicative method (1.2) is slightly better (Table 2).

The last models to comparison are the *multiple linear regression method*. Two different models were considered to the comparison. The first multiple linear regression method[18] considers the linear trend (t), seasonal dummies (Q_1, Q_2, and Q_3),[19] a dummy for the structural break of 4Q2007 (B_{4Q2007}), and a dummy for the structural break of 4Q2011 ($t*B_{4Q2011}$). The time breaks were pre-identified with a structural break analysis. The first break is related with the beginning of the digitalization, and the last break may be explained by the end of the first stage of digitalization and with the financial crisis, as mentioned in Sect. 3.

[17] General model is given by $\prod(B^S)\Phi(B)\Delta^D{}_S\Delta^d Y_t = \Psi(B)X_t + \Theta(B^S)\theta(B)\varepsilon_t$, where Xt are the exogenous variables.

[18] General model is given by $Yt = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_p X_{pt}$, where Xpt ($p = 1,\ldots,r$) are the independent variables, β_t are the parameters, and βr is the number of regressors.

[19] Q_4 was not included, to avoid multicollinearity problems.

Table 2 Exponential smoothing model parameters and model fit statistics for postal correspondence traffic

	Holt-Winters' additive (HWA) method		(1.2) Holt-Winters' multiplicative (HWM) method	
	Coefficients	SE	Coefficients	SE
Alpha (level – l)	0,548	*(0,117)*	0,553	*(0,117)*
Beta (trend – b)	0,001	*(0,016)*	0,002	*(0,016)*
Gamma (season – s)	0,811	*(0,264)*	0,807	*(0,266)*
R-squared	0,986		0,987	
RMSE	23,70		23,68	
MAPE	145,60		145,43	
MAE	18,08		18,06	
Normalized BIC	−7,263		−7,265	
Ljung-Box Q *Sig.*	*0,801*		*0,805*	

Source: The authors

$$\begin{aligned}\ln\_\text{corresp}_t &= 12,75 - 0,009t + 0,041Q_{1t} - 0,041Q_{2t} - 0,085Q_{3t} \\ &\quad - 0,036B4Q2007t - 0,003t * B_{4Q2011t}\end{aligned} \tag{1.3}$$

(Adj R-squared = 0,986)

The second model was estimated including exogenous variables (fixed broadband penetration) to capture the trend of the time series, seasonal dummies, and a dummy for the structural break that happened in 4Q2011. The regressions diagnostic does not identify heteroscedasticity, misspecification, or multicollinearity problems, and the residuals follow a normal distribution.[20]

$$\begin{aligned}\ln\_\text{corresp}_t &= 13,19 + 0,041Q_{1t} - 0,043Q_{2t} - 0,087Q_{3t} - 0,348\ln\_\text{penetBB}_t \\ &\quad + 0,931B_{4Q2011t} - 0,213\ \ln\_\text{penetBB}_t * \text{B}_{4Q2011t}\end{aligned} \tag{1.4}$$

(Adj R-squared = 0,979)

The analysis of different indicators related to the goodness-of-fit and error measures (in the estimation period and validation period) identifies model (1.2) – Holt-Winters' multiplicative method – as the best model to fit and forecast correspondence traffic in Portugal, in the period from the 1Q2005 to the 2Q2018 (Table 3).

Figure 5 shows the forecast of the correspondence traffic for the third and fourth quarters of 2018, as well as the respectively confidence limits. According to the results, postal correspondence traffic is better explained by its own behavior than by the explanatory variables considered, eventual because some important variables that might explain postal correspondence are not available or may not be measured.

[20] Breusch-Pagan test and White's test, Ramsey RESET test, and Shapiro-Wilk W test for normal data

Table 3 Error measures in the estimation and validation period, for postal correspondence traffic

Models	(1.1) SARIMA $(0,1,1)\ (0,1,1)_4$	(1.2) Holt-Winters' multiplicative	(1.3) MLR	(1.4) MLR + lnPenetrationBB
RMSE	25,5	**23,0**	23,3	26,7
AIC	−213,7	(∗)	−243,3	−222,6
BIC	−206,1	(∗)	−229,4	−208,7
Predictive capacity (RMSFE)∗∗	24,5	**20,3**	21,0	25,3

Source: The authors
Notes: (∗) not comparable indicators; (∗∗) out-of-sample analysis – comparison between predicted values to real values of the time series. The set forecasts to start from the first quarter of 2018 to the end of the second quarter of 2018

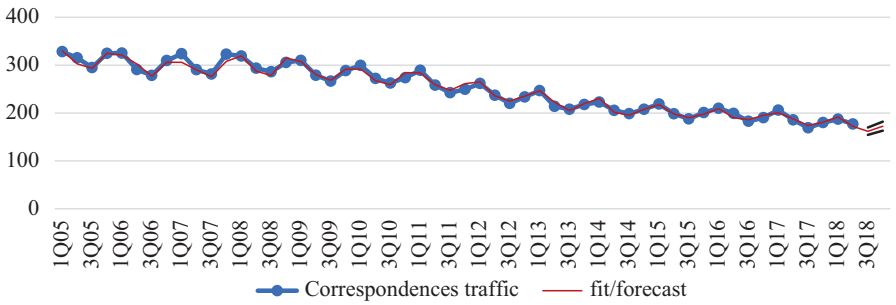

Fig. 5 Time series fit and forecast for postal correspondence traffic. Unit: Millions of objects. Note: Lower confidence limits (LCL) and upper confidence limits (UCL) for 95% confidence interval. (Source: Authors, with data from ANACOM)

4.2.2 Parcels

The time series of the parcels traffic starts in the 1Q2007 and goes to the 2Q2018 (46 observations). Although we have data since 2005, the first observations cannot be considered because of a strong variation in the results, explained by the break tests applied to the time series. The data is a nonstationary[21] time series, with a positive trend and a significant break that begins in the 1Q2012. This break is a consequence of digitalization and e-commerce, as well as the beginning of the new postal law and the lack of report of statistical information by some providers since 2014, as previously mentioned. The log transformation was applied to stabilize the variance of the time series and the first differences to obtain stationarity.

An *ARIMA model* was selected after residuals, ACF and PACF diagnostic,[22] and *ARIMAX models* were developed using the exogenous variables highly correlated

[21] Verified by the Augmented Dickey-Fuller test and Phillips-Perron test for unit root.

[22] General model is denoted as ARIMA (p, d, q) and given by $\Phi(B)\Delta^d y_t=\theta(B)\varepsilon_t$. The nonseasonal AR and MA components are represented by polynomials B and B of orders p and q, respectively, and d-difference components by $\Delta^d=(1\text{-}B)^d$; p, d, and q are the orders of nonseasonal AR, differentiation, and MA, respectively.

Table 4 Estimated results Box-Jenkins approach (compared ARIMA and ARIMAX model) for postal parcels traffic

Variable	(2.1) ARIMA	(2.2) ARIMAX | PenetrationBB
ln_Parcels		
ln_PenetraBB		
D1.		−2218**
_cons	0,016	0,058
ARMA		
ar		
L1.	−0,363*	
ARMA4		
ar		
L1.	0,514**	0,965***
ma		
L1.		−0,691***
Sigma		
_cons	0,059***	0,053***
Statistics		
rmse	61,9	57,7
aic	−116,9	−121,8
bic	−109,7	−112,8
Skewness	0,24	0,029
Pr(skewness)	0,46	0,928
Kurtosis	3,63	2,685
Pr(kurtosis)	0,21	0,901

Source: The authors
Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

with the parcels traffic presented in Sect. 3.3. The estimated results of the best models are (2.1) *ARIMA*(4, 1, 0) and (2.2) *ARIMAX* (0,1,0) $(1,0,1)_4$ with fixed broadband penetration as exogenous variable.[23] The *ARIMAX model* with export as exogenous variable has misspecification problems (Table 4).

The results of the exponential smoothing model with seasonal effect show that the additive method (2.3) has the best goodness-of-fit, compared with the multiplicative method (Table 5).

Two different *multiple regression models* were considered for comparison.

The first considers a linear trend (t), seasonal dummies (Q_2, Q_3, Q_4),[24] a dummy for the structural break in the 4Q2012 (B4Q2012), and a dummy for the 1Q2014 related to missing report of statistical information (D1Q2014).

[23] General model is given by $\Phi(B)\Delta^d y_t = \Psi(B)X_t + \theta(B)\varepsilon_t$ where *Xt* are the exogenous variables.

[24] Q1 is excluded to avoid multicollinearity problems.

Table 5 Exponential smoothing model parameters and model fit statistics for postal parcels traffic

	(2.3) Holt-Winters' additive (HWA) method		Holt-Winters' multiplicative (HWM) method	
	Coefficients	SE	Coefficients	SE
Alpha (level – l)	0,899	*(0,156)*	0,700	*(0,136)*
Beta (trend – b)	0,000	*(0,055)*	0,000	*(0,057)*
Gamma (season – s)	0,000	*(0,375)*	1000	*(0,513)*
R-squared	0,959		0,955	
RMSE	0,053		0,056	
MAPE	0,438		0,515	
MAE	0,039		0,046	
Normalized BIC	−5,609		−5,504	
Ljung-Box *Q Sig.*	0,397		0,755	

Source: The authors

$$\begin{aligned}\text{In\_parcels}_t &= 8{,}5 - 0{,}002t * \left(1 - D_{1Q2014t}\right) + 0{,}038\text{Q}_{2t} + 0{,}065Q_{4t} \\ &\quad + 0{,}198B_{4Q2012t} + 0{,}500D_{1Q2014t} \end{aligned} \tag{2.4}$$

(Adj. R-squared = 0,971)

The second model was estimated including fixed broadband penetration. Since this variable has a positive trend, in this model the time trend was not included to avoid multicollinearity.

$$\begin{aligned}\text{In\_parcels}_t &= 6{,}166 + 0{,}891\ln\_\text{penetBB} * \left(1 - D_{1Q2014t}\right) \\ &\quad + 0{,}052Q_{4t} + 0{,}236B4Q2012t + 2{,}81D_{1Q2014t} \end{aligned} \tag{2.5}$$

(Adj. R-squared = 0,9614)

The regressions diagnostic does not identify any problems related to heteroscedasticity, misspecification, multicollinearity, and normality of residuals.[25] The addition of exports as exogenous variable does not improve the model specification.

The results of the error measures return ARIMA as the best model if one considers the predictive capacity measure, while MRL model is the best model if one considers least RMSE measure (Table 6). Because parcels traffic presents an increase of its volume in the most recent data (1Q2018 and 2Q2018, when it had a constant trend since 2014), this affects negatively the estimation and forecast capabilities of the models. Since, at this time, we cannot determine if that growth is temporary or permanent, we assumed MRL model as the best model to fit parcels traffic in Portugal, in the period from the 1Q2007 to the 2Q2018.

[25] Breusch-Pagan test and White's test, Ramsey RESET test, and Shapiro-Wilk W test for normal data.

Table 6 Error measures in the estimation and validation period for postal parcels traffic

Models	(2.1) ARIMA	(2.2) ARIMAX (+lnPenetrationBB)	(2.3) Holt-Winters' additive	(2.4) MLR	(2.5) MLR (+lnPenetrationBB)
RMSE	61,69	57,70	50,27	**41,32**	48,06
AIC	−116,9	−121,8	(∗)	−150,61	−138,71
BIC	−109,7	−112,8	(∗)	−139,64	−129,57
Predictive capacity (RMSFE)∗∗	**20,55**	27,38	59,24	30,75	36,59

Source: The authors
Notes: (∗) not comparable indicators; (∗∗) out-of-sample analysis – comparison between predicted values to real values of the time series. The set forecasts to start from the first quarter of 2018 to the end of the second quarter of 2018

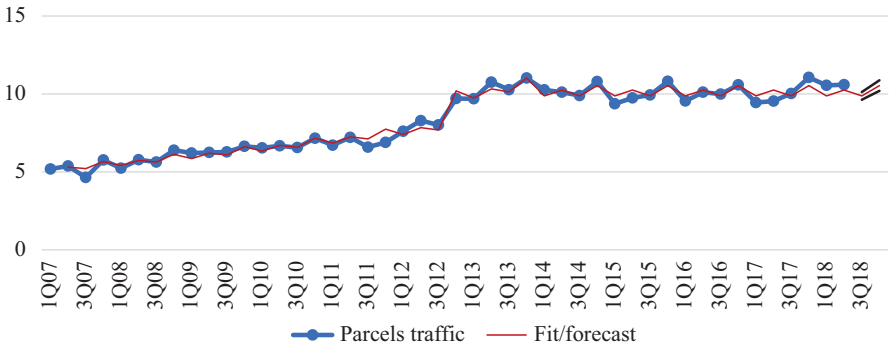

Fig. 6 Time series fit and forecast: Parcels traffic. Unit: Millions of objects. Note: Lower confidence limits (LCL) and upper confidence limits (UCL) for 95% confidence interval. (Source: the authors)

Figure 6 shows the forecast of the parcels traffic for the 3Q2018 and 4Q2018, considering MRL model (2.4).

5 Conclusions

The dynamics of the Internet technology sector has two different effects on the postal sector. On the one hand, digitalization, which includes e-government, e-substitution, and e-invoice, is a negative impact on the correspondence's postal traffic, and on the other hand, e-commerce helped the parcels traffic to grow, due to the delivery of physical product bought through the Internet. As a result, in Portugal, fixed broadband penetration and exports are highly correlated with the postal traffic, both in postal correspondences traffic and in parcels traffic (domestic and outgoing international), although with opposite effects.

At the time of the analysis, it was put in place by the law (Law no. 17/2012, of 26 April) that governs the provision of postal services under a full competition regime, on national territory, as well as of international services to or from national territory. Therefore, some definitions of some postal services changed. In particular, the maximum weight limit for postal parcels (previously, 20 kg) was removed.

In this study, three methodological approaches were compared: ARIMA and ARIMAX models, decomposition models, and multiple linear regressions.

The results show that, in the short run, postal traffic time series are better explained by their own behavior then when considering explanatory variables. Maybe that is because, apart from digitalization and e-commerce, other factors are affecting postal correspondences traffic, which are not being measured in this analysis (unavailability of the information, namely, variables that are more relevant than the trend itself, or inexistence of quarterly data).

With the digital single market goals, the decrease of the correspondence traffic is not expected to slow down soon, in Portugal. Regarding parcels traffic (domestic and outgoing international), the forecast models predict a stabilization of the series, even though the uncertainty related with the last two quarters of the time series, which might be explained by e-commerce but can also be the result of the use of different postal definition by some postal operators.

Some big questions remain not precisely answered, namely, what is the exact impact of digitalization on correspondences traffic and the direct effect of e-commerce on parcels traffic. For correspondence, data on big postal senders is necessary to understand the future evolution of the traffic. Concerning parcels, data from postal operators, disaggregated by origin (which traffic comes from e-commerce and which comes from other sources), would also be helpful to predict the future of parcels delivery.

Future works may concern the estimation of models with deseasonalized data, the consideration of relative values of the time series instead of absolute values, and the use of data regarding revenues, since they directly affect postal traffic. Considering that the causes affecting correspondences and parcels traffic are related, the estimation of a simultaneous equation models for correspondences and parcels traffic remains another promising possibility for future work.

References

ANACOM. (2011). *Challenges facing the postal sector as it moves to full liberalization*. https://www.anacom.pt/render.jsp?contentId=1092318

ANACOM. (2015). Study on the development of the postal sector. https://www.anacom.pt/render.jsp?contentId=1361251

ANACOM. (2018). *The communications sector 2017*. https://www.anacom.pt/render.jsp?categoryId=349440

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.

Box, G., Jenkins, G., & Reinsel, G. (1994). Time series analysis: Forecasting and control (3rd ed.). Hoboken: Prentice-Hall.

Brockwell, P., & Davis, R. (1991). Time series: Theory and methods (2nd ed.). New York: Springer.
Cazals, C., & Florens, J.-P. (2002). Econometrics of mail demand. In M. A. Crew & P. R. Kleindorfer (Eds.), Postal and delivery services (pp. 119–140). Boston: Kluwer Academic Publishers.
Cazals, C., Florens, J. P., Veruete-Mckay, L., Rodriguez, F., & Soteri, S. (2010). UK letter mail demand: A content based time series analysis using overlapping market survey statistical techniques. Available at https://www.tse-fr.eu/publications/uk-letter-mail-demand-content-based-time-series-analysis-using-overlapping-market-survey-statistical
Copenhagen Economics. (2018). *Main developments in the postal sector (2013–2016)*. Luxembourg: Publications Office of the European Union.
CTT. (2018). E-commerce report 2018.
EGIDE, IMR & ANACOM. (2015). *Study on the development of the postal sector*. https://www.anacom.pt/render.jsp?contentId=1361251
Enders, W. (1995). Applied econometric time series. New York: Wiley.
ERGP. (2019). *Developments in the postal sector and implications for regulation*. Available at https://ec.europa.eu/docsroom/documents/34881/attachments/1/translations/en/renditions/native
Holt, C. E. (1957). *Forecasting seasonals and trends by exponentially weighted averages (O.N.R. Memorandum No. 52)*. Pittsburgh: Carnegie Institute of Technology.
Jarosik, M., Nankervis, J., Pope, J., Soteri, S., & Veruete-McKay, L. (2002). Reforming the postal sector in the face of electronic competition. In M. A. Crew & P. R. Kleindorfer (Eds.), Postal and delivery services: delivering on competition (pp. 194–210). Boston: Springer.
Murteira, B., Muller, D., & Turkman, K. (1993). Análise de sucessões cronológicas. Lisboa: Mc Graw Hill.
Nikali, H., Mattila, J., Rintanen, I., & Huunhtanen, V. (2017). Demand and regulation for e-commerce in goods. In M. A. Crew & T. J. Brennan (Eds.), The future of the postal sector in a digital world (pp. 247–260). Cham: Springer Verlag.
StataCorp LLC. (2017). Stata time-series reference manual: Release 15 . Texas: College Station.
UPU. (2014). *Development strategies for the postal sector: an economic perspective*. Delft: Eburon Academic Publishers.
UPU. (2018). *Postal economic outlook 2018*. Universal postal union postal statistics. Available online in http://www.upu.int/en/resources/postal-statistics/latest-results.html
UPU. (2019). *Postal economic outlook 2019*. Universal postal union postal statistics. Available online in http://www.upu.int/en/resources/postal-statistics/latest-results.html
Wik. (2006). Main developments in the postal sector (2004–2006). New York: Springer.
Wik. (2009). The evolution of the European postal market since 1997. Available at https://ec.europa.eu/growth/sectors/postal-services/studies_en
Wik. (2019). *Development of cross-border e-commerce through parcel delivery*. Luxembourg: Publications Office of the European Union.
Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science, 6*, 324–342.
Yaffee, R., & McGee, R. (2000). Time series analysis and forecasting, with applications of SAS and SPSS . San Diego: Academic Press.