LONG-TERM RELATIONSHIP ESTIMATION AND COUPLING/DECOUPLING ANALYSIS BETWEEN MOTORWAY TRAFFIC AND GROSS VALUE ADDED. SPECIFICATION OF AN ARDL COINTEGRATION APPROACH AND APPLICATION TO THE ITALIAN CASE STUDY

Andrea POMPIGNA¹, Raffaele MAURO²
¹,² DICAM, University of Trento, Trento, Italy

Abstract:
As transportation is an activity derived from spatial complementarities between a certain supply at an origin and a certain demand at a destination, according to a general axiom it seems that economic activities entail transport demand. In this perspective, an essential analysis deals with the quantification of the relationships between transport demand and certain socioeconomic variables. Elasticity is a concept widely used in transport economics as a measure of the responsiveness of transport demand concerning different factors represented as independent variables in an econometric model and coupling/decoupling concepts have been proposed in literature. This paper deals with the estimation of elasticities of motorway traffic demand based on Gross Value Added (GVA), and the consequent investigation of coupling/decoupling situation. The analysis is based on the application of an Autoregressive-Distributed Lag (ARDL) cointegration model with the F-bound test and of the related Error Correction model. Starting from the general ARDL model and the methodology for the verification of its robustness, the same model is applied to the Italian toll road network. The time series of GVA for goods and services and the overall length of the toll network from 1995 to 2019 are considered as explanatory variables of the total annual distance traveled by light and heavy vehicles. The various tests in the ARDL framework show a cointegration between the variables, under the fulfillment of all the diagnostic requirements. In this way, the long-term elasticities and the short-term adjustment dynamics are estimated separately for the goods and services components of GVA, and light and heavy vehicles. Starting from stable estimates of elasticities, the long-term coupling and decoupling effects between motorway traffic of light and heavy vehicles and the national production of goods and services can be shown. The paper, as well as providing an updated picture of the Italian situation, identifies a methodological framework that can be transferred to other contexts for a sector of great interest to investors, such as the motorway sector. All this can be useful to meet the needs of numerous stakeholders, who want to deepen the links between the economic cycle and traffic demand on toll motorways.

Keywords: ARDL Cointegration, Error Correction Model, Toll Roads, Traffic Demand Elasticity, Coupling/Decoupling

To cite this article:

Contact:
1) andrea.pompigna@unitn.it [https://orcid.org/0000-0002-8473-7936] - corresponding author,
2) raffaele.mauro@unitn.it [https://orcid.org/0000-0002-1267-9335]

Article is available in open access and licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0)
1. Introduction

The conventional perspective in transport and economic geography leans on the assumption that transportation demand is a derived demand, both for passengers and freight transport (Rodrigue, 2006). Except for some exclusive cases of travels carried out simply to enjoy the ride, transportation is consequently an activity that depends on other activities derived from spatial complementarities between a certain supply at an origin and a certain demand at a destination, without which the trip loses its purpose and it has no more reason to take place (Rodrigue, 2006; Bamford, 2006). These general considerations lead to the formulation of the basic axiom that economic activities entail transport demand (Müller et al., 2016) and this transport demand generates impacts on the environment (Chamier-Gliszczyński, 2011; Chamier-Gliszczyński, 2012a; Chamier-Gliszczyński, 2012b).

Thus, the question that immediately arises is: how can this induction of traffic demand by economic activities be measured and represented? Is there a functional relationship between economic fundamentals and transportation demand? Several studies have addressed this issue, even recently, to answer these questions especially for air transport (e.g. Chi and Baek, 2013; Marazzo et al., 2010; Brida et al., 2016b; Brida et al., 2016a; Balsalobre-Lorente et al., 2021; Adedoyin et al., 2020) or freight transport (e.g. Nuzzolo et al., 2013; Yang, 2015; Mauro and Pompigna, 2019; Pompigna and Mauro, 2020a; Pompigna and Mauro, 2020b) and also deepening some recent trends that show, again in the case of freight transport, a decoupling (Profillidis and Botzoris, 2018) between economic growth and traffic demand (e.g. Kveiborg and Fosgerau, 2005; Ecola and Wachs, 2012; Alises et al., 2014; Alises and Vassallo, 2015). In this regard, it should be emphasized here that the link between transport demand and economic activities is of the cross-type. If on the one hand, there is the axiom of a transport demand that is generated from the presence of economic activities, on the other hand, several studies have proven the existence of a reciprocal relationship between the development of transport systems and economic dynamics (e.g. Banister, 2012; Pradhan and Bagchi, 2013; Mohmand et al., 2017; Vlahinić Lenz et al., 2018; Pradhan, 2019; Magazzino and Mele, 2020). Also, in this case, the questions above can be re-proposed with reversed roles. By including both these points of view, the possible bi-directionality of conditioning has been addressed by various studies, which have investigated the causal link between traffic and economic growth (e.g. Pradhan, 2010; Hakim and Merkert, 2016; Achour and Belloumi, 2016; Zhang and Graham, 2020; Flores and Chang, 2020), and then the nexus with atmospheric emissions, environmental alteration and climate change (Erdogan et al., 2020).

Beyond the specific interest in these research themes, which are engaging in understanding the cross-effects between economics and transport, a key aspect for the evaluation of policies and investments in the field of transport infrastructure economics regards the knowledge of the evolution of transport demand over time (Pompigna and Rupi, 2018), depending on the economic cycle of a certain region, for descriptive and forecasting needs (Profillidis and Botzoris, 2018). This is a critical aspect especially for tolled motorways, which over the years have seen an increase in the number of stakeholders involved in their planning, design, and management under innovative financing systems, with the participation of mixed entities in a Public-Private Partnership (PPP) (Pompigna et al., 2015). Actually, for these stakeholders, which are involved with different roles in the evaluation of investment initiatives in the road sector (e.g. central governments and national agencies, regional and local authorities, construction companies, concessionaires and managers of transport services, lenders, and investment funds), identifying the key parameters and their strength in influencing traffic demand on toll roads is essential (Gomez et al., 2016).

In this perspective, an essential analysis deals with the quantification of the relationships between transport demand and certain socioeconomic variables, which are representative of the reference context and by which analysts can set the forecast scenarios of its evolution (Pompigna et al., 2015; Gomez et al., 2016). Elasticity is a concept widely used in transport economics as a measure of the responsiveness of transport demand concerning different factors represented as independent variables in an econometric model. Transport demand elasticity can be defined as the percentage change in the transportation demand (i.e. the dependent variable in the econometric model) produced from a unit fluctuation of a certain explanatory variable (i.e. an independent variable in the econometric model), which
ofmacroeconomic and demographic fundamentals (Profillidis and Botzoris, 2018; Pratt, 2013).

Following the analysis of elasticities in long-term equilibrium relationships, their values also allow the evaluation of the coupling/decoupling effect between traffic demand and the economic system, which are currently of great attention for researchers, practitioners, and policymakers. The concept of coupling and decoupling in the transport sector was highlighted and theoretically defined by Tapio (Tapio, 2005), investigating the real entity and the sign of the link between traffic volumes, especially road traffic volumes, and economic activities, particularly the connections with the Gross Domestic Product (GDP).

Until a few years ago, the assumption shared all over the world considered that if a relationship between economic growth and traffic demand can be demonstrated, the trend in their rates must be coupled (Profillidis and Botzoris, 2018). This is what is called coupling, usually expressed between transport demand and GDP. On the contrary, decoupling reflects the de-linking between economic growth and growth in traffic demand, i.e. rates of change in transport demand don’t reflect rates of change in economic activity (Profillidis and Botzoris, 2018).

The decoupling of the demand for passenger and freight transport from economic activity, which in the motorway sector is reflected in the traffic of light and heavy vehicles, can be observed for developed economies thanks to technological progress (e.g. changes in the composition and weight of the economic sectors, increase in the use of the internet for remote communications and commerce) and efficiency increase in development patterns (e.g. new configurations of urban and territorial development both for residential and manufacturing activities, modern forms of integration of transport and logistics). Indeed, decoupling appears as the result of greater efficiency in managing economic growth, with no increase or even with a reduction of congestion in transport systems and pressure on environmental factors (Erdogan et al., 2020) in developing proecological transport systems (Jacyna et al., 2018). According to Tapio (Tapio, 2005), the decoupling of transport volume growth from economic growth can be seen when the elasticity values are less than 1.0. Therefore, with respect to different elasticity values, transport volume growth and economic growth can be coupled, negatively decoupled and decoupled. Ultimately, Tapio (Tapio, 2005) define different degrees of coupling and decoupling of transport volume growth (ΔVOL) from economic growth (ΔGDP) with reference to the elasticity E=%ΔVOL/%ΔGDP, which has now become a shared reference in the analyzes and discussions on this issue (Profillidis and Botzoris, 2018).

That said, as indicated in (Gomez et al., 2016), there are not many studies in the literature regarding toll motorways. Referring to this research and also to (Dunkerley et al., 2015) for a review of the updated literature on the subject, these studies were often carried out considering an aggregated motorway traffic description, i.e. without distinguishing light and heavy transport demand, and generally assuming a highly aggregated economic indicator like GDP as the only socioeconomic explanatory variable included in the analysis. Among these studies, we can mention (Gately, 1990) for the U.S.; McKinnon, 2007; Agnolucci and Bonilla, 2009) for the UK; (Libardo and Nocera, 2008) for Italy; (Li and Hensher, 2009) for Australia; (Matos and Silva, 2011) for Portugal; (Matos and Raymond, 2003), (Gomez and Vassallo, 2015) and (Gomez et al., 2016) for Spain. These analyses, in addition to not being very numerous, by their very nature strongly depend on the reference economic context. Although the models can be generalized, in terms of the type and of the specification of the variables, it is clear that the results, in terms of estimated values of the elasticities, cannot be transferred tout court in time (different time horizons) and space (different geographical and economic contexts) because countries go through different states of economic, social and technological development, constantly changing at different rates (Ecola and Wachs, 2012).

Nowadays, this type of analysis is requested for identifying the key parameters and their strength in influencing traffic demand on toll roads by an increasing number of stakeholders involved in the planning, design, financing, construction, maintenance, and operations of motorway facilities. Public contribution and potential PPP in road projects need to develop a long-term strategy, which must necessarily face the preparation of a detailed financial plan (Pompigna et al., 2015). Among other things, the financial plan should consider accurate traffic forecasts from a long-term perspective. Besides financial aspects, we also find other important issues
involving traffic growth forecasts on motorways, such as any needs related to the operation, management, and maintenance of the infrastructure and to the costs/benefits balance for the stakeholders and the local communities that are crossed and served. All these aspects need long-term perspectives by considering future scenarios that represent a certain vision on the progress of the socio-economic system. Once again, the problem of identifying an equilibrium relationship between traffic demand and the economic cycle greatly stands as a factor for understanding the underlying dynamics and for estimating future trends and forecast figures. Thus, the main objective of the paper is to define a general framework for exploring the relationships between the economic cycle and traffic demand on toll road systems. This interest is mainly aimed at providing a useful tool in the field of analysis and forecasts of motorway traffic, in consideration of a certain context, or rather a scenario, of conceivable economic evolution for a certain country in a medium-long term perspective.

Given the objective identified above, this research work is addressed using the methods of cointegration and error correction (Enders, 2014), and in particular the ARDL cointegration model with F-bound test (Pesaran and Shin, 1999; Pesaran et al., 2001), widely used in various sectors of econometric application. The purpose is to provide an analytical framework for estimating the elasticity values of the motorway traffic demand with respect to variables that are representative of the economic system and its cycle by preventing the problems of possible spurious regressions. These elasticities can be discussed according to (Tapio, 2005) for the characterization of the coupling/decoupling between transport and demand economic growth. This characterization provides further elements for the formulation of analysis and forecast evaluations in the motorway field, detailing the demand for the light and heavy traffic components and the economic system with respect to its different sectors. Regarding the latter, in this study the Gross Value Added (GVA) is assumed as a suitable explanatory macroeconomic variable for the economic activity to characterize the performance of the economic system instead of the most used GDP. Using the GVA, the model considers the actual production dynamics in the raw materials transformation process, with the possibility of considering the various sectoral contributions, instead of the necessary aggregation of GDP. In this way, the cointegration model assumes a greater articulation in the description of the effects of the economic cycle on the demand for motorway transport, allowing to study separately the elasticity and coupling/decoupling effects for different economic sectors.

The paper highlights the various steps of the ARDL cointegration and elasticities evaluation resorting to a real case study related to the Italian tolled motorway network by using the most recent data, i.e., up to 2019. In this perspective, because of this practical application, this research also aims to contribute to a better knowledge of the evolution of light and heavy vehicle demand on Italian toll roads. As far as we know, there is only one research by (Libardo and Nocera, 2008) based on a simple regression model and with data up to 2005, and therefore well before the economic and financial crisis that characterized the subsequent periods. In any way, beyond the estimates obtained for the model parameters in the Italian case study, the proposed model has general characteristics that allow it to be applied even in other contexts. This paper is organized as follows: section 2 describes the methodology and the model structure, section 3 shows the characteristics of the data sample for the case study, and section 4 and 5 reports the results and presents results’ discussion. Finally, the main conclusions are presented in section 6.

2. Methods
For the quantification of the elasticities of transport demand, aggregate econometric models are used (Mauro and Pompigna, 2019). These models express the relationship between transport demand and a set of explanatory variables, according to a generic function:

\[ y = f(x_1, x_2, \ldots , x_n) \]  

(1)

Where:
- \( y \) – the transport demand (dependent variable);
- \( x_i \) – the explanatory variables (independent variables) \( (i = 1, \ldots , n) \).

For the definition of the most appropriate functional form to describe the demand function \( y = f(x_1, x_2, \ldots , x_n) \) we can refer to the multiplicative function:
\[ y = e^{a_0} [x_1^{a_1} \cdot x_2^{a_2} \cdot \ldots \cdot x_i^{a_i} \cdot \ldots \cdot x_n^{a_n}] = e^{a_0} \prod_i x_i^{a_i} \]  \hspace{1cm} (2)

with \( \alpha_0 \) and \( \alpha_i \) as constant parameters, and in particular \( \alpha_i \) representing the elasticities of \( y \) with respect to independent variables \( x_i \). If we express \( y = f(x_1, x_2, \ldots, x_n) \) with the multiplicative model in Equation (2), by transforming all the variables according to the natural logarithm, we obtain a linear expression of the model, of the following type:

\[ \ln(y) = \alpha_0 + \sum_i \alpha_i \ln(x_i) \]  \hspace{1cm} (3)

The linear model in Equation (3) can be solved with the methods of linear regression, by introducing an appropriate error term \( \varepsilon \). Considering that we have for a certain interval \( T \) (e.g., for \( T \) years) the time series of the values actually recorded for the dependent variable \( y \) and for the independent variables \( x_i \) in each of the homogeneous sub-intervals \( t \) into which it can be subdivided (e.g., for every year \( t \), for the generic \( t \) we can write:

\[ \ln(y)_t = \alpha_0 + \sum_i \alpha_i \ln(x_i)_t + \varepsilon_t \]  \hspace{1cm} (4)

The application of the Ordinary Least Squares (OLS) linear regression method, considering the equations in the form of Equation (4) in each of the periods that form \( T \), allows to estimate the value of the intercept \( \alpha_0 \) and the elasticities \( \alpha_i \) with respect to each independent variable \( x_i \).

However, this simple OLS regression model, which at first glance appears easy and effective to use, hides some pitfalls and problems. As is well known, after the seminal work of Granger and Newbold (Granger and Newbold, 1974) and the literature that originated from this fundamental study, using OLS regression with data from non-stationary time series can lead to false inferences by configuring the so-called spurious correlation. A spurious correlation occurs in an OLS regression when non-stationary time series appear to be related according to the usual statistical criteria, but without any real sense. The concept of avoiding the spurious regression among variables was firstly explored by Engle and Granger (Engle and Granger, 1987) in their seminal paper, which introduced cointegration, i.e., the existence of a long-run relationship. These authors argued that spurious cointegration could be avoided if the I(1) time series are cointegrated, or in other words if the series are stationary once they have been transformed using first order differences, and if a linear combination of the original series is also stationary. Following these early discussions, cointegration tests and studies and empirical applications were carried out, defining the so-called classical cointegration methodology according to (Engle and Granger, 1987; Johansen and Juselius, 1990; Johansen, 1995).

An alternative methodology that has found considerable use in recent years is the so-called Bound Cointegration Test, introduced by Pesaran (Pesaran and Shin, 1999; Pesaran et al., 2001). This procedure employs the ARDL methodology and is also known as the proof-of-limits cointegration procedure. The ARDL cointegration approach presents several advantages in comparison with classical cointegration methods, such as:

- it is relatively more efficient in the case of small and finite sample data sizes;
- it can be applied whether the variables under the study are not integrated of the same order, with the only constraint that no variable is I(2) or higher;
- it can regard equal or different orders of number of lag length for all variables, without affecting the asymptotic distribution of the test statistic;
- a related Error Correction Model (ECM) provides short-run coefficients as an adjustment dynamic with respect to the coefficients describing the long-run equilibrium, in other words it is possible to assess the short-run and long-run relationship between the given variables simultaneously.

ARDL stands for Autoregressive-Distributed Lag. The model has an autoregressive component in the sense that the dependent variable \( Y \) (in this case \( Y = \ln(y) \)) is in part explained by lagged values of itself (according to a lag order \( p \)), and a distributed lag component, in the form of successive lags of a number of \( X_i \) (in this case \( X_i = \ln(x_i) \)) explanatory variables (each appearing with a lag order \( q_i \)). Considering the simplest case of a single explanatory variable (it is easy to extend it to the case of several explanatory variables), an ARDL regression model of order \((p, q)\) can be expressed with the following equation:
\[ Y_t + \sum_{i=1}^{p} \beta_i Y_{t-i} = \lambda + \sum_{j=0}^{q} \alpha_j X_{t-j} + \varepsilon_t \]  

(5)

Where:
\( \varepsilon_t \) – is a random error term which will be serially independent.

The coefficients can be estimated according to the Ordinary Least Squares (OLS) method, in compliance with the relative assumptions. Considering the more general case of the presence of a temporal trend, as well as a constant, the ARDL model can be written in the following form:

\[ \Delta Y_t = c_0 + c_1 t + (\theta_0 + \theta_1 t + \rho Y_{t-1} + \sigma X_{t-1}) + \sum_{j=0}^{q} \delta_j \Delta X_{t-j} + \varepsilon_t \]  

(6)

where \( \Delta \) is the first difference operator (\( \Delta Y_t = Y_t - Y_{t-1} \)). The parameters \( c_0 \) and \( c_1 \) are the so-called unrestricted intercept and unrestricted linear time trend; the term \( (\theta_0 + \theta_1 t + \rho Y_{t-1} + \sigma X_{t-1}) \) represents the long run relationship with \( \theta_0 \) the restricted intercept and \( \theta_1 \) the restricted linear time trend, \( \gamma_1 \) and \( \delta_j \) the coefficient of the short-run dynamic with respect to the optimal lags \( p \) and \( q \). At this point, it is necessary to make some considerations on the selection of the optimal lags and on the characterization of error terms. The appropriate values for the maximum lags \( p \) and \( q \) (or possibly \( q_1 \), ..., \( q_m \) if there are \( m \) explanatory variables in the model) can be obtained considering a maximum value for lags and one or more of the information criteria AIC (Akaike Information Criterion), SC (Schwarz Criterion) or BIC (Bayes Information Criterion), HQ (Hannan-Quinn criterion), etc. (Badshah and Bulut, 2020) based on a high log-likelihood value, which includes a penalization for more lags to achieve this, whose form varies from one to another. SC (or BIC), is considered as a consistent model-selector regarding optimal lags, providing slightly better estimates than the AIC criteria in small samples (Pesaran and Shin, 1999). For yearly data, for example, a maximum lag of 2 or 3 should not be exceeded. Once the optimal lags have been selected and the OLS estimate of Equation (6) has been provided, diagnostic tests on residuals must be performed. These must respect the normal distribution, be serially independent and homoscedastic. Usual error distribution tests can be performed to assess their normality (e.g. Jarque-Bera normality test), and then an LM (i.e. Breusch-Godfrey) test to evaluate the null hypothesis that the errors are serially independent, and an ARCH test for autoregressive conditional Heteroscedasticity. However, as non-normality, serial correlation, and Heteroscedasticity should not be present, the lag length should be adjusted for the possible biases with some care to be taken not to "over-select" the maximum lags (Pesaran et al., 2001). If the assumptions under which the OLS estimates are unbiased and consistent are respected, the stability of the estimated parameters can be tested resorting to CUSUM and CUSUMSQ tests proposed in (Brown et al., 1975). These tests consider the updated cumulative sum of the recursive residuals (CUSUM) and squared recursive residuals (CUSUMSQ) plotted against the breakpoints for the 5% significance line: if the plot of CUSUM and CUSUMSQ fall within the 5% significance band, the long-run and short-run estimated coefficients can be considered stable over the period.

At this point, it is possible to evaluate the existence of a long-term relationship, that is, that the variables are effectively cointegrated. The null hypothesis of no cointegration is that long-run coefficients are all equal to zero. Thus, for the testing of the existence of a long-run relationship, the null hypothesis is tested against the alternative hypothesis of cointegration, i.e. that the aforementioned coefficients are significantly different from zero. According to (Pesaran et al., 2001), the F test can be used to establish if a long-run relationship exists, considering a non-standard distribution that depends on the number of independent variables and their mix of I(0) and I(1) variables, and the presence or not of the intercept and/or trend term. The criteria in (Pesaran et al., 2001) consider upper and lower bounds of critical values and three different cases: computed F-statistic greater than the upper bound = rejection of the null hypothesis, i.e. existence of a long-run (level) relationship confirmed; computed F-statistic smaller than the upper bound = no rejection of the null hypothesis, i.e. the existence of a not-significant long-run (level) relationship; computed F-statistic between the upper and lower bound = inconclusive test. Now we must clarify the long-run relationship in the context of the ARDL model. If we consider Equation (6), a long-run relationship implies the existence of long-run equilibrium, with no tendency for change.
This implies that in the long-run equilibrium the first differenced variables in Equation (6) will be zero, then:

\[ \rho Y_{t-1} + \sigma X_{t-1} + \beta_0 + c_1 t + \epsilon_t = 0 \]  

(7)

where the final long-run coefficient for X is \(-\sigma/\rho\). Given a satisfactory result of the F-bound test, the long-run relationship is significantly a non-spurious regression, as a linear combination of the non-stationary variables is stationary in a simple OLS framework, that is \(\theta_0 + \theta_1 t + \rho Y_{t-1} + \sigma X_{t-1}\), and then:

\[ Y_t = \theta_0 + \theta_1 X_t + \theta_2 (t + 1) + \epsilon_t = 0 \]  

(8)

As \(\hat{\theta}_0\), \(\hat{\theta}_1\) and \(\hat{\theta}_2\) are the estimate OLS parameters for Equation (8), an error correction term is defined by:

\[ EC_{t-1} = Y_{t-1} - \hat{\theta}_0 - \hat{\theta}_1 X_{t-1} - \hat{\theta}_1 t \]  

(9)

that coincides with the residuals from Equation (8). Considering Equation (6) the term in the 1-lag variables \(Y_{t-1}\) and \(X_{t-1}\) can be replaced with the correction \(EC_{t-1}\), in the form:

\[ \Delta Y_t = c_0 + c_1 t + \sum_{i=1}^{p} \gamma_i \Delta Y_{t-i} \]

\[ + \sum_{j=0}^{q} \delta_j \Delta X_{t-j} + \lambda EC_{t-1} + \epsilon_t \]  

(10)

which can be estimated in the parameters \(\gamma_i\) and \(\delta_j\) that represent the short-run coefficients. In order to converge to equilibrium, \(\lambda\) must be negative and statistically significant, confirming the existence of a stable long-run relationship and cointegration between X and Y. The coefficient also represents a measure of the speed with which the adjustment towards equilibrium is expressed, in terms of the percentage of absorption of a shock in a period (e.g. % of return towards equilibrium in a year for annual time series).

3. Data

The choice of the variables is dictated by the objectives of the research, which are those already declared to investigate the existence of a long-term relationship between the macroeconomic fundamentals and the total annual distances of the journeys of light and heavy vehicles on the entire national motorway network. In applying to the case study, data relating to Italy were used for the period between 1995 and 2019.

- Total annual km traveled by light vehicles, TKLM, which represents the product of the volume of light vehicles in circulation for each year on the national motorway network and the distance in km covered by each of them (passenger cars from class A, according to the Italian classification).

- Total annual km traveled by heavy vehicles, TKMH, which represents the product of the volume of heavy vehicles in circulation for each year on the national motorway network and the distance in km covered by each of them (from classes B, 3, 4 and 5 according to the Italian classification).

These time series cover the period 1995-2019 and are expressed in millions of vehicles * km. These series were obtained from the database of the AISCAT organization (AISCAT, n.d.), which brings together all the motorway concessionaires operating on the Italian tolled network. Total annual km traveled (TKM) by light and heavy vehicles is a very suitable indicator to express the use of transport systems and therefore to express the aggregate traffic demand (Ecola and Wachs, 2012). In the case of motorway traffic on tolled infrastructures, this indicator can be calculated very easily in the case of kilometric tolls, since the entry and exit of each vehicle from the infrastructure is known for accounting purposes.

Based on the analysis in (Müller et al., 2016) for freight transport, GVA is assumed in this study as a suitable explanatory macroeconomic variable for the economic activity instead of the most used GDP both for light and heavy traffic and considering the effects due to the production components of goods and services. Since GVA is the difference between the final value of the goods and services produced and the value of the goods and services purchased to be used in the production process, it represents a measure of the gross increase in the resulting value of economic activity as the sum of the contributions of the individual production sectors. Thus, using GVA, the model considers the actual dynamics of national production in the process of transformation of raw materials, with the possibility of considering the various sectoral contributions, instead of the necessary aggregation of GDP. In this way, the
Cointegration model takes a greater articulation in the description of the effects of the economic cycle on the demand for motorway transport, allowing to study the elasticity and coupling/decoupling effects of the goods and services components of GVA separately. The explanatory variables used in the case study for the Italian economy and motorway network are defined as follows:

- Total extension of the toll motorway network for each year, expressed in km, MNKM;
- Gross Value Added of the national economy relating to the production of goods (including agricultural and industrial sectors), AVG;
- Gross Value Added of the national economy relating to the production of services, AVS.

---

**Fig. 1.** Log-transformed Time series for: (a) Total annual km traveled by light vehicles; (b) Total annual km traveled by heavy vehicles; (c) Gross Value Added of the national economy relating to the production of goods; (d) Gross Value Added of the national economy relating to the production of services; (e) Total extension of the toll motorway network for each year.
All the time series cover the period 1995-2019. MNKM time series were obtained from the database of the AISCAT organization (AISCAT, n.d.), while AVG and AVS were obtained from the database of ISTAT, which is the Italian national statistical institute (ISTAT, n.d.). Time series for AVG and AVS are expressed in millions of euros per year as chained-linked values with 2015 as the reference year. It must be highlighted that a logarithmic transformation was carried out for each of the original variables. The natural logarithm has been applied to all the data to account for the direct expression of the elasticity by regression coefficient as in Equation (2) and (3). The natural logarithms of TKML, TKMH, MNKM, AVG and AVS are denoted as LTKML, LTKMH, LTKMT, LMNKM, LAVG and LAVS respectively. Table 1 shows some descriptive statistics for all transformed variables for time series in Figure 1.

### 4. Results

As we said in Section 2 the F-bounds ARDL cointegration does not work with I(2) variables. Thus, we must investigate if the given time series are I(0) and/or I(1). This means that it is necessary to identify how many differentiations must be carried out on the initial series to have stationarity, i.e. that the series of differences does not have a unit root. Therefore, the stationarity of a time series can be examined by unit root tests. These tests provide statistical evidence on the stationarity of a given series using different methods, among which the most used are Augmented Dickey-Fuller, Phillips-Perron, and KPSS procedures. It should be emphasized that time series, and in particular economic time series, can present structural breaks, which are sudden upward or downward changes due to contingent situations of various kinds (policy changes, external shocks, crises, etc.). In the presence of structural break, conventional unit root test methods may show a time series to be non-stationary (Perron, 1989) and this can generate problems in investigating the order of integration of a series. In the case of ARDL cointegration, being able to proceed in the presence of series I(0) or I(1), the problem could arise if the presence of structural breaks led to erroneously assigning orders I(2) or higher. This would lead to erroneously consider the ARDL model not applicable, due to incorrect attribution of the order of integration of the series. This could also happen in the case study, as appears from the first analysis of the progressions of the series in the period 1995-2018, which are represented in Figure 1.

Given this, Perron (Perron, 1989) has developed a unit root test method, which accommodates a known structural break in the time series, with extended discussions in (Vogelsang and Perron, 1998), (Zivot and Andrews, 2002) and (Banerjee et al., 1992), with further contributions from other studies reviewed in (Perron, 2006). In this case, we proceeded by applying the model in Eviews 10 (EViews, n.d.) and in particular the innovational outlier (IO) model, which assumes that the break occurs gradually unlike the alternative additive outlier (AO) model (i.e. assuming the breaks occur immediately).

For the automatic break selection method, we assume the Dickey-Fuller t-statistic minimization in Eviews 10 (EViews, n.d.) considering trend and intercept specification. Applying the unit root test with breakpoint, the results of the Augmented Dickey-Fuller (ADF) test statistic leads to clear evidence that all the time series are I(1). Thus, we can reject the hypothesis that the data are I(2), which is important for the legitimate application of the F-bounds test. Table 2 shows the results of the test conducted for the time series defined in section 3, both for the levels and for the 1st differences. In consideration of the objectives of the study,
therefore, the ARDL equations for our study are presented as follows:

**Model 1 – light vehicles**

\[
\Delta LTKML_t = \left( \theta_1 + \rho_1 LTKML_{t-1} + \sigma_1 LAVG_{t-1} \right) + \left( \sum_{j=0}^{q.1} \gamma_{1,j} \Delta LTKML_{t-j} \right) + \left( \sum_{j=0}^{q.2} \delta_{1,j} \Delta LAVG_{t-j} \right) + \left( \sum_{j=0}^{q.2} \delta_{2,j} \Delta LAVS_{t-j} \right) + \left( \sum_{j=0}^{q.3} \delta_{3,j} \Delta LMNMK_{t-j} + \epsilon_{Lt} \right)
\]

\[\text{(11)}\]

**Model 2 – heavy vehicles**

\[
\Delta LTKMH_t = \left( \theta_H + \rho_H LTKMH_{t-1} + \sigma_2 LAVG_t + \sum_{j=1}^{q.1} \gamma_{2,j} \Delta LTKMH_{t-j} \right) + \left( \sum_{j=0}^{q.2} \delta_{1,j} \Delta LAVG_{t-j} \right) + \left( \sum_{j=0}^{q.3} \delta_{2,j} \Delta LAVS_{t-j} \right) + \left( \sum_{j=0}^{q.3} \delta_{3,j} \Delta LMNMK_{t-j} + \epsilon_{Ht} \right)
\]

\[\text{(12)}\]

We used the ARDL F-bound cointegration test in Microfit 5.50 (Pesaran and Pesaran, n.d.) to examine the existence of a cointegration relationship for the two models. Considering a maximum lag 3 for all the variables, the optimal lag orders of the models are identified using the Schwarz Bayesian Criterion (SBC), which operates better in small samples in the ARDL framework than the alternative AIC criteria (Pesaran and Shin, 1999).

The results shown in Table 3 confirm that long-run cointegration exists for the two cases considered, fully respecting the required confidence limits. Table 4 shows the tests for normality (Jarque-Bera), absence of serial correlation (Breusch-Godfrey), Heteroscedasticity (White test and Auto-Regressive Conditional Heteroscedasticity, ARCH, test) and functional form misspecification (Ramsey Regression Specification Error Test, RESET (Ramsey, 1969), for non-linear combinations). All these tests are satisfied and therefore no problems appear in the OLS estimation of the three models. As indicated in (Pesaran et al., 2001) and (Pesaran and Pesaran, 2009) the cumulative sum of recursive residuals (CUSUM) and the CUSUM of square (CUSUMQ) test (Brown et al., 1975) are used to assess parameter stability in estimated models. Figures 2 and 3 show the results for both tests and both models and indicate the absence of any instability of the coefficients, with the plots falling inside the critical bands of the 5% confidence interval of parameter stability. Table 5 shows the long-run coefficients and the error correction representation for each of the three models estimated with Microfit 5.50 (Pesaran and Pesaran, n.d.). Finally, Figure 4 represents the in-sample prediction values from the ARDL models versus the actual values for total distances traveled by light and heavy vehicles during the whole period 1995-2019 and the residuals, revealing a satisfactory fitting of the actual values.

**Table 2. Unit Root Tests with Breakpoints – levels and 1st differences**

<table>
<thead>
<tr>
<th>Series</th>
<th>Levels</th>
<th>1\textsuperscript{st} differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t-Statistic</td>
<td>Prob.\textsuperscript{1}</td>
</tr>
<tr>
<td>LTKML</td>
<td>-4.764893</td>
<td>0.135</td>
</tr>
<tr>
<td>LTKMH</td>
<td>-3.520597</td>
<td>0.8041</td>
</tr>
<tr>
<td>LMNMK</td>
<td>-3.893841</td>
<td>0.5922</td>
</tr>
<tr>
<td>LAVG</td>
<td>-2.510286</td>
<td>&gt; 0.99</td>
</tr>
<tr>
<td>LAVS</td>
<td>-3.147357</td>
<td>0.9366</td>
</tr>
</tbody>
</table>

Null Hypothesis: Series has an unit root; Trend and Break Specification: Trend and intercept; Break Type: Innovational outlier; Break selection: Minimize Dickey-Fuller t-statistic.

**Table 3. Results of ARDL F-bound test (Null Hypothesis: No long-run relationships exist)**

<table>
<thead>
<tr>
<th>Model</th>
<th>Selected lags</th>
<th>F-stat</th>
<th>95%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td>1:</td>
<td>(3,3,3,2)</td>
<td>12.917</td>
<td>3.9855</td>
<td>5.4519</td>
</tr>
<tr>
<td>2:</td>
<td>(2,2,2,3)</td>
<td>5.634</td>
<td>3.9855</td>
<td>5.4519</td>
</tr>
</tbody>
</table>
5. Discussion

The estimated coefficients of the long-run relationships are significant in each of the two models for all the variables. In model 1 we can observe that the extension on the motorway network and both goods and services components of the national added value have a positive significant impact on the total distances traveled by light vehicles. With a coefficient 1.20, a 1% increase in the total length of the national motorway network will cause the total distance traveled by light vehicles to increase by 1.20% in the long run. About the components of national added value, the long-run coefficients highlight a greater elasticity for the component which relates to services: an increase of 1% leads to an increase of 1.18% in the total mileage of light vehicles, against a corresponding increase of 0.45% due to a 1% increase in the added value for goods.

Table 4. Diagnostic tests

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Breusch-Godfrey Serial Correlation Test:</td>
<td>F(1,6) = .28804</td>
<td>0.611</td>
</tr>
<tr>
<td></td>
<td>Jarque-Bera Normality test</td>
<td>CHSQ(2) = .18515</td>
<td>0.912</td>
</tr>
<tr>
<td></td>
<td>Heteroscedasticity White test</td>
<td>F(1,20) = .74490</td>
<td>0.398</td>
</tr>
<tr>
<td></td>
<td>Heteroscedasticity ARCH test</td>
<td>F(1,6) = .6786E-3</td>
<td>0.980</td>
</tr>
<tr>
<td></td>
<td>RESET</td>
<td>F(1,6) = .4790E-3</td>
<td>0.983</td>
</tr>
<tr>
<td>2</td>
<td>Breusch-Godfrey Serial Correlation Test:</td>
<td>F(1,8) = 3.8668</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>Jarque-Bera Normality test</td>
<td>CHSQ(2) = 3.0437</td>
<td>0.218</td>
</tr>
<tr>
<td></td>
<td>Heteroscedasticity White test</td>
<td>F(1,20) = .56674</td>
<td>0.460</td>
</tr>
<tr>
<td></td>
<td>Heteroscedasticity ARCH test</td>
<td>F(1,8) = .016628</td>
<td>0.901</td>
</tr>
<tr>
<td></td>
<td>RESET</td>
<td>F(1,8) = .065564</td>
<td>0.804</td>
</tr>
</tbody>
</table>

**Fig. 2.** CUSUM (a) and CUSUMSQ (b) for parameter stability tests of Model 1 in equation (11)

**Fig. 3.** CUSUM (a) and CUSUMSQ (b) for parameter stability tests of Model 2 in equation (12)
Table 5. Estimated Long Run Coefficients and EC Representation for the Selected ARDL Model

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMNKM</td>
<td>1.2011</td>
<td>4.0465[.005]</td>
</tr>
<tr>
<td>LAVG</td>
<td>0.45039</td>
<td>4.5147[.003]</td>
</tr>
<tr>
<td>LAVS</td>
<td>1.1765</td>
<td>11.5246[.000]</td>
</tr>
<tr>
<td>EC</td>
<td>LTKML - 1.2011<em>LMNKM - .45039</em>LAVG - 1.1765<em>LAVS + 21.5290</em>C</td>
<td>LTKMH - 1.1292<em>LMNKM - .79159</em>LAVG - 1.4858<em>LAVS + 30.8206</em>C</td>
</tr>
</tbody>
</table>

Regressor | Model 1 | Model 2 |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>∆LTKML1</td>
<td>0.57203</td>
<td>4.5452[.001]</td>
</tr>
<tr>
<td>∆LTKML2</td>
<td>0.24623</td>
<td>1.5184[.160]</td>
</tr>
<tr>
<td>∆LTKMH1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>∆LTKMT1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>∆LMNKM</td>
<td>-1.2278</td>
<td>-1.7044[.119]</td>
</tr>
<tr>
<td>∆LMNKM1</td>
<td>1.1879</td>
<td>2.0155[.072]</td>
</tr>
<tr>
<td>∆LMNKM2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>∆LAVG</td>
<td>0.084773</td>
<td>.91974[.379]</td>
</tr>
<tr>
<td>∆LAVG1</td>
<td>-0.43703</td>
<td>-4.3183[.002]</td>
</tr>
<tr>
<td>∆LAVG2</td>
<td>-0.18858</td>
<td>-2.0961[.062]</td>
</tr>
<tr>
<td>∆LAVS</td>
<td>0.57266</td>
<td>1.9160[.084]</td>
</tr>
<tr>
<td>∆LAVS1</td>
<td>-0.61495</td>
<td>-2.2673[.047]</td>
</tr>
<tr>
<td>∆LAVS2</td>
<td>-0.67112</td>
<td>-1.9455[.080]</td>
</tr>
<tr>
<td>EC(-1)</td>
<td>-1.1048</td>
<td>-6.4434[.000]</td>
</tr>
</tbody>
</table>

Fig. 4. Residual, actual and fitted time series for model 1 (a) and model 2 (b)
The situation is similar in model 2 for heavy vehicles, with an increase of about 1.13%, 1.49%, and 0.79% in the total annual traveled due to a 1% increase respectively for the total length of the national motorway network, added value for services and added value for goods. The elasticity values obtained in heavy vehicles compared to light vehicles are lower in terms of the length of the network, while they are higher for the added value components. In both cases (models 1 and 2) the coefficient relating to the added value of services appears substantially higher than that of the goods produced. The lagged error correction term in the short-run models is statistically significant and it appears with a negative sign, as requested.

In model 1, which regards light vehicles, EC(-1) appears with a coefficient -1.1048, which implies that the error correction process fluctuates around the long-run values, converging quickly to the equilibrium once the dampened fluctuations are complete (Narayan and Smyth, 2006).

Model 2 for heavy vehicles shows a negative and significant lagged error correction term, and it indicates that any deviation from the long-run equilibrium is corrected about 84% for each year. In summary, we can say that the long-run relationships estimated with the ARDL cointegration model show a response in the distance traveled that is always positive, and:

- more than proportional to the variation in the added value of the services produced by the national economy for both cases (i.e. light and heavy vehicles), and therefore elastic;
- less than proportional to the variation in the added value of the goods produced by the national economy for both cases (i.e. light and heavy vehicles), and therefore inelastic;
- more than proportional (elastic) for light and heavy vehicles with respect to the variation in the extension of the national motorway network.

Considering the distinction between coupling and decoupling by (Tapio, 2005) for the estimated elasticity values \( \hat{\varphi}_i \) (i.e. coupling for \( 0.8 \leq \hat{\varphi}_i \leq 1.2 \) and decoupling for \( \hat{\varphi}_i > 1.2 \) or \( \hat{\varphi}_i < 0.8 \)) we can identify

- for light traffic demand: a long-run coupling between motorway traffic demand and services component of GVA; a long-run decoupling between motorway traffic demand and goods component of GVA;
- for heavy traffic demand: a long-run decoupling between motorway traffic demand and goods component of GVA, even if very close to the limit; a long-run decoupling between motorway traffic demand and services component of GVA.

If it is clear a decoupling effect for heavy traffic with respect to the growth of domestic production, this effect appears to be different for the production of goods and the production of services. For goods, there is a less than proportional effect, which could be linked, for example, to the improvement of the logistic chain of inter-sectoral exchanges or the progressive increase in the unit value of the goods produced. For services, the effect is more than proportional, and in this case it could be linked, for example, to greater fragmentation of deliveries due to an increase in e-commerce.

For light traffic, on the other hand, a coupling was found between the demand for light traffic and the production of services, attributable to a not yet expressed incidence of remote communication technologies in the period, and a less than proportional decoupling compared to the production of goods, which could be linked to greater efficiency and integration of competitive transport systems and urban and territorial policies. Finally, the values obtained can be compared with the results of (Libardo and Nocera, 2008), albeit with the difference in the model considered, first of all for the reference to the sectoral GVA in this study and the GDP by the authors cited above.

Within the limits of the above, we observe that the elasticity values we have obtained for motorway traffic against economic growth are lower than those in (Libardo and Nocera, 2008). In the study, the elasticities are equal to 2.33 for passengers and 1.59 for goods, as a variation of the total distances against the variation in GDP in the period 1980-2005. In addition to the differences in the model and the variables, the different entity of the elasticity values obtained must necessarily be linked to the analysis period. The study does not consider, for chronological reasons, the effects of the international financial market crisis of 2008 and its implications on the Italian economy also in subsequent years, which are instead included in this work. On the other hand, by moving the time window from the 1980-2005 interval to the 1995-2019 interval, we must consider the change in the
boundary conditions that may have occurred, and which show themselves as determinants of a different characterization of the traffic trend on the road network and the economic cycle.

We can observe that these analyzes will have interesting implications also in the future, especially considering the effects induced by the COVID-19 pandemic, which began to spread around the world at the beginning of 2020 and which also seriously affected the Italian population. Despite having demonstrated good stability of the parameters in the analysis period through the statistical tests, the continuation of this research will also consider the effects of the pandemic and its evolution, also by comparing the situation of various European and non-European countries.

6. Conclusions

This research tried to estimate the relationships between light and heavy traffic, in terms of annual distances traveled, on the motorway network and the economic cycle. Using a multiplicative formulation for these relations between traffic demand and macroeconomic fundamentals, elasticity has been used as a responsiveness measure. To solve the problems related to the possibility of spurious regressions in the estimation of the aforementioned relationships, a cointegration approach has been adopted, which sees the elasticities as parameters of the long-term relationship between the variables. For a description of the economic cycle, the national GVA was examined. This macroeconomic fundamental was broken down into two components relating to goods and services, to allow a greater characterization of the effects of the economic cycle on the demand for motorway traffic.

Thus, an ARDL F-bound cointegration approach with the related Error Correction Model has been specified, to estimate the dynamic relationship between Gross Value Added components for goods and services production and total kilometers traveled. The model was applied to a case study considering the Italian tolled motorway network allowing a full and detailed discussion of its specification, calibration, and diagnosis phases. Annual data in the period 1995-2019 have been used for the total annual km traveled by light and heavy vehicles on the national motorway network, assuming explanatory variables related to the GVA concerning goods and services production. Since the dependent variables are in any case conditioned by the length of the motorway network and since this has varied during the period under examination, this additional variable has been considered as an explanatory variable in the model. The stationarity of the series to be used in this analysis has been examined by unit root tests in presence of structural breaks. As a result, growth variables appeared to be I(1), i.e. stationary in first difference. The ARDL F-bound cointegration approach has been used to examine the existence of an equilibrium relationship between variables for two distinct models, related to light and heavy vehicle traffic demand. An ECM approach allowed estimating short and long-run coefficients, the latter of which represent the elasticity of transport demand in relation to the explanatory variables for both models. The ARDL model estimations ensured the cointegration between variables, in compliance with the requirements of no autocorrelation, no Heteroscedasticity, correct functional identification, normally distributed errors, and parameters stability both for light and heavy traffic demand. ECM estimation highlighted a negative and statistically significant error correction coefficient either way. The results showed a dynamic of returning to long-term equilibrium after any shock. After a shock, this dynamic showed a very fast annual recovery for heavy traffic and fluctuations, that converges to equilibrium with the damping of oscillations for light traffic.

Thus, the long-run relationships that have been proven to exist show a response in the distance traveled that is for light and heavy vehicles: elastic in relation to services component of GVA; inelastic in relation to goods component of GVA; elastic in relation to the extension of the national motorway network. The elasticity values that have been estimated are stable and positive, highlighting: a long-run coupling between light transport and services component of GVA; a long-run decoupling between light transport and goods component of GVA and between heavy transport and both services and goods component of GVA.

Thus, heavy traffic revealed a decoupling effect in relation to the growth of domestic production, but this effect appeared different for goods and services components. The effect appeared to be less than proportional for goods, due for example to the occurrence of situations such as the improvement of the logistic chain of inter-sectoral exchanges or the pro-
gressive increase in the unit value of the goods produced. The same effect appeared to be more than proportional for services, linked for example to situations such as a greater fragmentation of deliveries due to an increase in e-commerce. For light traffic, on the other hand, we observed a coupling between the demand for light traffic and the production of services, probably linked to a not yet significant incidence of remote communication technologies, and a less than proportional decoupling compared to the production of goods, which could be linked to greater efficiency and integration of competitive transport systems and urban and territorial policies. These findings defined an updated picture of the effects that the Italian economic cycle exerts on the motorway traffic of both light and heavy vehicles, compared to previous studies, by moving towards a more updated time horizon. This allowed to take into account the change in the boundary conditions that may have occurred in the last years, and which showed themselves as determinants of a different characterization of the relationship between traffic demand on the road network and the economic cycle. Although the numerical estimates are limited to the specific case of Italy, in more general terms the ARDL cointegration methodology shows its appropriateness and usefulness for this type of analysis, aimed at identifying the values of the elasticity of the road transport demand in relation to a set of explanatory variables. As pointed out, these analyzes are required by an increasing number of stakeholders for identifying the key parameters and their strength in influencing toll roads traffic demand. The specified models also allow the evaluation of the coupling/decoupling effect between traffic and the economic system’s sources, which are currently of great attention. These analyses will have interesting implications also in the future, especially considering the effects induced by the COVID-19 pandemic, which began to spread around the world at the beginning of 2020 and which also seriously affected the Italian population. Although having proved the statistical stability of the parameters estimated through the most used tests, an in-depth study in this sense will require further investigations and studies. The continuation of this research will consider the effects of the pandemic and its evolution, also by comparing the situation of various European and non-European countries, and the use of more robust stability analysis techniques to define the dynamics of their variability over time and the possible consequent changes in the coupling/decoupling effects.

References


