



## Un metodo per la stima del costo del trasporto pubblico su gomma

### *A methodology to assess the road public transit cost*

Prof. Ing. Umberto PETRUCELLI<sup>(\*)</sup>  
Dott. Ing. Salvatore CARLEO<sup>(\*\*)</sup>

#### 1. Introduzione

È nota l'importanza attribuita, nei paesi che sovvenzionano il trasporto pubblico locale, al costo standard di questi servizi, sia come strumento per regolare i rapporti economici fra ente affidante ed azienda affidataria, sia come riferimento per un'equa ripartizione delle risorse finanziarie per sovvenzionare il servizio, fra gli enti locali che ne hanno competenza.

La bibliografia sui costi del trasporto contempla diversi studi che hanno individuato funzioni di costo per stimare le risorse necessarie a produrre un dato servizio di trasporto. In questo percorso bibliografico s'inserisce la presente ricerca, che propone un metodo sintetico di determinazione del costo chilometrico standard per l'esercizio di un servizio di trasporto pubblico su gomma, valido sia per l'ambito urbano quanto per quello extraurbano ma applicato solo a quest'ultimo per via della limitatezza dei dati che si sono resi disponibili. Il metodo proposto si basa su modelli appositamente costruiti che contemplano parametri desumibili dai programmi di esercizio, comunque noti anche al di fuori delle aziende esercenti. Nonostante la ristrettezza del campione disponibile, la validità del metodo risulta verificata nella capacità di adattamento ai dati a disposizione, come dimostrato anche da alcune applicazioni pratiche sviluppate.

#### 2. Il tema e lo stato dell'arte

La stima del costo standard, cioè di quel costo che un'azienda efficiente economicamente sostiene per fornire un determinato servizio di trasporto nell'ambito di un contesto operativo ben definito, presenta una certa complessità. Secondo la classificazione consolidatasi nella letteratura scientifica, i modelli a tal fine costruiti si possono suddividere in due principali categorie: ingegneristici o "bottom up" e statistici o "top down". I primi si avvalgono

#### 1. Introduction

*Is known the importance ascribed, in the countries that subsidize local public transit, to the standard cost of these services, both as a tool to regulate the economic relations between entrusting authority and entrusted company, and as a reference for a fair sharing of financial resources to subsidize the service, among local authorities having jurisdiction.*

*The bibliography on transport costs consists of several studies that have identified cost functions to estimate the resources required to produce a given transport service. This research fits in the aforesaid bibliographic path. We propose a synthetic method to assess the standard cost per kilometer of a road public transit service, valid both for urban and suburban and intercity transit, but here only applied to the latter because of the restricted available data. The methodology involves the built of specific models requiring the knowledge of parameters to be derived from operating programs or otherwise known even outside of providing companies. Despite the low available sample size, the validity of this method, in the adaptability to known data, is verified, as demonstrated by some developed applications.*

#### 2. The topic and the state of the art

*The assessment of the standard cost, that is the cost incurred by an efficient company to provide a specific transport service within a clearly defined operational context, is rather difficult. According to the classification used in the scientific literature, these models can be divided into two main categories: engineering or "bottom up" and statistical or "top down". The first ones make use of analysis of the costs incurred in the various business areas and this implies the knowledge of the production processes of each factor and their costs. These models, in addition to require parameters known only within the companies, can lead to ideal*

<sup>(\*)</sup> Università della Basilicata – Scuola di ingegneria – Potenza.

<sup>(\*\*)</sup> Ingegnere dei trasporti – libero professionista – Stigliano (Matera).

<sup>(\*)</sup> University of Basilicata – School of Engineering – Potenza.

<sup>(\*\*)</sup> Mobility engineer – freelance – Stigliano (Matera).

## POLITICA E ECONOMIA

di analisi dei costi sostenuti nelle varie aree aziendali e questo presuppone una conoscenza dei processi produttivi dei singoli fattori e dei rispettivi costi. Detti modelli, oltre a richiedere la conoscenza anche di parametri noti solo agli esercenti, risentono del rischio concreto di condurre a valori di costo ideali, discosti quindi dalle specificità degli operatori, poiché non fanno riferimento direttamente ai costi storici realmente sostenuti. Inoltre risulta spesso complicato, per alcune attività aziendali, modellare il legame tecnologico tra input e output, con il rischio di condurre a scelte discrezionali nella scelta dei fattori produttivi. I modelli "top down" si avvalgono invece di assunzioni e di analisi statistiche che inferiscono il costo sulla base di informazioni, cross-section o aggregate, provenienti in parte da aziende di trasporto operanti sul territorio la cui efficienza economica tuttavia non è comprovata, il che porta ad un costo che non è sempre di massima efficienza. Questi modelli presentano il rischio concreto di una stima dei costi che risente delle specificità delle aziende produttrici comprese nel campione indagato piuttosto che delle specificità dell'ambito in cui il servizio viene prodotto. Esistono poi anche modelli che hanno caratteristiche intermedie tra le due tipologie nominate e vengono definiti ibridi.

Molti degli studi analizzati sono stati indirizzati alla verifica di eventuali economie di scala, di densità e di scopo, più che alla costruzione e validazione di modelli di costo unitario di produzione, sebbene abbiano di fatto individuato di volta in volta, con metodologie di analisi differenti, le funzioni di costo più rappresentative per la casistica esaminata (frequentemente di forma trans-logaritmica). I risultati a cui sono giunti sono spesso controversi.

Nella disamina della letteratura esistente, si citano brevemente gli studi considerati più attinenti all'argomento qui trattato. Tra questi figura l'analisi di FILIPPINI et al. (2001) [7] che, a partire dai dati forniti dalle aziende di trasporto svizzere, sono giunti alla definizione di una funzione di costo translogaritmica stimata con i minimi quadrati generalizzati (GLS) che, tra le variabili considerate, tiene conto dei fattori di input (lavoro, capitale, carburante) e di due variabili *dummy* indicanti le caratteristiche territoriali e di rete: DEN (densità di fermate) e REGIO (orografia). I risultati ottenuti consistono in costi unitari di produzione in funzione dei valori assunti dalle due variabili *dummy* e della produzione in bus-km, utilizzabili in una politica di *yardstick*<sup>(1)</sup> così come specificato dall'autore.

Completamente differente nell'approccio e nello scopo è lo studio di DALEN e GOMEZ LOBO (2002) [6] che, basandosi sui dati derivanti da compagnie di trasporto norvege-

*cost values, unrelated to the company's features, since it does not refer directly to historical costs actually incurred. Furthermore, it is often difficult for some business activities, shaping the technological link between input and output, and this entails the risk of discretionary choice of inputs to take into account. The "top down" models make use instead of assumptions and statistical analyzes that infer the cost on the basis of information, cross-section or aggregated, partly from transport companies operating in the area which efficiency, however, is not demonstrated, and so these models can assess a cost that is not always of maximum efficiency. These models present the real risk to estimate a cost reflecting the specific characteristics of the investigated sample companies rather than the specific context in which the service is produced. Then there are also hybrids models that have intermediate characteristics between the two types named.*

*Many of the analyzed studies have been aimed to verify possible economies of scale, density and scope, more than the construction and validation of models of production unit cost, although they have in fact identified from time to time, by different analysis methodologies, the most representative cost functions (frequently of translog form) for the cases examined. The results that have come are often controversial.*

*In the existing literature, the more relevant studies on the topic here discussed are cited briefly. Among these, we place the analysis of FILIPPINI et al. (2001) [7] that, from data provided by Swiss transport companies, have come to define a translog cost function estimated by generalized least squares (GLS). This function takes into account, among the variables, the input factors (labor, capital, fuel) and two dummy variables indicating the territory and network features: DEN (density of stops) and REGIO (orography). So the authors obtained unit production costs, related to the values of the two dummy variables and the bus-km production variables, that can be used in a political yardstick<sup>(1)</sup> as specified by the same authors.*

*Quite different approach and purpose has the study of DALEN and GOMEZ LOBO (2002) [6], based on data from the Norwegian transport companies. It attempted to assess the costs but not the efficiency of companies over time by analyzing a term present in the translog cost function,  $\mu_{imp}$ , indicating the specific ineffectiveness of the company. Using a multinomial logit model, the authors found the absence of relationship between business efficiency and the transition to an incentive contract (subsidy cap or standard cost) and this result contradicts much of the existing literature.*

<sup>(1)</sup> *Yardstick* è un termine inglese ed è inteso come "parametro di valutazione", riferendosi all'analisi comparata di imprese operanti nello stesso settore (nel caso in esame, TPL su gomma), finalizzata alla riduzione delle asimmetrie informative tra regolatore e soggetto regolato e allo stimolo della concorrenza in mercati con alto grado di concentrazione.

<sup>(1)</sup> *Yardstick* is a reference value in the comparative analysis of companies operating in the same sector (in this case, road local public transit) meant to reduce the information asymmetry between regulator and regulated subject and stimulate competition in markets with high level of concentration.

## POLITICA E ECONOMIA

si, hanno cercato di valutare non i costi ma l'efficienza economica delle aziende nel tempo attraverso l'analisi di un termine presente nella funzione di costo translogaritmica,  $\mu_{\text{int}}$ , indicante l'inefficienza economica specifica dell'azienda. Tramite un modello logit multinomiale, gli autori hanno scoperto la mancata relazione tra l'efficienza economica aziendale e il passaggio ad un contratto incentivante (*subsidy cap* o modello a costo standard) e questo risultato contraddice gran parte della letteratura preesistente.

Il lavoro di CUBUCKU (2006) [5] ha avuto invece come base empirica una *cross-section* di dati provenienti dalle agenzie di trasporto statunitensi. Nella funzione di costo, anche in questo caso translogaritmica, sono state considerate tre variabili geografiche e di rete:  $H_{\text{Slope}}$ , indicante la percentuale di terreno pianeggiante,  $H_{\text{Density}}$ , la densità di popolazione in [persone/miglia quadrate], e  $H_{\text{Street}}$  che rappresenta la lunghezza media del tronco di strada in miglia. I risultati a cui l'autore è giunto sono dei grafici che mostrano la variazione dell'elasticità di costo del prodotto, reciproco dell'economia di scala, per tre diverse categorie di valori delle tre variabili suddette (minimo, medio e massimo) prese contemporaneamente o separatamente, in funzione dello stesso output espresso in passeggeri-miglia. In questo modo l'autore ha ricavato, nei vari casi, le soglie di produzione oltre le quali il costo unitario dell'offerta aumenta.

Un più recente e interessante studio di ABRATE et al. (2014) [1], che ha riguardato aziende di trasporto italiane, ha fatto uso di una funzione di costo diversa da quelle dei precedenti lavori poiché si tratta di una specificazione multi-prodotto composita, proposta da PULLEY e BRAUNSTEIN [8] nel 1992. Infatti, lo scopo di questo studio è stato quello di trovare potenziali economie di scala e di scopo nell'eventualità di fusioni tra le varie compagnie di trasporto o nel caso di integrazioni orizzontali tra le tre tipologie di servizi considerate (urbani, extraurbani e a noleggio). I risultati trovati hanno evidenziato che le aziende integrate di piccole dimensioni, che operano nello stesso ambiente, presentano economie di scala e di scopo al contrario delle aziende integrate di grandi dimensioni, che presentano evidenti diseconomie.

Il modello ASSTRA<sup>(2)</sup> (2014) [3] è di tipo "bottom up", poiché ricostruisce analiticamente il costo attraverso l'analisi dei processi che avvengono nelle varie aree aziendali. Sulla base dei dati raccolti dalle aziende consociate mediante questionari preimpostati, si sono formulati opportuni postulati, considerando due scenari differenti. Il modello ha permesso di definire delle formule per il calcolo del costo standard e del ricavo standard e, per differenza, anche del corrispettivo standard<sup>3</sup>, ponendolo in funzione

<sup>(2)</sup> ASSTRA sta per "Associazione Trasporti" ed è una associazione nazionale delle aziende di trasporto pubblico locale in Italia, sia di proprietà degli enti locali, sia private.

<sup>(3)</sup> Il corrispettivo standard corrisponde alla differenza tra il costo standard e il ricavo standard presunto, nel caso di contratti "net-cost", mentre è pari all'intero costo standard nel caso di contratti "gross cost".

*The work of CUBUCKU (2006) [5] had instead as empirical basis cross-section data from the US transit agencies. In the cost function, also in this case translog, were considered three geographic and network variables:  $H_{\text{Slope}}$  indicating the percentage of flat land,  $H_{\text{Density}}$  the population density in [persons / square miles], and  $H_{\text{Street}}$  which represents the average length of the road trunk in miles. The results that this author has come are graphs showing the change in the elasticity of the product cost, mutual of the scale economy, for three different categories of values of the above three variables (minimum, mean and maximum) taken simultaneously or separately, according to the same output expressed in passenger-miles. In this way, the author obtained, in the various cases, the production thresholds beyond which the unit cost of the transport supply increases.*

*An interesting and more recent study of ABRATE et al. (2014) [1], which covered the Italian transit companies, used a cost function different from those of previous works, namely it is a multi-composite product specification, given by PULLEY and BRAUNSTEIN [8] in 1992. In fact, the purpose of this study was to find potential economies of scale and scope in the event of mergers between the various transport companies, or in the case of horizontal integrations between the three types of services considered (urban, intercity and rental). The results found showed that merged small companies working in the same environment have scale and scope economies, unlike merged large companies which have clear diseconomies.*

*The ASSTRA<sup>(2)</sup> model (2014) [3] is a "bottom up" type, as analytically reconstructs the cost through the analysis of the processes occurring in the various business areas. Based on data collected from associated companies through preset questionnaires, they have formulated appropriate postulates, considering two different scenarios. The model allowed to build formulas to calculate the standard cost and revenue and, by difference, also of the standard compensation<sup>(3)</sup>, placing this latter in function of the only cost driver, that is the commercial speed in km/h.*

*The ANAV<sup>(4)</sup> model [2] has a macro-analytic approach that relates the cost of the service by both exogenous and endogenous parameters of the company, for which two end values (best and worst) have been defined, by suitable analysis performed on associate companies operating the road local public service.*

*The model of AVENALI et al. (2014) [4], which is basically an evolution of the ANAV model, is a "top down" one,*

<sup>(2)</sup> ASSTRA stands for "Transport Association" and is an association of local public transport companies in Italy, both owned by local authorities, and private.

<sup>(3)</sup> The standard compensation corresponds to the difference between the standard cost and estimated standard revenue, in the case of net-cost contracts, while it is equal to the standard cost in the case of gross-cost contracts

<sup>(4)</sup> ANAV stands for "National Association of Passenger Transport" and is an Italian association of bus passengers transport companies.



## POLITICA E ECONOMIA

dell'unico driver di costo: la velocità commerciale media in km/h.

Il modello ANAV<sup>(4)</sup> [2] ha un approccio macroanalitico che mette in relazione il costo del servizio con parametri sia esogeni che endogeni dell'azienda, per i quali sono stati definiti due valori di estrema (best e worst), attraverso opportune analisi effettuate su aziende consociate che gestiscono il servizio di TPL su gomma.

Il modello di AVENALI et al. (2014) [4], che è praticamente una evoluzione del modello ANAV, è di tipo "top down", seppur con la presenza di vincoli normativi tipici dei modelli ingegneristici. Infatti esso parte da alcune assunzioni fatte sulle variabili scelte (produzione chilometrica, velocità commerciale media, grado di ammortamento del parco rotabile) sulla base delle indicazioni provenienti dai dati raccolti tramite questionario e, mediante procedimenti di inferenza statistica dei dati a disposizione, giunge a determinare il costo standard chilometrico del servizio.

### 3. La metodologia per la costruzione del modello

#### 3.1. I parametri considerati

Il modello proposto è finalizzato a stimare un costo di massima efficienza che può essere di riferimento sia come prezzo a base d'asta nelle procedure di gara per l'assegnazione dei servizi di trasporto pubblico automobilistico (approccio micro), sia in una valutazione del fabbisogno finanziario di ciascun ente preposto alla organizzazione degli stessi servizi (approccio macro). Il modello in questione è del tipo "top down" ma supera il limite tipico di questo approccio, accennato in precedenza, rappresentato dal rischio di stimare costi non efficienti. Infatti esso è costruito e calibrato attraverso dati tratti da servizi di trasporto pubblico locale prodotti in condizioni di massima efficienza globale accertata. Detta condizione è conseguenza del fatto che il data base costruito comprende soltanto bandi di gara per l'affidamento di servizi di trasporto che sono risultati vinti con ribassi d'asta insignificanti a conferma che i costi posti a base di gara ed utilizzati nella calibrazione del modello sono di massima efficienza.

Il modello si basa su tre indicatori di efficienza successivamente inglobati in un unico indicatore complessivo e si avvale di altri parametri noti.

Il rapporto tra le ore di guida e le ore di servizio giornaliero<sup>(5)</sup> prodotte da ciascun autista può essere definito

<sup>(4)</sup> ANAV sta per "Associazione Nazionale Autotrasporto Viaggiatori" ed è una associazione di imprese di trasporto passeggeri su autobus.

<sup>(5)</sup> Le ore di guida giornaliero dipendono dal programma di esercizio, stabilito generalmente dalle stazioni appaltanti, dal tipo di servizio e dalle ore di servizio giornaliero, stabilite nei contratti aziendali o nei CCNL. Le ore di guida sono solo una parte delle ore di servizio giornaliero prestate dall'autista, poiché quest'ultime comprendono anche i tempi di attesa al capolinea tra una corsa e la successiva.

*even with the presence of local regulatory constraints of engineering models. In fact it is based on some assumptions about the choices variables (mileage production, commercial speed, degree of depreciation of the fleet) based on indications from the data collected through questionnaire and, by methods of statistical inference of the data available, it comes to assess the standard cost per kilometer of service.*

### 3. The methodology to build the model

#### 3.1. Selected parameters

*This model is designed to estimate a maximum efficiency cost that can be both as reference price in the tender procedures for the award of the road public transit services (micro approach), both in an evaluation of the financial requirements of each local authority managing the same services (macro approach). The model concerned is a "top down" but overcomes the typical limits of this approach, already mentioned above, representing the risk to estimate inefficient costs. In fact it is built and calibrated by data from local public transport services produced in proven maximum efficiency conditions. This circumstance is a result of the fact that the built database includes only calls for tenders for the award of transport services that were won with insignificant auction markdowns confirming that costs based auction and used in the model calibration are of maximum efficiency.*

*The model is based on three performance indicators, later incorporated into one overall indicator, and makes use of other known parameters.*

*The relationship between the driving hours and the daily hours of service<sup>(5)</sup> produced by each driver can be defined as the average utilization coefficient of drivers  $U_A$ . It is an exogenous indicator characterizing the efficiency of the operating program, on which the transport company cannot generally make change:*

$$U_A = \frac{H_G}{H_L} \quad (1)$$

*where  $H_G$  is the number of driving hours and  $H_L$  is the number of working hours for each driver provided in the average daily or in one year.*

*A second index, to be instead obtained from the business data, is the impact of the drivers cost on the work total cost  $I_{AP}$ , that is the ratio of the annual cost of only drivers ( $C_A$ ), and the annual cost of all staff including drivers ( $C_p$ ):*

<sup>(5)</sup> *The daily hours of service depend on the operating program, generally established from managing authorities, type of service and hours of daily service, defined in business contracts or in National Collective Labour Contracts. The daily hours are only some of the hours of the daily service provided by the driver, as the latter include waiting times at the terminus, between the rides.*

## POLITICA E ECONOMIA

come *coefficiente di utilizzazione medio degli autisti*  $U_A$ . Si tratta di un indicatore esogeno, su cui l'azienda di trasporto non può generalmente intervenire, che caratterizza l'efficienza del programma di esercizio:

$$U_A = \frac{H_G}{H_L} \quad (1)$$

dove  $H_G$  è il numero di ore di guida e  $H_L$  il numero di ore di lavoro prestate in un giorno medio per ciascun autista.

Un secondo indice determinabile invece dai dati aziendali è l'*incidenza del costo degli autisti sul costo complessivo del lavoro*  $I_{AP}$ , dato dal rapporto tra il costo annuale dei soli autisti ( $C_A$ ), e il costo annuale di tutto il personale compresi gli autisti ( $C_P$ ):

$$I_{AP} = \frac{C_A}{C_P} \quad (2)$$

Questo indice misura l'efficienza aziendale in modo indiretto, in quanto l'azienda è tanto più economicamente efficiente quanto meno spende per le voci di costo diverse dal personale di guida e dal materiale rotabile, e quindi quanto più l'indicatore si avvicina all'unità. Si tratta di un indicatore endogeno su cui l'azienda ha possibilità di manovra, tuttavia non è esente da scelte extra-aziendali dovute a particolari esigenze definite nel contratto di servizio.

Un altro indice endogeno che misura l'efficienza aziendale, e in parte anche del programma di esercizio svolto, è l'*incidenza del costo totale del lavoro sul costo complessivo del servizio*  $I_{PS}$ , dato dal rapporto tra il costo annuo di tutto il personale ( $C_P$ ) e il costo annuo del servizio ( $C_S$ ):

$$I_{PS} = \frac{C_P}{C_S} \quad (3)$$

Viste le definizioni date per i tre indici, si può comprendere che la determinazione del costo standard si dovrebbe basare su valori di  $U_A$  tratti da servizi realmente esercitati e su valori di  $I_{AP}$  e  $I_{PS}$  teorici, tipici di una azienda massimamente efficiente, oppure ricavati da aziende la cui efficienza sia effettivamente comprovata e non supposta a priori come invece avviene per i modelli ASSTRA e ANAV.

Partendo dalla percorrenza annua sviluppata dal servizio in bus-km ( $P$ ), desumibile in genere dal bando di gara o dal programma di esercizio, si può definire il *costo chilometrico del servizio*  $C_{SK}$  come:

$$C_{SK} = \frac{C_S}{P} \quad (4)$$

A partire dalla (2) e dalla (3), si ottengono rispettivamente  $C_P$  e  $C_S$  tramite le formule inverse:

$$I_{AP} = \frac{C_A}{C_P} \quad (2)$$

*This index measures business efficiency in an indirect way, because the company is more efficient as it is less expensive for the cost items different from the drivers and the buses ones and then as more the indicator approaches to unity. It is an endogenous indicator on which the company has room for maneuver, however, it is not free from off-farm choices due to special requirements defined in the service contract.*

*Another endogenous index measuring business efficiency, and also partially the operating program, is the incidence of total labor cost on the overall cost of the service  $I_{PS}$ , that is the ratio between the annual cost of all staff ( $C_P$ ) and the annual cost of service ( $C_S$ ):*

$$I_{PS} = \frac{C_P}{C_S} \quad (3)$$

*Given the three indexes, we can understand that the determination of the standard cost should be based on  $U_A$  values get form really produced services and on  $I_{AP}$  and  $I_{PS}$  theoretical values, typical of a maximally efficient company, or obtained by companies whose efficiency is actually proven and not supposed a priori as is the case of ASSTRA and ANAV models.*

*Starting from the annual mileage produced by the service, in bus-km ( $P$ ), generally inferable from the tender notice or the operating program, we can define the cost per kilometer of the service  $C_{SK}$  as:*

$$C_{SK} = \frac{C_S}{P} \quad (4)$$

*From (2) and from (3) we obtain respectively  $C_P$  and  $C_S$  through the reverse formulas:*

$$C_P = \frac{C_A}{I_{AP}} \quad (5)$$

$$C_S = \frac{C_P}{I_{PS}} \quad (6)$$

*Substituting (5) into (6) and this in (4), we obtain the formula of the cost per-kilometer reworked in the following way:*

$$C_{SK} = \frac{C_A}{I_{AP} \cdot I_{PS} \cdot P} \quad (7)$$

*The annual cost of the drivers ( $C_A$ ) is the product of the number of needed drivers and the average annual cost of each driver ( $C_{AU}$ ):*

## POLITICA E ECONOMIA

$$C_P = \frac{C_A}{I_{AP}} \quad (5)$$

$$C_S = \frac{C_P}{I_{PS}} \quad (6)$$

Sostituendo la (5) nella (6) e questa nella (4), si ottiene la formula del costo chilometrico rielaborata nel seguente modo:

$$C_{SK} = \frac{C_A}{I_{AP} \cdot I_{PS} \cdot P} \quad (7)$$

Il costo annuo degli autisti ( $C_A$ ) si può calcolare come il prodotto tra il numero di autisti necessario e il costo medio annuo di ciascun autista ( $C_{AU}$ ):

$$C_A = \text{num autisti} \cdot C_{AU} \quad (8)$$

Il numero degli autisti necessario è ricavabile come rapporto tra le ore di servizio annue complessivamente rese al pubblico ( $H_{ST}$ ) e le ore di guida annue di ciascun autista:

$$\text{num autisti} = \frac{H_{ST}}{\text{ore guida annue x autista}} \quad (9)$$

Le ore di guida annue di ciascun autista sono date dal prodotto delle giornate effettivamente lavorate nell'anno ( $N_g$ ) per le ore di lavoro giornaliero di ciascun autista ( $H_L$ ) e per il coefficiente di utilizzazione medio degli autisti ( $U_A$ ):

$$\text{ore guida annue x autista} = N_g \cdot H_L \cdot U_A \quad (10)$$

Sostituendo la (10) nella (9), per poi sostituirla nella (8) e questa nella (7), si ottiene la nuova formula rielaborata per il costo chilometrico:

$$C_{SK} = \frac{H_{ST} \cdot C_{AU}}{N_g \cdot H_L \cdot U_A \cdot I_{AP} \cdot I_{PS} \cdot P} \quad (11)$$

Al denominatore della (11), compare il prodotto dei tre indici menzionati in precedenza ( $U_A$ ,  $I_{AP}$  e  $I_{PS}$ ) che si può assumere come un nuovo indicatore di efficienza globale  $E_G$ :

$$E_G = U_A \cdot I_{AP} \cdot I_{PS} \quad (12)$$

quindi la relazione (11) diventa:

$$C_{SK} = \frac{H_{ST} \cdot C_{AU}}{N_g \cdot P \cdot H_L \cdot E_G} \quad (13)$$

La relazione (13) così definita permette di valutare il costo chilometrico standard unitario di un dato servizio, note

$$C_A = \text{number of drivers} \cdot C_{AU} \quad (8)$$

The number of needed drivers is obtainable as the ratio between the total annual hours of service provided to the people ( $H_{ST}$ ) and annual hours of driving for each driver:

$$\text{number of drivers} = \frac{H_{ST}}{\text{yearly driving hours for eac driver}} \quad (9)$$

The hours of annual driving of each driver is given by the product of actual worked days in the year ( $N_g$ ) for daily work-shift of each driver ( $H_L$ ) and the coefficient of drivers utilization ( $U_A$ ):

$$\text{yearly driving hours for eac driver} = N_g \cdot H_L \cdot U_A \quad (10)$$

Substituting (10) into (9), and then replacing it in (8) and this last in (7), we obtain the new revised formula for the cost per kilometer:

$$C_{SK} = \frac{H_{ST} \cdot C_{AU}}{N_g \cdot H_L \cdot U_A \cdot I_{AP} \cdot I_{PS} \cdot P} \quad (11)$$

The denominator of (11) shows the product of the three previously mentioned indexes ( $U_A$ ,  $I_{AP}$  and  $I_{PS}$ ) which can assume as a new global efficiency indicator  $E_G$ :

$$E_G = U_A \cdot I_{AP} \cdot I_{PS} \quad (12)$$

then the relation (11) becomes:

$$C_{SK} = \frac{H_{ST} \cdot C_{AU}}{N_g \cdot P \cdot H_L \cdot E_G} \quad (13)$$

The so defined relation (13) allows to evaluate the standard unit cost of a given service, known the annual mileage and the global efficiency indicator. From its inverse formula, we precisely get  $E_G$ :

$$E_G = \frac{H_{ST} \cdot C_{AU}}{N_g \cdot P \cdot H_L \cdot C_{SK}} \quad (14)$$

Hence the (14) expresses the global efficiency indicator depending on the standard cost elaborated on a specific category of trusted services and other known parameters.

In summary, using a specially built database including transport services whose costs are of maximum efficiency, we calibrated a model that provides the  $E_G$  value as a function only of the service characteristics that are note to all.

The model can be used to calculate the global efficiency index  $E_G$  for a service for which we know operating characteristics and, by inserting this value in (13), we can get the estimated standard cost  $C_{SK}$  of this service.

## POLITICA E ECONOMIA

la percorrenza annua dello stesso e l'indicatore di efficienza globale. Dalla sua forma inversa, si ricava proprio  $E_G$ :

$$E_G = \frac{H_{ST} \cdot C_{AU}}{N_g \cdot P \cdot H_L \cdot C_{SK}} \quad (14)$$

Quindi la (14) esprime l'indicatore di efficienza globale in funzione del costo standard elaborato su una determinata categoria di servizi messi a gara e di altri parametri noti.

In sintesi, avvalendosi di un database appositamente costruito con servizi di trasporto i cui costi sono di massima efficienza, si è messo a punto un modello che fornisce il valore di  $E_G$  in funzione delle sole caratteristiche del servizio note a tutti.

Il modello può essere utilizzato per calcolare l'indicatore di efficienza globale  $E_G$  per un servizio di cui sono note le caratteristiche di esercizio e, inserendo tale valore nella (13) è possibile ottenere il costo standard  $C_{SK}$  stimato per tale servizio.

Per costruire e calibrare il modello che esprime  $E_G$ , si è provveduto a mettere in relazione, attraverso regressioni, il valore di  $E_G$ , calcolato con la (14) per ciascun bando compreso nel database, con i parametri noti che caratterizzano il programma di esercizio.

### 3.2. I valori utilizzati

Il data-base costruito raccoglie i dati relativi ad un certo numero di gare per l'appalto dei servizi di trasporto pubblico locale extraurbano espletate recentemente in Italia ed è caratterizzato dalle specificità di seguito richiamate.

- 1) Il costo posto a base di gara per ogni servizio in appalto rappresenta il costo di massima efficienza per lo specifico servizio dal momento che le gare sono state vinte con ribassi d'asta insignificanti. Pertanto il costo del servizio corrisponde al corrispettivo riconosciuto all'azienda, in presenza di contratti "gross-cost", ed alla somma del corrispettivo e dei ricavi da traffico, in contratti "net-cost".
- 2) I corrispettivi riconosciuti alle aziende affidatarie ed i costi dei servizi posti a base di gara (entrambi al netto dell'IVA), sono stati tutti aggiornati allo stesso orizzonte temporale, cioè l'anno 2014, in misura del tasso di inflazione ISTAT-FOI<sup>(6)</sup> o con un tasso diverso se previsto dal capitolato di gara.
- 3) Le aziende producono in condizioni tali da mantenere i coefficienti  $I_{AP}$  e  $I_{PS}$  ai livelli di massima efficienza, per via del meccanismo selettivo rappresentato dalla

<sup>(6)</sup> L'ISTAT è l'istituto nazionale di statistica che, tra i vari compiti, compila e pubblica le tabelle dell'indice FOI, indice di inflazione dei prezzi al consumo per le famiglie di operai e impiegati.

To build and calibrate the model that expresses  $E_G$ , we put in relation, through regressions, the  $E_G$  value, calculated with (14) for each tender notice included in the database, with known parameters characterizing the operating program.

### 3.2. Adopted values

The built database collects data for a number of calls for tenders, recently completed in Italy, for the award of local suburban public transit services, and it is characterized by below mentioned specific features.

- 1) The cost based auction for each contracted service is the most efficient cost for the specific service because the auctions were won with insignificant markdowns. Therefore the cost of the service corresponds to the amount to the company, in the presence of gross-cost contracts, and to the sum of the amount to the company and the fee and traffic revenues, in net-cost contracts.
- 2) Amounts to the companies and the costs based auction for the services (both excluding VAT), have all been updated at the same time horizon, i.e. the year 2014, by the inflation rate ISTAT-FOI<sup>(6)</sup> or by a different rate if stated in the tender document.
- 3) The companies produce in conditions such as to maintain the  $I_{AP}$  and  $I_{PS}$  coefficients at maximum efficiency, because of the selective mechanism represented by public tender, therefore the global efficiency indicator  $E_G$  differs from one contract to another only for the different characteristics of the service operating program.
- 4) The contracts proposed in auction does not involve charges for the company beyond those generally required, like sales and verification of tickets, or additional revenues than those arising from the sale of tickets, in the case of net cost contracts. Any revenues from the commercial exploitation of movable and immovable property are considered negligible since the considered services are only intercity ones.

The information collected for each call for tender were as follows:

- call for tender identification code (IGC);
- year in which the call was issued;
- type of service (urban, suburban, intercity);
- type of contract (net or gross cost);
- amount to the company per-kilometer (updated to 2014);
- cost per kilometer (updated to 2014);
- percentage of costs covered by traffic revenues.

<sup>(6)</sup> ISTAT: Italian Statistics National Institute; it fills out and publishes table of the FOI index; FOI-index is the monetary inflation index of consumer price for workers and employees; VAT is the Value Added Tax.



## POLITICA E ECONOMIA

gara pubblica, pertanto l'indicatore di efficienza globale  $E_G$  differisce da un appalto all'altro solo per le differenti caratteristiche del programma di esercizio.

- 4) I contratti proposti in appalto non prevedono oneri accessori per la ditta appaltatrice oltre quelli generalmente richiesti di vendita e verifica dei titoli di viaggio, né ricavi accessori oltre quelli derivanti dalla vendita dei titoli di viaggio, nel caso di contratti *net cost*; gli eventuali ricavi derivanti dalla valorizzazione commerciale di beni mobili e immobili vengono considerati trascurabili trattandosi di servizi solo extraurbani.

Le informazioni raccolte per ciascun bando di gara sono state le seguenti:

- Codice Identificativo di Gara (CIG);
- anno in cui è stato emanato il bando;
- tipologia di servizio (urbano, suburbano, extraurbano);
- tipologia di contratto (*net cost* o *gross cost*);
- il corrispettivo chilometrico (attualizzato al 2014);
- il costo chilometrico (attualizzato al 2014);
- la percentuale di costo coperta dai ricavi da traffico.

In aggiunta a queste informazioni identificative della gara e dei servizi, sono state inserite nel database le caratteristiche del programma di esercizio, per ciascuna tipologia di servizio, sintetizzate nei seguenti parametri, appositamente calcolati, da cui dipende  $E_G$ :

- percorrenza annuale del servizio ( $P$ ), indicata sul bando di gara in bus-km/anno. Il suo legame con  $E_G$  è ambiguo, dato che in letteratura si è giunti spesso a risultati contrastanti. Gli studi più recenti, tra cui quello di AVENALI et al. [4], sembrano propendere per un andamento ad "U" della curva del costo unitario all'aumentare della produzione, per cui oltre una certa soglia di produzione, variabile a seconda degli studi, non si hanno più economie di scala. Il segno che ( $P$ ) assumerà nel modello di costo sarà quindi conseguenza dei contesti e dei valori presi in considerazione nel database per questo parametro.
- velocità commerciale media del servizio ( $V_C$ ) in km/h, calcolata come rapporto tra la percorrenza chilometrica annua e la durata temporale delle corse. Numerosi studi internazionali hanno verificato che con l'aumento della velocità commerciale si ha una riduzione del costo di esercizio e quindi un aumento dell'indice  $E_G$ , specialmente in ambito urbano, per effetto di una maggiore produttività degli autisti in termini di ore di guida rapportate alle ore di servizio ( $U_A$ ). Ci si aspetta quindi un segno positivo per il relativo coefficiente di regressione nella formula.
- durata temporale media delle corse ( $T_C$ ) in ore/corsa, che è la media dei tempi di viaggio, da un estremo all'altro, di tutte le corse realizzate, ottenuta dividendo le ore di servizio annue rese al pubblico per il numero di corse effettuate in un anno. I percorsi di lunga durata,

*In addition to these information identifying the auctions and services, we have included in the database the features of the operating program for each type of service, summarized in the following specially calculated parameters affecting  $E_G$ :*

- *annual mileage ( $P$ ), in bus-km/year, reported in the call for tenders. His connection with  $E_G$  is ambiguous, since the literature has often come to conflicting results. More recent studies, including that of Avenali et al. [4], seem to favor an "U" trend for the unit cost of production as a function of the annual mileage, for which more than a certain threshold of mileage, variable depending on the studies, there has been no economies of scale. The sign of ( $P$ ) in the model will then be a consequence of contexts and values considered in the database for this parameter;*
- *average commercial speed of service ( $V_C$ ) in km/h, i.e the ratio between the annual mileage and the running time of the rides. Many international studies have observed that the increase in the commercial speed follows a reduction of the operating cost, and then a rise of  $E_G$  index, especially in urban areas. It's due to an higher productivity of the drivers in terms of driving hours related to the service hours ( $U_A$  ratio). Therefore we expect a positive sign for the coefficient of ( $V_C$ ) in the formula of  $E_G$ ;*
- *average travel time of the rides from one terminus ( $T_C$ ), in hours/rides; that is the average travel time, from one terminus to the other one, of all the rides performed in one year. The long routes, typical of the intercity transit services, tends to low values of  $U_A$ , and then of  $E_G$ , because of the lowest weight of the driving hours on the total service hours. Therefore we expect a negative sign for this variable in the formula of  $E_G$ ;*
- *service-density ( $D_S$ ), in km / square km, that is the ratio between the annual mileage and the served area. A larger area to serve results in longer transport lines and then, according to that said about the  $T_C$  variable, a lower value of  $E_G$ . The expected sign for this variable, in the formula of  $E_G$ , is therefore positive (the area appears in the denominator);*
- *intensity of service ( $F_S$ ) indicate how many times the path between the two terminuses is covered in one year, for each line or for all the network; it is obtained by dividing the overall mileage for the sum of the distances of the routes between the terminuses. More rides on the lines result in a rise of the  $U_A$  and then of  $E_G$  index, particularly in urban areas. The expected sign is positive;*
- *peak hour factor of the service ( $P_F$ ) calculated as the ratio between the number of rides in the peak period of an average day and the average number of rides in the moderate period of the same width. The set average day is the 14 October 2015<sup>(7)</sup> and the width of the reference period*

<sup>(7)</sup> We choose, as the analysis day, a school and work weekday, in a period that usually is not affected by extraordinary flows, like the autumn.



## POLITICA E ECONOMIA

tipici delle linee extraurbane, conducono tendenzialmente a valori di  $U_A$ , e quindi di  $E_G$ , più bassi per la minore incidenza delle ore di guida sul totale delle ore di servizio. Ci si aspetta quindi un segno negativo per questa variabile nella specificazione di  $E_G$ .

- densità del servizio ( $D_S$ ) in [km/km quadrati], data dal rapporto tra la percorrenza chilometrica annua e l'area del bacino servito. Una maggiore superficie da servire implica una maggiore lunghezza delle linee di trasporto e quindi, visto quanto detto per la variabile  $T_C$ , un minor valore di  $E_G$ . Il segno previsto nella formula per questa variabile è quindi positivo (la superficie compare al denominatore dell'indicatore).
- intensità del servizio ( $F_S$ ) indica quante volte viene percorsa in un anno la tratta compresa tra i due capilinea, per una singola linea, oppure l'intera rete complessivamente e si ottiene dividendo la percorrenza complessiva per la distanza o per la somma delle distanze comprese tra i capilinea. Con una rete maggiormente utilizzata, si avrebbe un aumento dell'indice  $U_A$  e quindi di  $E_G$ , specialmente in ambiti urbani. Il segno atteso è positivo.
- fattore dell'ora di punta del servizio ( $P_F$ ) calcolato come rapporto tra il numero di corse che si hanno in un giorno medio nel periodo di punta e il numero medio di corse che si hanno in un qualunque altro periodo di morbida ma della stessa ampiezza. Il giorno medio considerato è stato il 14/10/2015<sup>(7)</sup> mentre l'ampiezza dell'intervallo di riferimento è di un'ora per il servizio urbano, di un'ora e mezza per il servizio suburbano e di due ore per quello extraurbano. Con un fattore di punta più alto, si ha bisogno di un maggior numero di autisti, poiché si ha un'alta concentrazione delle corse in un tempo ristretto ed a ciò consegue una maggiore incidenza del costo del personale di guida su quello del personale nel suo complesso  $I_{AP}$  e quindi un aumento di  $E_G$ . Di converso, con un fattore di punta basso si ha una maggiore uniformità nella distribuzione delle corse durante la giornata, quindi una maggiore incidenza  $U_A$  delle ore di guida sul totale delle ore di servizio degli autisti e pertanto ancora un aumento di  $E_G$ . Per queste ragioni, la previsione sul segno che il fattore assumerà nella formula è incerta e dipenderà da quale delle due cause avrà un peso maggiore.

Purtroppo, per vari motivi, spesso i dati rilasciati dalle stazioni appaltanti riguardo alle gare espletate sono risultati incompleti o addirittura inutilizzabili. Di conseguenza, la numerosità del campione risulta ridotta (9 unità<sup>(8)</sup>)

<sup>(7)</sup> Si è voluto scegliere come giorno di analisi, un giorno infrasettimanale scolastico, lavorativo, in un periodo che di solito non risente di flussi straordinari, come quello autunnale.

<sup>(8)</sup> Nel campione sono stati considerati: la Regione Valle d'Aosta (suddivisa in tre sottobacini con i rispettivi lotti assegnati singolarmente: Alta Valle, Centro Valle, Bassa Valle), la Provincia di Cuneo, l'area metropolitana di Torino, la Regione Molise, l'area metropolitana di Bologna e la Provincia di Perugia suddivisa in due sottobacini.

*is one hour for the urban transit, one hour and half for the suburban and two hours for the intercity one. In presence of an higher peak factor, we need a greater number of drivers, since we would have a high concentration of the rides in a short time, which results in a higher incidence of the cost of the drive staff on the total personnel ( $I_{AP}$ ) and then an higher  $E_G$  value. Conversely, a low peak factor, that is a more uniform distribution of the rides in the day, results in a greater incidence of driving hours on the total hours of service ( $U_A$ ) and therefore an higher  $E_G$  value again. For these reasons, the prediction on the sign that this variable in the formula is uncertain and will depend on which of the two cases will have a major weight.*

*Unfortunately, for various reasons, often the data released to us by the contracting authorities with regard to the carried-out call for tenders were incomplete or even unusable. Consequently, the sample size is reduced (9 units<sup>(8)</sup>) and related to the only intercity and suburban transit services, however the methodology developed is generally valid. The availability of data about public tender of local public transit services will be filled shortly in Italy, as soon as the (Italian) National Observatory of the Public Transport Policies (established at the Infrastructures and Transports Ministry by the article 1, paragraph 300 of the Law 244/2007) will have built and made available its data-base.*

### 3.3. The model formulations

*We have conducted a multiple linear regression based on data collected into the database, having  $E_G$  as dependent variable in relation with the previous six explanatory variables of operating program (full model). In table 1 there are the empirical values of these parameters and the  $E_G$  indicator calculated for each lot of the tender competition.*

*You may notice that the  $E_G$  values totaled on average around 0.33, with the maximum tip of 0.431 for Turin and the slightest hint of 0.275 for the first sub-basin of Perugia. You can take also other useful information: the mileage  $P$ , as well as the density  $D_s$  and the intensity  $F_s$  of the service are greatly variable and this was to be expected given the different territorial features of the various areas as well as the variability of user needs to satisfy (more rides in productive areas, etc.); the  $V_c$  range is between 30 and 40 km/h, this also indicative of the serviced area;  $T_c$  lies on average at around 0.70 while the peak hour factor  $P_F$  is substantially uniform, excepting for two tips (Bologna and Cuneo).*

*In tables 2, 3 e 4 are explained the results of the multiple regression analysis. Table 2 shows that the R-Squared coefficient of multiple determination of complete regression*

<sup>(8)</sup> In the sample we have considered the following areas: Aosta Valley (subdivided into three sub-areas singularly entrusted: Lower Valley, Center Valley, High Valley), the suburbs of Cuneo, the metropolitan area of Turin, the Molise region, the metropolitan area of Bologna, the province of Perugia subdivided in two sub-areas.

## POLITICA E ECONOMIA

con analisi fatte solo per i servizi di TPL extraurbano, tuttavia la metodologia messa a punto ha validità generale. Peraltro il problema della disponibilità dei dati sugli appalti dei servizi di trasporto pubblico locale sarà a breve superato non appena l'Osservatorio Nazionale sulle Politiche del Trasporto Pubblico Locale (istituito, presso il Ministero delle Infrastrutture e Trasporti, dall'art.1, comma 300 della L.244/07) avrà costruito e reso disponibile il suo data-base.

### 3.3. Le formulazioni del modello

Sulla base dei dati raccolti, è stata condotta una analisi di regressione multipla lineare avente come variabile dipendente  $E_G$  e come variabili esplicative i sei parametri del programma di esercizio citati poc'anzi (*modello completo*).

Nella tabella 1 sono indicati i valori empirici di questi parametri e l'indicatore  $E_G$  calcolati per ciascun lotto di gara bandito. Si può notare che i valori di  $E_G$  si attestano mediamente intorno a 0,33, con la punta massima di 0,431 per Torino e la punta minima di 0,275 per il sottobacino 1 di Perugia. Si possono trarre inoltre altre utili informazioni: la percorrenza  $P$  così come la densità di servizio  $D_s$  e l'intensità di servizio  $F_s$  sono notevolmente variabili e ciò era prevedibile data la diversità di contesto territoriale dei vari ambiti oltre che la variabilità delle esigenze degli utenti da soddisfare (maggiori corse in contesti produttivi, ecc.), l'intervallo di  $V_c$  è compreso tra i 30 e i 40 km/h, anche questo indicativo del contesto servito,  $T_c$  si mantiene mediamente intorno ai 0,70 mentre il fattore dell'ora di punta  $P_F$  è sostanzialmente uniforme, eccettuando due picchi (Bologna e Cuneo).

Nelle tabelle 2, 3 e 4 sono esposti invece i risultati dell'analisi di regressione. Dalla tabella 2 si vede che il *coefficiente di determinazione multiplo*  $R^2$  del modello di regressione completo è molto alto (0,95), cioè il 95% della variabilità di  $E_G$  viene spiegato, attraverso questo modello, dalle variabili esplicative scelte. Anche il *coefficiente di determinazione multiplo corretto*  $R^2_C$  è abbastanza alto (0,80). Per il *coefficiente di correlazione multiplo*  $R$ , che misura la correlazione lineare tra i valori osservati di  $E_G$  e i suoi valori stimati, il modello dimostra un'alta bontà di adattamento, pari al 97%. La *varianza dei residui*, indicata nella tabella 3 come MQE (Media dei Quadrati degli Errori) è molto bassa, così come sono bassi gli *errori standard* dei coefficienti di regressione, indicando quindi una minima differenza tra valori stimati e valori osservati. A corredo della verifica della bontà del modello, è stata condotta una *analisi dei residui*, indagando su una possibile

*model is very high (0.95), i.e. 95% of the variability of  $E_G$  is explained, through this model, by the set explanatory variables. Also the  $R^2_C$  correct coefficient of multiple determination is quite high (0.80). For the multiple correlation coefficient  $R$ , which measures the linear correlation between the observed values of  $E_G$  and its estimated values, the model demonstrates a high goodness of fit, equal to 97%. The variance of residuals, indicated in table 3 as MQE (Medium of Squares of Errors) is very low, and also standard errors of the regression coefficients are low, thus indicating a minimal difference between the estimated values and the observed ones. In support of the verification of the model's goodness, it was conducted an analysis of residues, investigating a possible relationship between the residuals and the explanatory variable. From the analysis, it was possible to exclude a linear relation, logarithmic or quadratic and then the model would seem appropriate for the provision of  $E_G$ .*

Afterwards, to verify the significance of the model, we conducted an *F-test*, by which we investigated whether there was a significant relationship between the dependent variable and the set of considered explanatory variables. This test requires that you make two hypotheses, one null  $H_0$  and the alternative  $H_1$ :

$$H_0: \beta_1 = \beta_2 = \dots = \beta_6 = 0 \quad (15)$$

$$H_1: \text{at least one of } \beta_i \neq 0 \quad (16)$$

That is, the null hypothesis predicts that all regression coefficients are zero and therefore there is no linear relationship between the  $E_G$  and the explanatory variables, while the alternative hypothesis is that there is a linear relationship between the  $E_G$  and at least one explanatory variable. In order to verify which hypothesis is valid, we used the *F statistic* having a Fisher distribution with  $p$  and  $n-p-1$  degrees of freedom, where  $p$  is the number of explanatory variables in the model and  $n$  is the sample size:

TABELLA 1 - TABLE 1

Valori empirici delle variabili calcolati per ciascun ambito territoriale studiato  
*Empirical values of variables calculated for each area studied*

Ambiti Area	$E_G$	$P$	$V_c$	$T_c$	$D_s$	$F_s$	$P_F$
Alta Valle High Valley	0,299	363.656,6	33,6	0,61	362,9	3.206,7	1,51
Centro Valle Center Valley	0,363	596.116,8	33,2	0,65	624,7	3.909,9	1,05
Bassa Valle Lower Valley	0,381	973.960,0	32,6	0,79	746,4	3.797,1	1,57
Cuneo	0,350	10.641.940	36,83	0,78	1.541,6	3.318,8	2,38
Torino	0,431	3.871.819	29,11	0,63	770,7	4.603,0	1,51
Molise	0,294	11.391.150,4	41,1	0,73	2.566,7	2.188,5	1,76
Bologna	0,311	10.241.569,8	35,58	0,63	2.766,3	1.141,7	3,73
Perugia 1	0,275	5.220.022,4	36,9	0,85	1.262,9	2.278,9	1,85
Perugia 2	0,285	2.264.127	39	0,62	1.028,9	1.958,5	1,76

## POLITICA E ECONOMIA

relazione tra i residui e le variabili esplicative. Dalle analisi, si è potuta escludere una relazione di tipo lineare, logaritmica o quadratica e quindi il modello sembrerebbe adeguato per la previsione di  $E_G$ .

Successivamente, per verificare la significatività del modello, è stato condotto un test F, tramite il quale si è voluto verificare se c'è una relazione significativa tra la variabile dipendente e l'insieme considerato delle variabili esplicative. Questo test prevede che si facciano due ipotesi, una nulla  $H_0$  e l'alternativa  $H_1$ :

$$H_0: \beta_1 = \beta_2 = \dots = \beta_j = 0 \quad (15)$$

$$H_1: \text{almeno un } \beta_j \neq 0 \quad (16)$$

cioè l'ipotesi nulla prevede che tutti i coefficienti di regressione siano nulli e che quindi non vi sia alcuna relazione lineare tra la  $E_G$  e le variabili esplicative, mentre l'ipotesi alternativa prevede che ci sia una relazione lineare tra la  $E_G$  e almeno una variabile esplicativa.

Per poter verificare quale ipotesi sia valida, si utilizza la statistica F avente una distribuzione di Fisher con p ed n-p-1 gradi di libertà, dove p è il numero di variabili esplicative del modello ed n è la numerosità del campione:

$$F = \frac{MQR}{MQE} \quad (17)$$

dove MQR è la Media dei Quadrati della Regressione ed è pari 0,0035 mentre MQE è la Media dei Quadrati degli Errori, nominata prima, e pari a 0,0006, quindi la statistica F sarà pari a 6,29 (tabella 3). Il valore trovato di F va confrontato con  $F_U$  che è il valore critico sulla coda di destra di una distribuzione di Fisher con p e n-p-1 gradi di libertà e con un livello di significatività pari a 0,05, cioè rappresenta il valore della distribuzione F di Fisher per cui la probabilità di trovare valori maggiori o uguali ad esso risulta pari al 5%.

Si avranno quindi una regione di accettazione di  $H_0$  e una di rifiuto:

$$\text{Regione di rifiuto di } H_0: F > F_U \quad (18)$$

$$\text{Regione di accettazione di } H_0: F < F_U \quad (19)$$

Il valore critico  $F_U$  in corrispondenza di 6 e 2 gradi di libertà è pari a 19,3. Dal confronto con F viene fuori che:

$$F < F_U \rightarrow H_0 \text{ va accettata.} \quad (20)$$

Questo significa che non c'è nessuna relazione tra  $E_G$  e le altre variabili esplicative e che quindi il modello studiato nel complesso non risulta significativo, ma ciò non pregiudica la metodologia alla sua base. Infatti è molto probabile che questo risultato sia dovuto alla scarsa numerosità del campione che influisce sul grado di libertà n-p-1 rendendolo troppo basso, tanto è vero che per valori più alti di n si riscontrano valori molto più bassi per il valore critico  $F_U$ .

Dati i valori stimati dei coefficienti della regressione nella tabella 4, l'espressione che permette di valutare  $E_G$  in funzione delle variabili esplicative è la seguente:

$$F = \frac{MQR}{MQE} \quad (17)$$

*In the (17), MQR is the average of the squares regression and equals 0.0035 while MQE is the average of squares of errors, nominated before, and equal to 0.0006; so the F statistic is equal to 6.29 (table 3). The calculated F value must be compared with  $F_U$  which is the critical value in the right tail of a Fisher distribution with p and n-p-1 degrees of freedom and with a significance level equal to 0.05. It represents the value of the Fisher F distribution for which the probability of finding values greater or equal to it is equal to 5%. Then you will have an  $H_0$  acceptance region and a rejection one:*

$$\text{Rejection region of } H_0: F > F_U \quad (18)$$

$$\text{Acceptance region of } H_0: F < F_U \quad (19)$$

*FU critical value, corresponding to 6 and 2 degrees of freedom, is equal to 19.3. By comparison with F it comes out that:*

$$F < F_U \rightarrow H_0 \text{ is to accept.} \quad (20)$$

*This means that there is no relationship between  $E_G$  and other explanatory variables and that the model studied in the complex is not significant, but this will not affect the methodology at its base. In fact it is very likely that this result is due to the low sample size that influences the degree of freedom n-p-1 making it too low. So much so that, for higher values of n are found much lower values for the critical value  $F_U$ .*

TABELLA 2 - TABLE 2

Statistiche della regressione multipla (modello completo)  
Multiple regression statistics (full model)

R multiplo - Multiple R	0,97
$R^2$	0,95
$R^2_c$	0,80
Errore standard - Standard Error	0,02
Osservazioni - Observations	9

TABELLA 3 - TABLE 3

Analisi della varianza (modello completo)  
Analysis of variance (full model)

	gdl dof	SQ SS	MQ MS	F	Signif. F
Regressione Regression	6	0,0208	0,0035	6,29	0,143
Residuo Residual	2	0,0011	0,0006		
Totale Total	8	0,0219			



## POLITICA E ECONOMIA

TABELLA 4 - TABLE 4

Coefficienti stimati della regressione (modello completo)  
*Estimated coefficients of the regression (full model)*

	<i>Coefficienti Coefficients</i>	<i>Errore Standard Standard Error</i>	<i>Stat t</i>
Intercetta <i>Intercept</i>	-0,0949	0,43916	-0,21617838
P	-8,91E-09	8,037E-09	-1,10823702
V <sub>C</sub>	0,0025	0,0082031	0,30577296
TC	-0,0310	0,1040357	-0,29779567
DS	6,11E-05	3,717E-05	1,64227749
F <sub>S</sub>	8,20E-05	3,464E-05	2,36793884
P <sub>F</sub>	0,0449	0,0373433	1,20245019

$$E_G = -0,0949 - 8,91 \cdot 10^{-9} P + 0,0025 V_C - 0,0310 T_C + 6,11 \cdot 10^{-5} D_S + 8,20 \cdot 10^{-5} F_S + 0,0449 P_F \quad (21)$$

Una volta ricavato  $E_G$  dalla (21), è possibile calcolare il costo chilometrico unitario di massima efficienza con la (13).

Nonostante le statistiche di regressione espresse nella tabella 2 siano da considerarsi positive, tuttavia si è potuto accertare, con una successiva verifica, la presenza di variabili collineari attraverso il *Variance inflation factor* (VIF), indicatore che permette di misurare la multicollinearità di ciascuna variabile esplicativa  $X_j$  con le altre, dato dalla seguente espressione:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (22)$$

dove  $R_j^2$  è il coefficiente di determinazione multiplo del modello di regressione nel quale la variabile  $X_j$  dipende dalle altre variabili esplicative. Il valore minimo del  $VIF_j$  è 1 e indica che la variabile  $X_j$  è incorrelata alle altre. Con un valore superiore a 5, la variabile esplicativa  $X_j$  viene considerata altamente correlata alle altre e quindi viene eliminata dal modello. È stato trovato che l'unica variabile con assenza di collinearità è la durata temporale delle corse  $T_C$  ( $VIF = 1,3 < 5$ ) e quindi è l'unica che andrebbe "salvata" mentre le altre andrebbero eliminate, ma è stato verificato che ciò porta ad un modello statisticamente poco significativo.

Questi risultati sono riconducibili, con ogni probabilità, alla numerosità molto ridotta del campione, a cui si potrebbe ovviare con una maggiore disponibilità di dati, ma la sua validità resta confermata dalle statistiche della regressione presentate nella tabella 2.

Successivamente, considerando per alcune o per tutte le variabili esplicative la non linearità (variazione quadratica, cubica, logaritmica, esponenziale), sono stati costruiti modelli alternativi di regressione che potessero fornire valori stimati dell'indice di efficienza globale (e di

Given the estimated coefficients of the regression in table 4, the expression that allows the evaluation of  $E_G$  is as follows:

$$E_G = -0,0949 - 8,91 \cdot 10^{-9} P + 0,0025 V_C - 0,0310 T_C + 6,11 \cdot 10^{-5} D_S + 8,20 \cdot 10^{-5} F_S + 0,0449 P_F \quad (21)$$

Once obtained  $E_G$  from the (21), it's possible to calculate the unit cost per kilometer of maximum efficiency with the (13).

Despite the regression statistics shown in table 2 are to be regarded as positive, however we have verified the presence of collinear variables through the Variance inflation factor (VIF), index to measure the multicollinearity of each explanatory variable  $X_j$  with the other, given by the expression:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (22)$$

Where  $R_j^2$  is the coefficient of multiple determination of the regression model in which the variable  $X_j$  depends from the other explanatory variables. The minimum value of  $VIF_j$  is 1 and indicates that the variable  $X_j$  is uncorrelated to the other. With a value greater than 5, the explanatory variable  $X_j$  is considered highly correlated to the other and then is removed from the model. We found that the only variable with the absence of multicollinearity is the time duration of the rides  $T_C$  ( $VIF = 1.3 < 5$ ) and therefore is the only one that should be "saved" while the others should be eliminated, but it has been verified that this leads to a statistically insignificant model.

These results are attributed to the very small size of the sample, which can be overcome with greater availability of data but its validity remains confirmed by regression statistics presented earlier. Subsequently, considering the nonlinearity for some or all explanatory variables (quadratic, cubic, logarithmic, exponential variation), we have built alternative models of regression that could provide estimates of global efficiency index (and therefore the standard cost per kilometer) altogether closer to those directly derivable from the data available.

Among the different tested models, the one who provided more encouraging results, considers the following variables (alternative model): annual mileage  $P$ , commercial speed of service of 1st and 3rd grade ( $V_C$  e  $V_C^3$ ), intensity of service of 1st and 3rd grade ( $F_S$  e  $F_S^3$ ), peak hour factor of 1st and 3rd grade ( $P_F$  e  $P_F^3$ ). In tables 5, 6, 7 and 8 are exposed the results of the analysis of regression.

Using the values of the coefficients in table 8, we can express the relation between the global efficiency indicator  $E_G$  and the explanatory variables with the following formula:

## POLITICA E ECONOMIA

conseguenza del costo standard chilometrico) complessivamente più vicini a quelli direttamente ricavabili dai dati a disposizione.

Tra i vari modelli sperimentati, quello che ha fornito i risultati più incoraggianti è quello che considera le seguenti variabili esplicative (*modello alternativo*): percorrenza annuale  $P$ , velocità commerciale di grado 1 e 3 ( $V_C$  e  $V_C^3$ ), intensità di servizio di grado 1 e 3 ( $F_S$  e  $F_S^3$ ), fattore dell'ora di punta di grado 1 e 3 ( $P_F$  e  $P_F^3$ ). Nelle tabelle 5, 6, 7 e 8 sono esposti i risultati dell'analisi di regressione.

Utilizzando i valori dei coefficienti della tabella 8, si può esprimere, con la formula che segue, la relazione che intercorre tra l'indice di efficienza globale  $E_G$  e le variabili esplicative considerate:

$$E_G = -0,02627 - 4,543 \cdot 10^{-9} P + 0,00051 V_C + 2,933 \cdot 10^{-6} V_C^3 + 2,457 \cdot 10^{-5} F_S + 2,285 \cdot 10^{-12} F_S^3 + 0,03317 P_F + 0,0015 P_F^3 \quad (23)$$

Dalla tabella 6, si può notare che questo modello ha una elevata bontà di adattamento ai dati a disposizione (99%), dove il 98% della variabilità di  $E_G$  viene spiegato dalle variabili esplicative scelte. La verifica di adattamento ottimale del modello ai dati del database si può vedere dalla tabella 9, dove le differenze tra i valori stimati e quelli effettivi dei costi chilometrici sono minime.

Dal confronto delle statistiche del "modello completo" di regressione (tabella 2), caratterizzato dalla linearità in tutti i suoi parametri, con il "modello alternativo" (tabella 6), che considera invece la non linearità di quasi tutti i parametri considerati, si può vedere che il migliore risulta proprio quest'ultimo, seppur caratterizzato da una collinearità intrinseca dovuta alla presenza ripetuta di diverse variabili ma con diverso grado di potenza. In particolare, l'indicatore che permette di poter fare il confronto statistico tra modelli che hanno variabili diverse, è il coefficiente di determinazione multiplo corretto  $R^2_c$ , legger-

$$E_G = -0,02627 - 4,543 \cdot 10^{-9} P + 0,00051 V_C + 2,933 \cdot 10^{-6} V_C^3 + 2,457 \cdot 10^{-5} F_S + 2,285 \cdot 10^{-12} F_S^3 + 0,03317 P_F + 0,0015 P_F^3 \quad (23)$$

*The results in table 6 shows that this model has a high goodness of fit to the data available (99%), where the 98% of variability of  $E_G$  is explained through the explanatory variables chosen. Verifying optimal adaptation of the model to the data in the database can be seen from the table 9, where the differences between estimates and actual mileage costs are minimal.*

*From a comparison of statistics of the "full model", characterized by the linearity of all its parameters, with the "alternative model", that instead considers the non linearity of almost all the considered parameters, we have seen that the best is precisely this last, although it is characterized by an intrinsic collinearity due to the repeated presence of several variables but with different degree of power. Particularly, the indicator that allows to make the statistical comparison between models having different variables, is the coefficient of correct multiple determination  $R_c^2$ , that is slightly higher for the alternative model. It should be added that all regression coefficients, both for the complete model and for the alternative one, for which the value was predicted as a sign, did not give different results from the expected ones.*

*Later attempts have been made omitting some variables with lower incidence and / or considering a different level of power for the existing variables, to arrive at new formulations that could provide satisfactory results, comparable with the real values.*

*Among the various "partial" models tested, those giving the more acceptable results are two (model 2 and model 3). Model 2, which has an  $R^2$  of 0.98 and a  $R_c^2$  of 0.83, has as explanatory variables the mileage  $P$ , the commercial speed  $V_C$  (at the grades 1 and 2), the intensity of service  $F_S$  (at the grades 1 and 2) and the peak hour factor  $P_F$  (at grades 1 and*

TABELLA 5 - TABLE 5

Valori empirici delle variabili calcolati per ciascun ambito territoriale studiato  
Empirical values of variables calculated for each area studied

Ambiti Area	$E_G$	$P$	$V_C$	$V_C^3$	$F_S$	$F_S^3$	$P_F$	$P_F^3$
Alta Valle High Valley	0,299	363.656,6	33,6	37933,056	3.206,7	3,3E+10	1,51	3,442951
Centro Valle Center Valley	0,363	596.116,8	33,2	36594,368	3.909,9	6,0E+10	1,05	1,157625
Bassa Valle Lower Valley	0,381	973.960,0	32,6	34645,976	3.797,1	5,5E+10	1,57	3,869893
Cuneo	0,350	10.641.940	36,83	49958,01299	3.318,8	3,7E+10	2,38	13,481272
Torino	0,431	3.871.819	29,11	24667,58403	4.603,0	9,8E+10	1,51	3,442951
Molise	0,294	11.391.150,4	41,1	69426,531	2.188,5	1,0E+10	1,76	5,451776
Bologna	0,311	10.241.569,8	35,58	45042,01711	1.141,7	1,5E+09	3,73	51,895117
Perugia 1	0,275	5.220.022,4	36,9	50243,409	2.278,9	1,2E+10	1,85	6,331625
Perugia 2	0,285	2.264.127	39	59319	1.958,5	7,5E+09	1,76	5,451776

## POLITICA E ECONOMIA

mente più grande per il modello alternativo. Va aggiunto che tutti i coefficienti di regressione, sia per il modello completo che per il modello alternativo, per i quali era stato previsto il valore in segno, non hanno dato esiti diversi da quanto ci si aspettava.

Successivamente sono stati fatti dei tentativi, omettendo alcune variabili in base alla loro minore incidenza e/o considerando un diverso grado di potenza delle variabili esistenti, per giungere a nuove formulazioni che potessero fornire dei risultati soddisfacenti e confrontabili con quelli reali, sulla scorta dei dati a disposizione.

Tra i vari “modelli parziali” sperimentati, quelli che sono risultati più accettabili nei risultati trovati sono due (modello 2 e modello 3). Il modello 2, che ha un  $R^2$  pari a 0,98 e un  $R^2_c$  pari a 0,83, ha come variabili esplicative la percorrenza  $P$ , la velocità commerciale  $V_C$  (di grado 1 e 2), l'intensità del servizio  $F_S$  (di grado 1 e 2) e il fattore dell'ora di punta  $P_F$  (di grado 1 e 2), mentre il modello 3, che ha gli stessi valori di  $R^2$  e di  $R^2_c$  del modello precedente, ha le stesse variabili del modello precedente con l'aggiunta della percorrenza chilometrica di secondo grado e con la rimozione della variabile del fattore dell'ora di punta, sempre di secondo grado.

Il modello 2 è caratterizzato dalla seguente relazione che permette di determinare l'indice di efficienza globale  $E_G$ :

$$E_G = 0,0029 - 3,82 \cdot 10^{-9} P + 0,0011 V_C + 0,00014 V_C^2 - 6,30 \cdot 10^{-5} F_S + 2,5 \cdot 10^{-8} F_S^2 + 0,0315 P_F + 0,0037 P_F^2 \quad (24)$$

Per il modello 3, invece, la relazione che permette di valutare  $E_G$  è la seguente:

$$E_G = -0,1641 - 5,04 \cdot 10^{-9} P + 1,42 \cdot 10^{-16} P^2 + 0,013 V_C - 2,93 \cdot 10^{-5} V_C^2 - 8,44 \cdot 10^{-5} F_S + 2,83 \cdot 10^{-8} F_S^2 + 0,043 \cdot P_F \quad (25)$$

I costi chilometrici unitari per i modelli 2 e 3 sono stati ottenuti utilizzando rispettivamente le relazioni (24) e (25) e sostituendole nell'espressione (13).

In aggiunta ai modelli sperimentati, in particolare il “modello completo” espresso dalla (21) e il “modello alternativo” definito dalla (23), nella ricerca della formulazione migliore, si è tentato di trovare una specificazione che risultasse il più possibile semplificata nella forma e nel numero di variabili considerate, senza alcuna possibilità di sconto sulla correttezza del metodo e dei risultati trovati. Seguendo questo ragionamento, sono stati effettuati alcuni tentativi ed è stato individuato il modello più confacente sulla scorta dei dati del database a disposizione. Individuando come unica variabile il rapporto elevato al quadrato tra l'intensità di servizio e la velocità commerciale, si è trovata l'espressione che definisce  $E_G$  in questo modello semplificato attraverso una regressione semplice lineare:

$$E_G = 6 \cdot 10^{-6} \cdot \left( \frac{F_S}{V_C} \right)^2 + 0,2768 \quad (26)$$

TABELLA 6 - TABLE 6

Statistiche della regressione (modello alternativo)  
*Multiple regression statistics (alternative model)*

R multiplo - <i>Multiple R</i>	0,99
$R^2$	0,98
$R^2_c$	0,82
Errore standard - <i>Standard Error</i>	0,02
Osservazioni - <i>Observations</i>	9

TABELLA 7 - TABLE 7

Analisi della varianza (modello alternativo)  
*Analysis of variance (alternative model)*

	gdl <i>dof</i>	SQ <i>SS</i>	MQ <i>MS</i>	F	Signif. F
Regressione <i>Regression</i>	7	0,021415	0,00305	6,076	0,3029135
Residuo <i>Residual</i>	1	0,000504	0,00050		
Totale <i>Total</i>	8	0,021919			

TABELLA 8 - TABLE 8

Coefficienti stimati della regressione (modello alternativo)  
*Estimated coefficient of the regression (alternative model)*

	Coefficienti <i>Coefficients</i>	Errore Standard <i>Standard Error</i>	Stat t
Intercetta <i>Intercept</i>	-0,026271846	1,235171	-0,0213
P	-4,54305E-09	5,93E-09	-0,7659
$V_C$	0,000514094	0,055218	0,0093
$V_C^2$	2,93307E-06	1,42E-05	0,2060
$F_S$	2,4565E-05	7,17E-05	0,3424
$F_S^2$	2,28496E-12	2,59E-12	0,8826
$P_F$	0,033167362	0,056669	0,5853
$P_F^2$	0,001496419	0,002941	0,5089

2). The third model, which has the same values of  $R^2$  and  $R^2_c$  of the previous one, has the same variables as the previous model with the addition of the mileage at the second degree and with the removal of the peak hour factor variable, always at the second degree.

The model 2 is characterized by the following equation which determines the overall efficiency index  $E_G$ :

$$E_G = 0,0029 - 3,82 \cdot 10^{-9} P + 0,0011 V_C + 0,00014 V_C^2 - 6,30 \cdot 10^{-5} F_S + 2,5 \cdot 10^{-8} F_S^2 + 0,0315 P_F + 0,0037 P_F^2 \quad (24)$$

For the 3 model, however, the relationship that allows the assessment of  $E_G$  is the following:



## POLITICA E ECONOMIA

In fig. 1 è presentata graficamente la relazione tra  $E_G$  e la variabile suddetta, che ha un buon coefficiente di determinazione multiplo ( $R^2 = 0,84$ ).

L'uso di questa unica variabile indipendente nella specificazione ci è stata suggerita dalle analisi di correlazione fatte mettendo in relazione  $E_G$  con la velocità commerciale  $V_C$  (fig. 2) e con l'intensità di servizio  $F_S$  (fig. 3), una volta per ognuno di essi. Si è visto così che, sulla base dei dati a disposizione, l'indice di efficienza globale  $E_G$  cresce in maniera grosso modo lineare, con un  $R^2 = 0,64$ , a fronte di una riduzione di  $V_C$  (proporzionalità inversa), mentre  $E_G$  cresce, anche qui linearmente con una buona approssimazione ( $R^2 = 0,69$ ), all'aumentare dell'intensità di servizio (proporzionalità diretta).

La proporzionalità inversa trovata tra l'indice di efficienza globale e la velocità commerciale, e quindi la proporzionalità diretta con il costo chilometrico per la (13), sembra andare contro quanto asserito da gran parte degli studi internazionali già richiamati in bibliografia.

Tuttavia vanno fatte delle considerazioni che giustificano questa relazione. Va tenuto in conto infatti che il servizio preso in considerazione è di tipo extraurbano e pertanto presuppone l'esistenza di tempi di attesa fra l'arrivo della corsa di andata e la partenza di quella del ritorno che aumentano con la crescita della velocità commerciale per effetto della riduzione della durata di ciascuna corsa. Di conseguenza, con l'aumento della velocità commerciale, si riduce il numero di ore di guida mentre il numero di ore di lavoro rimane praticamente lo stesso (così come fissato dal timetable), e ciò porta alla riduzione del coefficiente di utilizzazione medio degli autisti  $U_A$  e quindi, a parità delle altre condizioni, anche di  $E_G$  per la (12) e quindi all'aumento del costo chilometrico  $C_{sk}$  per la (13). Tuttavia, nonostante la riduzione di  $E_G$  al crescere della velocità commerciale, la presenza indiretta, e con peso nettamente maggiore, di quest'ultimo parametro nella espressione del costo  $C_{sk}$  attraverso il parametro delle ore di servizio annue  $H_{ST}$  (rapporto tra la percorrenza annua e la suddetta velocità commerciale), garantisce la proporzionalità inversa, intuitiva nonché confermata in tutti gli studi, fra costo unitario  $C_{sk}$  e velocità commerciale.

Sia il "modello alternativo", definito dalla (23), che il "modello semplifi-

TABELLA 9 - TABLE 9

Verifica della bontà di adattamento del modello alternativo  
Check of goodness of fit (alternative model)

Ambiti	Eg model	Eg reale	Cskmod	Csk reale	Differ.	Differ. %
Alta Valle High Valley	0,310	0,299	3,175	3,291	-0,116	-3,7
Centro Valle Center Valley	0,365	0,363	2,731	2,743	-0,012	-0,4
Bassa Valle Lower Valley	0,364	0,381	2,787	2,662	0,125	4,5
Cuneo	0,355	0,35	2,529	2,565	-0,036	-1,4
Torino	0,435	0,431	2,613	2,635	-0,022	-0,8
Molise	0,291	0,294	2,765	2,736	0,028	1,0
Bologna	0,310	0,311	2,994	2,988	0,006	0,2
Perugia 1	0,270	0,275	3,316	3,258	0,058	1,7
Perugia 2	0,289	0,285	2,931	2,975	-0,044	-1,5
				Media	-0,002	-0,04

$$E_G = -0,1641 - 5,04 \cdot 10^{-9} P + 1,42 \cdot 10^{-16} P^2 + 0,013 V_C - 2,93 \cdot 10^{-5} V_C^2 - 8,44 \cdot 10^{-5} F_S + 2,83 \cdot 10^{-8} F_S^2 + 0,043 \cdot P_F \quad (25)$$

The unit costs per kilometer for models 2 and 3 were respectively obtained by using the relations (24) and (25) and replacing in the expression (13).

In addition to the tested models, especially for the "full model" expressed by (21) and the "alternative model" defined by (23), we tried to find a specification that might be as simplified as possible in shape and number of variables, without any possibility of discount on the correctness of the method and results found. For this purpose, some attempts have been made and we found the most suitable model on the basis of the available databases. Identifying as the sole variable the squared ratio between the intensity of the ser-

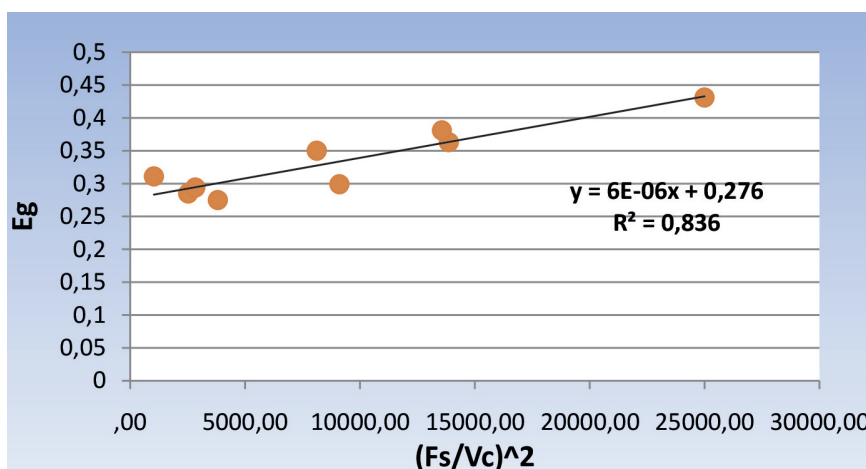


Fig. 1 - Regressione semplice lineare tra  $E_G$  e  $\left(\frac{F_S}{V_C}\right)^2$ .

Fig. 1 - Simple linear regression between  $E_G$  and  $\left(\frac{F_S}{V_C}\right)^2$ .

## POLITICA E ECONOMIA

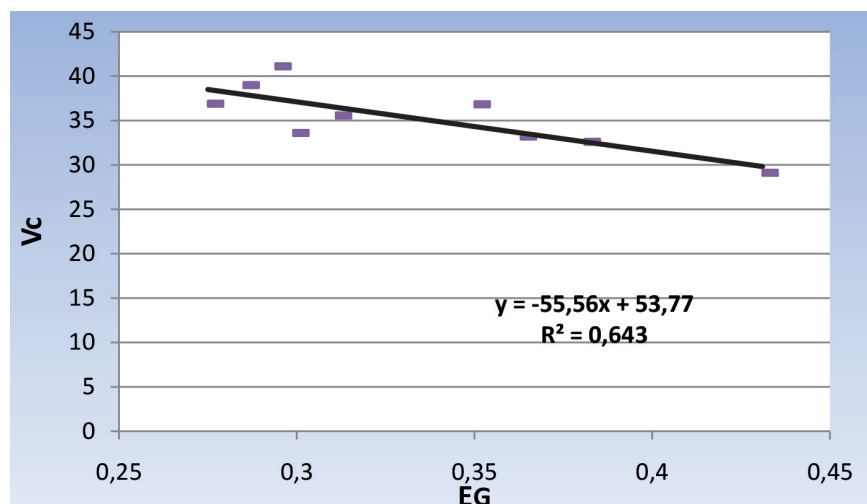


Fig. 2 - Correlazione lineare tra  $E_G$  e  $V_C$ .  
Fig. 2 - Linear correlation between  $E_G$  and  $V_C$ .

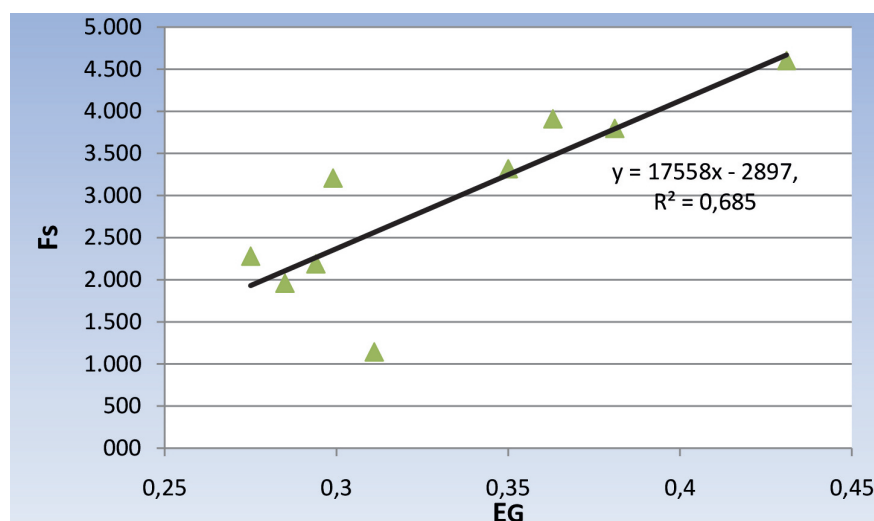


Fig. 3 - Correlazione lineare tra  $E_G$  e  $F_S$ .  
Fig. 3 - Linear correlation between  $E_G$  e  $F_S$ .

cato", espresso dalla (26), sono stati testati utilizzando i dati del database per diversi valori della velocità commerciale. Dall'analisi dei risultati in termini di costo chilometrico, si è potuto appurare che il "modello alternativo" presenta dei risultati accettabili per velocità superiori ai 30 km/h ma contenuti entro opportuni limiti di velocità caratteristici del trasporto pubblico locale su gomma extraurbano (intorno ai 50 km/h), mentre il "modello semplificato" presenta risultati più idonei per velocità commerciali basse, seppur al di sopra di un ragionevole limite stabilito dalle condizioni d'esercizio (20-25 km/h). In tabella 10 sono esposti i risultati di un caso specifico, in termini di coefficiente di efficienza globale e di costo chilometrico, per i due modelli citati e per diversi valori della velocità commerciale.

vice and the commercial speed, we found the expression that defines  $E_G$  in this simplified model through a simple linear regression:

$$E_G = 6 \cdot 10^{-6} \cdot \left(\frac{F_S}{V_C}\right)^2 + 0,2768 \quad (26)$$

In fig. 1 it's graphically presented the relationship between  $E_G$  and the said variable, which has a good coefficient of multiple determination ( $R^2 = 0,84$ ).

The use of this unique variable in the function is suggested by the correlation analysis made linking  $E_G$  index with the commercial speed  $V_C$  (fig. 2) and with the intensity of service  $F_S$  (fig. 3), once for each of them. So we have seen, basing on the available data, that the global efficiency indicator  $E_G$  grows roughly linear ( $R^2 = 0,64$ ) with a reduction of  $V_C$  (inverse proportionality), while  $E_G$  grows, here linearly with good approximation ( $R^2 = 0,69$ ), with the increase of intensity of service (proportionality).

The inverse proportionality found between the global efficiency index and the commercial speed, and therefore the direct proportionality with the cost per kilometre for the (13), seems to go against what many international studies, here already mentioned, have asserted. However must be made some considerations that justify this relationship. Should be taken into account that the service considered is the suburban – long-distance one. This type of service presupposes the existence of waiting time between the arrival of forward ride and the departure of return one, increasing with the growth of commercial speed, for the less duration of each rides. Therefore, with the increase of commercial speed, we have the reducing of the number of hours of driving while the number of hours worked remains virtually the same. This reduces the average coefficient of drivers utilization  $U_A$  and therefore, at the same other conditions, the  $E_G$  index for the (12) and increases the cost per kilometre  $C_{sk}$  for the (13).

However, despite of the reduction of  $E_G$  with the increase of commercial speed, we have noted a greater weight of this last parameter inside the hours of service  $H_{ST}$  into the (13) formula. So this result, supported by many studies, confirms what we expected, that is the inverse proportionality between unit cost  $C_{sk}$  and the commercial speed.

Both the "alternative model" (expressed by the (23) and the "simplified model", expressed by (26), were tested using

## POLITICA E ECONOMIA

Visti i risultati, la metodologia da adottare più opportuna consiste nel prendere in considerazione entrambi i modelli (23) e (26) definendo, per entrambi, i campi di applicazione in termini di velocità commerciale per non incorrere in risultati poco attendibili. Detto ciò, il “modello semplificato” può essere applicato per velocità comprese tra i 25 e i 30 km/h mentre il “modello alternativo” per velocità superiori ai 30 km/h ed inferiori ad un limite opportuno che può essere individuato in 50 km/h. Naturalmente detti limiti scaturiscono dalla velocità commerciale dei servizi compresi nel database utilizzato per la costruzione dei modelli. Avvalendosi di un data-base comprendente servizi con velocità commerciali diverse è possibile ottenere altre calibrazioni dei modelli con differenti intervalli di validità.

I modelli tengono in considerazione anche la eventuale presenza di Sistemi Intelligenti di Trasporto (ITS) dal momento che questi ultimi influiscono sulla variabile “velocità commerciale  $V_c$ ”. Detta variabile, infatti, è un indice indiretto delle caratteristiche della strada e del traffico nonché della eventuale presenza di facilitazioni finalizzate alla regolarità e alla velocizzazione del servizio di trasporto pubblico.

### 4. I casi studiati

Per valutare la affidabilità dei modelli, li abbiamo applicati alla stima del costo unitario di alcuni servizi di trasporto pubblico locale su gomma per i quali risultano verificate le assunzioni del data-base. Abbiamo poi confrontato il valore ottenuto dal modello con il costo chilometrico posto a base della gara per l'affidamento.

#### 4.1. Il trasporto pubblico extraurbano delle province di Potenza e Matera

Nel 2007 sono stati emanati due bandi di gara relativi all'assegnazione dei servizi extraurbani di trasporto pubblico riguardanti rispettivamente i due bacini provinciali di Matera e Potenza.

I costi chilometrici “reali” per i due ambiti sono stati ottenuti a partire dai corrispettivi chilometrici unitari posti a base di gara e aggiornati al 2014 mediante tasso ISTAT FOI, sommati al contributo integrativo chilometrico fisso (0,07 euro/km) e ai rispettivi ricavi chilometrici unitari effettivi relativi al 2011 (trattasi di contratti net-cost).

Per una migliore comparazione, oltre ai modelli descritti dalla (21) e dalla (23) e dalla (26), sono stati considerati dei modelli alternativi di regressione lineare ottenuti valutando combinazioni e gradi di potenza diversi

TABELLA 10 - TABLE 10

Valori di costo chilometrico per i modelli (23) e (26) nel caso di “Perugia 2”  
Cost per kilometer values for the models (23) and (26) in the case of “Perugia 2”

Ambito Perugia 2 - Area Perugia 2					
$V_c$ (km/h)	Csk reale Csk (real value)	Eg sempl Eg (simplified model)	Csk sempl Csk (simplified model)	Eg alter Eg (alternative model)	Csk alter Csk (alternative model)
10	11,601	0,507	<b>6,522</b>	0,103	<b>31,999</b>
20	5,801	0,334	<b>4,945</b>	0,129	<b>12,815</b>
30	3,867	0,302	<b>3,645</b>	0,190	<b>5,805</b>
40	2,900	0,291	<b>2,839</b>	0,304	<b>2,723</b>
50	2,320	0,286	<b>2,312</b>	0,488	<b>1,356</b>

data from the database for different values of commercial speed. We were able to ascertain that the “alternative model” shows acceptable results for speeds above 30 km/h but kept within appropriate limits of speed characteristic for local public road intercity transit (about 50 km/h). Instead the “simplified model” presents most suitable results for lower commercial speeds, but over a reasonable limit imposed by operating conditions. The table 10 exposes the results of a particular case, in terms of global efficiency and cost per kilometer, for the two models mentioned and for different values of commercial speed.

The more appropriated methodology to be adopted is to consider both models (23) and (26) defining the ranges in terms of commercial speed to avoid unreliable results. Following this indications, the “simplified model” can be applied to speeds between 25 and 30 km/h while the “alternative model” for speeds above 30 km/h and below a limit that should be located at 50 km/h.

Of course, these limits are correlated to the commercial speeds of the services contemplated into the database used for the construction of the model. Using a database that includes services with different commercial speeds we can get other calibrations of models with different validity ranges.

The thus set models already take into account the possible presence of Intelligent Transport Systems (ITS), since the effects of these last ones are already included within the “commercial speed”  $V_c$  variable. The latter, in fact, is an indirect indicator of road and traffic features and also of possible facilities to the public transport aimed to regularize and speed up the service.

### 4. Case studies

To evaluate the reliability of the models, we applied them to estimate the unit standard cost of road public transit services for which the data-base assumptions are verified. Then we compared the values calculated by the models, with the unit costs of the tenders in the data-base.



## POLITICA E ECONOMIA

delle variabili esplicative e quindi costruendo il “modello 2” descritto dalla (24) e il “modello 3” descritto dalla (25).

Sia per l'ambito di Matera che di Potenza, la velocità commerciale risulta superiore ai 30 km/h (43,2 km/h per Potenza, 51,9 km/h per Matera<sup>(9)</sup>) quindi si rientra nel campo di applicazione del “modello alternativo”. I risultati ottenuti in termini di costo chilometrico per i vari modelli citati sono esposti in fig. 4 per l'ambito extraurbano di Potenza e in fig. 5 per quello di Matera.

Dai due istogrammi, si può vedere che i modelli 2 e 3 propongono dei costi più bassi rispetto a quello “reale”, ma presentano dei valori in segno inaspettati e anomali per più di un parametro (intensità di servizio di grado 1 per il modello 2; velocità commerciale di grado 2, intensità di servizio di grado 1 per il modello 3), per le motivazioni spiegate nel paragrafo 2.2 quando si sono descritte le variabili esplicative. Il “modello completo” risulta invece controverso, poiché da una parte presenta un costo inferiore rispetto a quello “reale” (Potenza), mentre dall'altra parte risulta superiore (Matera), senza possibilità di una razionale spiegazione, mancando quindi di uniformità e coerenza. Più affidabile si presenta invece il “modello alternativo” visto che presenta dei ribassi di costo abbastanza simili per i due ambiti (-31% per Potenza e -25,8% per Matera, che di fatto presentano servizi tra loro molto simili) e data la differenza molto ridotta tra i due costi chilometrici calcolati mediante lo stesso modello (7 centesimi di euro, circa il 5%).

### 4.2. I servizi sostitutivi ed integrativi di FAL e Trenitalia (Basilicata - Italia)

Nell'anno 2010 la Regione Basilicata ha sottoscritto con le società di trasporto Trenitalia e Ferrovie Appulo-Lucane FAL, i contratti di affidamento dei servizi di trasporto che riguardano non solo le corse su ferro ma anche i servizi sostitutivi ed integrativi di quelli ferroviari di interesse locale su gomma. Ad oggi il contratto in essere è in proroga, nelle more della procedura di gara che verrà successivamente espletata, fissata dalle norme per il 2017.

Il costo unitario al chilometro concertato non è quindi da considerarsi di massima efficienza non essendo scaturito da alcuna procedura concorsuale. Pertanto sono stati considerati come costi di massima efficienza per le due categorie di servizi su gomma, di competenza di FAL e di Trenitalia, quelli derivanti da una media aritmetica dei costi standard calcolati con i metodi ANAV, ASSTRA e quello definito dalla Regione Basilicata con la L.R. 34/88 e successive modifiche (oggi abrogata) [10], in osservanza della legge nazionale 151/81 [9].

<sup>(9)</sup> Nonostante la velocità commerciale per l'ambito extraurbano di Matera sia superiore al limite fissato di 50 km/h, si è voluto comunque testare il modello per appurarne il comportamento, vista anche l'esigua entità dello sfioramento di un limite che non è rigido ma comunque indicativo del tipo di servizio e frutto dei dati presi in considerazione nel database.

### 4.1. The suburban and intercity public transport in the provinces of Potenza and Matera (Italy)

*In 2007, two public calls for tenders related to the award of public road transit services respectively for the two provincial areas of Potenza and Matera, are issued.*

*The “real” costs per kilometer for the two areas were obtained from the unit fees per kilometer fixed in the tenders and discounted to 2014 by ISTAT FOI rate, added to a fixed additional contribution (0.07 €/ km) and their respective unit revenues for 2011 (net-cost contracts).*

*For a better comparison, in addition to the models described by the (21), the (23) and by the (26), we have considered alternative models of linear regression obtained by evaluating combinations and degrees of different power of the explanatory variables. Among the different attempts, the models that gave more acceptable results are two: the model 2 expressed by the (24) and the model 3 expressed by the (25).*

*Both for Matera than for Potenza, the commercial speed is more than 30 km/h (43,2 km/h for Potenza, 51,9 km/h for Matera<sup>(9)</sup>), then these cases fall within the validity range of the “alternative model”. The results achieved in terms of cost per kilometer for the various models are shown in fig. 4 for Potenza and in fig. 5 for Matera.*

*From the two histograms, we can see that the models 2 and 3 produce lower costs than that “real”, but are provided with values in unexpected and abnormal sign for more than one parameter (intensity grade of service factor of 1st grade in the model 2; commercial speed of 2th grade and 1st grade for service intensity in the model 3), for the reasons explained in section 2.2 when the explanatory variables are described. The “full model” is controversial because by one hand has a lower cost than the “real” one (Potenza), while by the other side is higher (Matera) without the possibility of a rational explanation, therefore lacking uniformity and consistency. The “alternative model” is more reliable for the presence of cost reductions fairly similar for the two areas (-31% for Potenza, -25,8% for Matera) which actually have very similar services and a very low difference in unit cost per kilometer calculated with the same model (7 euro-cents, about 5%).*

### 4.2. The substitute and additional services of FAL and Trenitalia (Basilicata - Italy)

*In the year 2010 the Basilicata Region signed with the two Italian railway transport companies Trenitalia and FAL (Ferrovie Appulo Lucane), the contracts to entrust transport services, affecting not only the rail transit but also the substitute road transit services. Currently, the contract is ex-*

<sup>(9)</sup> Although the commercial speed of the suburban transit service for Matera is slightly superior to the limit of 50 km/h, we have still wanted to test the model to verify the behaviour, considering also the small amount of overshooting of this limit that is only indicative of the type of service and the result of the data taken into account in the database.

## POLITICA E ECONOMIA

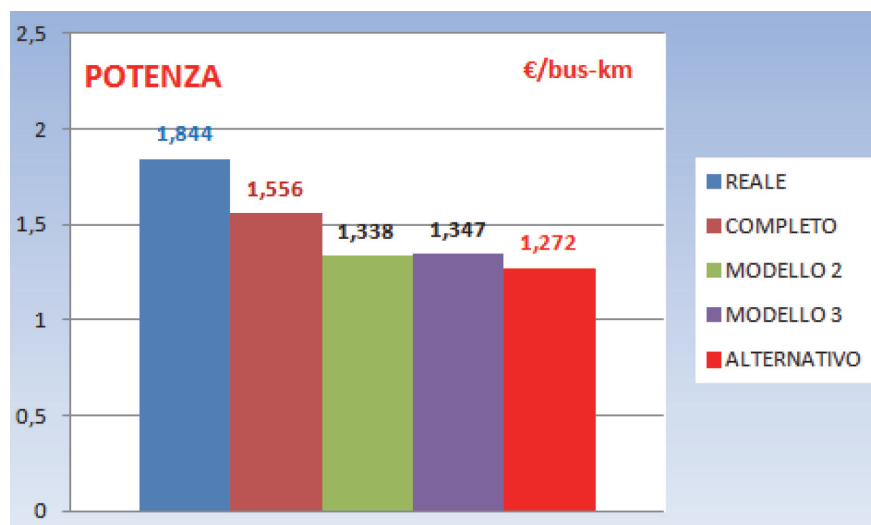


Fig. 4 - Confronto tra i costi chilometrici unitari calcolati secondo vari modelli per l'ambito extraurbano di Potenza (in euro/bus-km).

Fig. 4 - Comparison of unit costs per kilometre calculated according to various models for the suburban service of Potenza.

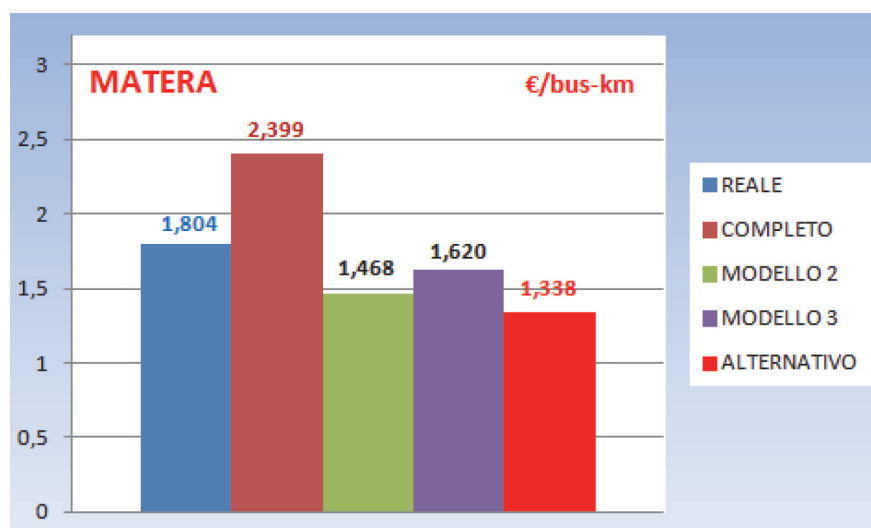


Fig. 5 - Confronto tra i costi chilometrici unitari calcolati secondo vari modelli per l'ambito extraurbano di Matera (in euro/bus-km).

Fig. 5 - Comparison of unit costs per kilometer calculated according to various models for the suburban service of Matera.

La metodologia definita dalla Regione Basilicata, con la L.R. 34/88 e successive modifiche, prevede che il costo del personale di movimento venga dedotto dai documenti contabili di servizio, con una presunzione di massima efficienza aziendale, che risulta però non verificata. Si è introdotta perciò una modifica, calcolando il costo del personale di movimento sulla base dei programmi di esercizio, attraverso la velocità commerciale e le ore di servizio complessivamente rese nell'anno valutabili dagli orari di servizio. I costi relativi al personale amministrativo e ausiliario si ottengono invece come percentuale, definita per

tended, pending the tendering procedure that will be carried out by 2017, as fixed by the laws. The here agreed unit cost per kilometer cannot be regarded as a maximum efficiency cost, not resulting from any tendering.

So we have considered other types of calculation of the cost as costs of maximum efficiency for the two categories of road services appertaining to FAL and Trenitalia. The above cost is the result of the arithmetical mean of the standard costs calculated by AS-STRADA and ANAV methods and the method defined by the Basilicata Region with the Regional law 34/88 and following modifications [10] (now repealed), in accordance with the National law 151/81 [9].

The methodology, defined by the Basilicata Region by the Regional Law 34/88 and following modifications provides that the movement personnel costs were inferred by records of service, with a presumption of maximum efficiency for the firm which is not verified. So we've introduced an amendment calculating the movement personnel costs, based on operating programs, through the commercial speed and the hours of service made during the year evaluated with the same operating programs. Auxiliary and administrative personnel costs are obtained instead as a percentage, defined by law, of personnel movement cost, while the service director costs are defined from the National Collective Labour Agreement. The Regional law 34/88 also provides the values of parameters to be taken into calculation procedure that reference, in addition to the personnel costs, also the costs of traction of buses, the amortization costs of the plants and overhead costs.

The ANAV method [2] is based on the pattern that has already mentioned in the state of the art (section 1), which expresses the cost as a function of endogenous and exogenous parameters taken with its minimum and maximum values, so there are two types of cost (best and worst). The values of these parameters have been validated through empirical observations on a representative sample of transport companies associate to ANAV, distinguished according to the type of service (urban and intercity). The methodology considers the initial evaluation of the driver cost per kilometer obtained mostly by exogenous parameters (working

## POLITICA E ECONOMIA

legge, del personale di movimento mentre i costi relativi al responsabile di esercizio sono desunti dai CCNL. La legge regionale 34/88 fornisce inoltre i valori dei parametri da adottare nella procedura di calcolo che fanno riferimento, oltre al costo del personale, anche ai costi di trazione, degli autobus, di ammortamento degli impianti e delle spese generali.

Il metodo dell'ANAV [2] si basa sul modello a cui si è già accennato nello stato dell'arte, che esprime il costo in funzione di parametri endogeni ed esogeni presi con valori minimi e massimi, in modo da avere due tipologie di costo (migliore e peggiore). I valori di questi parametri sono stati validati attraverso osservazioni empiriche su di un campione rappresentativo stabilito da ANAV di aziende di trasporto consociate distinte in base al tipo di servizio (urbano ed extraurbano). La metodologia prevede la valutazione iniziale del costo chilometrico per autista ricavato per lo più da parametri esogeni (giornate lavorative annue, velocità commerciale minima, ore di guida giornaliere per autista, costo annuo di un autista). Da questo valore di costo chilometrico si ricava il costo complessivo del lavoro e, da quest'ultimo, il costo complessivo del servizio mediante opportuni coefficienti. Aggiungendo a tale costo l'utile d'impresa (10%) e assumendo un rapporto ricavi/costi totali pari al 35%, si giunge a definire il compenso di esercizio. Nel costo totale sono inserite anche le tasse (IRAP) con una incidenza del 3,9% e le quote di ammortamento dei veicoli a carico dell'azienda (valori medi quantificati in 345.039 euro/anno per il servizio extraurbano e 345.424 euro/anno per il servizio urbano).

Il modello dell'ASSTRA (già citato nello stato dell'arte) è stato costruito avvalendosi dei dati provenienti da un campione di aziende di trasporto pubblico consociate (approccio microanalitico). La metodologia si basa su una suddivisione dei costi per processi aziendali, valutando l'efficienza non solo del singolo fattore produttivo ma dell'intera area di processo aziendale. Come nel modello ANAV, viene effettuata la stima del numero di addetti alla guida in base alle percorrenze, alla velocità commerciale e ai parametri di produttività dello stesso personale (giornate lavorative annue e ore di guida giornaliere). Il personale addetto agli altri settori (marketing, commerciale e amministrativo, ecc.) viene calcolato come percentuale del personale di guida. Altre caratteristiche in comune con il modello precedente sono la presenza di due scenari, migliore e peggiore, in base ai valori assunti dai parametri considerati, un tasso di remunerazione del capitale al 10% e l'aver considerato, per la valutazione del costo, l'IRAP e le quote di ammortamento a carico del gestore.

I modelli ANAV e ASSTRA, pur essendo caratterizzati da un approccio diverso, sono strutturalmente simili. Entrambi inoltre, pur presentandosi concettualmente ineccepibili, assumono parametri difficilmente valutabili all'esterno dell'azienda e quindi non validati.

*days per year, minimum speed, daily riding hours per driver, annual cost of a driver). From this value of cost per kilometer is derived the total cost of the work and, from this last, the total cost of the service through appropriate coefficients. Adding to this last cost the business profit (10%) and assuming a ratio incomes/total cost equals to 35%, we can get the operating fee. In the total cost are also included local taxes (IRAP) with a percentage of 3,9% and amortization shares of the vehicles at the expense of the company (average value estimated at 345,039 Euro/year for the intercity service and 345,424 Euro/year for the urban service).*

*The ASSTRA model, already mentioned in the state of art (section 1), was built using data from a sample of public transport companies of the Association (micro-analytical approach). The methodology is based on a breakdown of costs for business processes, assessing the efficiency of not only the single factor of production but of the entire area of business process. As in the ANAV model, the ASSTRA model estimates the number of drivers depending on speed, distance and parameters of productivity of the same driver personnel (annual working days and hours of driving per day). Personnel assigned to other sectors (marketing, sales and administrative) is calculated as a percentage of the driver personnel. Other features in common with the previous cost evaluation model are the presence of two scenarios, best and worst, based on the values of the parameters considered, and also to take into account a rate of return on capital equal to 10%, the local tax and the amortization shares at the expense of the company.*

*ANAV and ASSTRA models, despite of a different approach in their methodology, are structurally similar. They both present a flawless conception, taking into account hardly evaluable parameters outside the company and therefore not validated.*

*As a point of comparison and to check the validity of the "alternative model", we have applied its methodology to FAL and Trenitalia services using the (23) to calculate their global efficiency indicators. The results have included in the (13) for the calculation of the unit standard costs. So we have collected the unit costs calculated with different models in fig. 6 for FAL services and in fig. 7 for Trenitalia services.*

*In fig. 6 we can see that "alternative model" performs very well in the case of FAL road transport services, showing a unit cost per kilometer of 2,76 euro/km that is largely in line with cost values calculated with the ANAV and the Basilicata Region models, while it's lower than about 40 cents from the value calculated by the ASSTRA model. The same consideration cannot be said about the results in relation to the Trenitalia services, since the "alternative model" leads to an abnormal cost value as decidedly small (just over 50 cents) certainly unlikely, compared with the unit costs calculated with other models included in the gap between 1.50 euro/km and 2 euro/km. This result takes explanation in the different operating characteristics of the lines operated by Trenitalia that cannot be defined of local transport in the strict sense, given the high average commercial speed (66 km/h against the 38 km/h of FAL services) and given the low number of stops, typical instead of an intercity long-dis-*



## POLITICA E ECONOMIA

Come termine di confronto e per verificarne la validità, è stato applicato il “modello alternativo” ai suddetti servizi di FAL e Trenitalia utilizzando la (23) per il calcolo dei rispettivi indici di efficienza globale, e i valori trovati sono stati inseriti poi nella (13) per il calcolo dei costi chilometrici standard. I risultati per i vari modelli citati sono rappresentati in fig. 6 per i servizi FAL e in fig. 7 per i servizi Trenitalia su gomma.

Dalla fig. 6, si può notare che “il modello alternativo”, formulato tramite la (23), si comporta molto bene nel caso dei servizi di trasporto su gomma eserciti da FAL, mostrando un costo chilometrico unitario di 2,76 euro/km che è sostanzialmente in linea con i valori di costo calcolati con il modello ANAV e con quello della Regione Basilicata, mentre risulta più basso di circa 40 cent rispetto al valore calcolato con il modello ASSTRA. Lo stesso non si può dire dei risultati relativi ai servizi eserciti da Trenitalia, poiché il “modello alternativo” porta ad un valore di costo anomalo, in quanto decisamente esiguo (poco più di 50 cent), non confrontabile quindi con i costi unitari calcolati con gli altri modelli e compresi nella forbice tra 1,50 euro/km e 2 euro/km e certamente poco verosimile. Questo risultato trova spiegazione nelle diverse caratteristiche di esercizio delle linee gestite da Trenitalia che non possono definirsi di trasporto locale in senso stretto, data l'elevata velocità commerciale media (66 km/h contro i 38 km/h dei servizi FAL) e il basso numero di fermate tipiche invece di un servizio intercity. Pertanto il modello espresso tramite la (23), che è stato calibrato a partire da un database riferito a servizi di trasporto locale di accessibilità, mal si adatta a valutare il costo per questa tipologia di servizi di collegamento.

### 5. Conclusioni

Il modello proposto vuole essere una valida alternativa per la stima dei costi di massima efficienza economica dei servizi di trasporto su gomma al fine di supportare la determinazione delle sovvenzioni di esercizio da corrispondere alle aziende di trasporto.

I vantaggi più lampanti si possono riscontrare nella semplicità della metodologia, nella facilità di calcolo dei

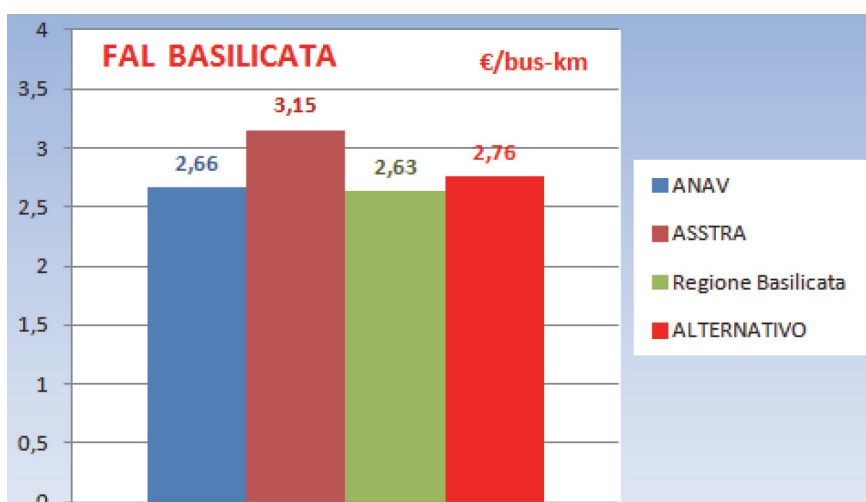


Fig. 6 - Confronto tra i costi chilometrici unitari calcolati secondo vari modelli per i servizi su gomma di FAL (in euro/bus-km).

Fig. 6 - Comparison of the unit standard costs per kilometer calculated according to various models for FAL road transit services.

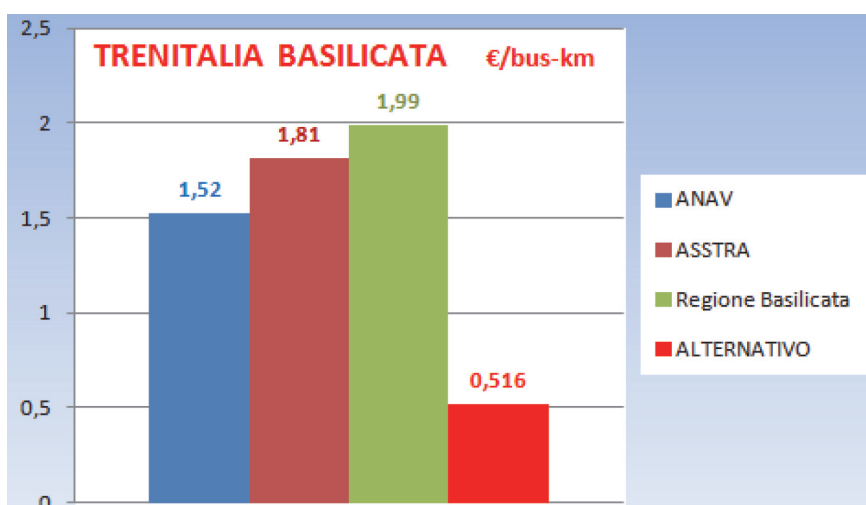


Fig. 7 - Confronto tra i costi chilometrici unitari calcolati secondo vari modelli per i servizi su gomma di Trenitalia (in euro/bus-km).

Fig. 7 - Comparison of the unit standard costs per kilometer calculated according to various models for Trenitalia road transit services.

tance service. Therefore the model expressed by the (23), which is calibrated from a database that refers to local accessibility transit services, is not suitable to estimate the cost of link transit services like the Trenitalia ones.

### 5. Conclusions

The built model aims to be a viable alternative to assess the maximum efficiency cost of the road transit services in order to support the determination of the operating subsidy to be paid to the transport companies.

The most obvious advantages can be found in the sim-

## POLITICA E ECONOMIA

parametri di esercizio, desumibili dai quadri orari e dai documenti di gara, e nella trasparenza e confrontabilità dei risultati. Tuttavia va riconosciuto che le formulazioni trovate risentono della scarsità di dati resi disponibili dagli enti appaltanti i servizi di trasporto che ha portato ad avere un database limitato nel numero di casi. Disponendo di un ben più esteso archivio è possibile ottenere calibrature più raffinate dei modelli individuati.

Nonostante i limiti evidenziati, si sono individuati i campi di validità per i modelli calibrati, che sono rappresentati da valori medio-alti delle velocità commerciali per il modello descritto dalla (23) e valori medio-bassi per il modello espresso dalla (26).

Le formulazioni qui messe a punto, seppure con le limitazioni evidenziate, forniscono comunque risultati accettabili e una bontà di adattamento elevata per i servizi di trasporto con spiccate caratteristiche di accessibilità. In presenza di un database esteso sarebbe inoltre opportuno operare la classificazione dei servizi di trasporto, differenziandoli in base alle loro caratteristiche di esercizio (soprattutto fra servizi di collegamento, che presentano poche fermate e velocità commerciali elevate, e servizi di accessibilità, caratterizzati da fermate frequenti e velocità commerciali basse), in modo da avere più modelli, ognuno adatto ad una specifica categoria di servizi.

*plicity of the methodology, in the computational ease of the operational parameters that can be derived from timetables and the tender documents, and in the transparency and comparability of the results. However it should be recognized that the built formulations are affected by the lack of data made available by the transport services contracting authorities that has led to a limited number of cases in the database. Having a much more extensive archive you can get more refined calibration of the identified models.*

*Despite the highlighted limits, the validity fields for calibrated models, which are represented by medium-high values of commercial speeds for the two models described by (23) and medium-low values for the model expressed by (26), are identified.*

*The formulations here developed, however, although with the highlighted limitations, provide acceptable results and a high goodness of fit for transport services with outstanding accessibility features. In the presence of an extensive database it would also be appropriate to make the classification of transport services, differentiating them according to their operating characteristics (especially between connecting services, which have few stops and high commercial speed, and accessibility services, characterized by frequent stops and low commercial speed), so we could have more models, each suited to a specific category of service.*

### BIBLIOGRAFIA - REFERENCES

- [1] ABRATE G., ERBETTA F., FRAQUELLI G., VANNONI D. (2014) – “Cost function estimation of multi-service firms. Evidence from the passenger transport industry”, Carlo Alberto Notebooks, n. 380, <http://www.carloalberto.org/assets/working-papers/no.380.pdf>.
- [2] ANAV, La Sapienza University (2013) – “La determinazione del costo standard nei servizi di trasporto pubblico locale su autobus: aspetti metodologici e prime esperienze applicative”, Residenza di Ripetta, Roma, <http://www.ferpress.it/wp-content/uploads/2013/06/Studio-su-Costo-Standard2.pdf>.
- [3] ASSTRA (2013) – “Un modello di calcolo del costo standard per il trasporto pubblico locale e regionale automobilistico”, Audizione presso la IX Commissione trasporti, poste e telecomunicazioni della Camera dei Deputati, Roma, [http://www.sindacatofast.it/sites/default/files/doc\\_pdf\\_inf/audizione%20commissione%20trasporti.pdf](http://www.sindacatofast.it/sites/default/files/doc_pdf_inf/audizione%20commissione%20trasporti.pdf).
- [4] AVENALI A., BOITANI A., CATALANO G., D'ALFONSO T., MATTEUCCI G. (2014) – “Un modello per la determinazione del costo standard nei servizi di trasporto pubblico locale su autobus in Italia”, Economia e politica industriale, n. 4, 181-213.
- [5] Cubukcu K. MERT (2006) – “Cost of Urban Bus Transit Operations and Geography of Service Territory”, Conference of Regional and Urban Modeling, Free University of Brussels, <http://ecomod.net/sites/default/files/document-conference/ecomod2006-rum/1149.pdf>.
- [6] DALEN D.M., GOMEZ LOBO A. (2002) – “Regulatory contracts and cost efficiency in the Norwegian Bus Industry: Do high-powered contracts really work?”, 7th Annual Conference of the Latin American and Caribbean Economic Association, Discussion Paper 6, <http://brage.bibsys.no/xmlui/bitstream/handle/11250/94048/1/Dalen1dp2002-06.pdf>.
- [7] FILIPPINI M., MAGGI R., PRIONI P. (2001) – “Cost-based Yardstick Regulation in the Swiss Regional Public Bus Industry”, Conference paper, 1st Swiss Transport Research Conference, Monte Verità/Ascona, [https://www.researchgate.net/publication/228603705\\_Cost-based\\_Yardstick\\_Regulation\\_in\\_the\\_Swiss\\_Regional\\_Public\\_Bus\\_Industry](https://www.researchgate.net/publication/228603705_Cost-based_Yardstick_Regulation_in_the_Swiss_Regional_Public_Bus_Industry).
- [8] PULLEY L.B., BRAUNSTEIN Y.M. (1992) – “A Composite Cost Function for Multiproduct firms with an application to Economies of Scope in Banking”, Review of Economics and Statistics, 74, 221-230.
- [9] Repubblica Italiana, Legge 10/4/1981, n. 151. [http://www.retecivica.trieste.it/new/admin/Immagini\\_up/html/20050608172221.pdf](http://www.retecivica.trieste.it/new/admin/Immagini_up/html/20050608172221.pdf)
- [10] Repubblica Italiana, Regione Basilicata, Legge Regionale n. 34/1988.