

# The Argentinean North Central Railway, A State Funded line: Some Results About its Economic Performance for the period from 1889 to 1920

María de las Mercedes Abril  
Universidad Nacional de Tucumán.  
Argentina

María Beatriz Blanco  
Universidad Nacional de Tucumán.  
Argentina

At the beginning of the 20<sup>th</sup> century, the state funded railways in Argentina became one the most important enterprises in the country, partly because of its political role as well as its prominent part in the national economy. Our goal is to analyze the behavior of the North Central Railway. In order to do that, we will work with a series of variables that describe its economic performance, such as the annual series of rail kilometers, transported passengers, transported cargo (measured in tons) and a measure of the invested capital, which we are going to called *return*. We are going to use some techniques related to time series analysis. These techniques will lead us to check if the transport policy carried out by the Argentine State was done with a true sense of economic gain or not.

*Keywords:* State funded enterprises, railways, economic performance, time series.

## Introduction

### The railways as an agent of change and development

From the second half of the 19<sup>th</sup> century, we can see that the railways help transform Argentina from a rural country to a more developed and integrated one. The railways arrived late to our country, they were finally ready to operate in late 1857, when the West Railway, built with national funds, was inaugurated. The introduction of this mean of transport changed forever the argentine physiognomy and with this began a development stage for our country. But it was necessary to bring the basic machinery and working labor from Europe for its correct installation.

According to its level of economic development, we can easily identify two large zones in our country. The littoral with a production required in external markets, was the zone that attracted private funding. And there were other places far away from the country's main harbor, specially the ones in the north that did not

have the same production so the expected return was lesser. It was there where the government, with great effort, invested in railways and other means of transport, despite this serious setback.

The building of a railway is an enterprise that demands large amounts of money, because of its size and complexity. The state funded network grew in periods of economic stability. This expansion was made using Treasury resources and with internal as well as external loans. These loans were supposed to be paid with taxes incomes and using, in some way the expected railway benefits. The State was largely indebted as a result of the construction and the equipment of its railways.

One of these state funded lines was the North Central Railway, projected in late 1868 according to Law N° 280. Its construction as well as its mechanical performance was filled with serious errors. Among them was the choosing of the trocha. They eventually decided to use a narrow one, mainly because of economical reasons. Other issue was the lamentable state of the infrastructure because of the reduction in construction costs and the little knowledge of the terrain where the rail track was supposed to pass. The government tried several times to lease it, but finally had to take over its exploitation. Nevertheless, it cannot be denied that the State provided the north part of the country, a place far away from the distribution points and the larger population centers, with a vast railway network that crossed the country. At the same time, this modern form of transport allowed communication as well as trade exchange with nearby regions and enabled the integration of the north to the economic development of the country.

But in 1914 everything changed. The construction of every line was suspended because of the lack of funding due to the First World War. This caused the decline of the product trade and it stopped all the branches of the rail activity. All of these factors created an intense crisis that affected in a crucial way the performance of every line and this one in particular. Things slowly began to take its normal course by the beginning of the 1920s.

## Method

### Some ideas on time series modeling and volatility

The general idea behind a time series is quite simple. It simply is the register of any quantity that evolves through time. Certainly the data from the railway activity can be circumscribed into this definition.

Specifically a time series can be defined as data set, such as  $y_1, y_2, \dots, y_n$  that is ordered in time. The most basic and general model used to represent any time series is given by

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \quad t = 1, 2, \dots, n, \quad (1)$$

where  $\mu_t$  is a component that changes smoothly over time known as *tendency*,  $\gamma_t$  is a component with a fixed period called *seasonality* and  $\varepsilon_t$  is an irregular component called *error*.

The main reason for modeling any time series is to allow the prediction of future values, although this is not the case, since we deal with historical data. The hallmark of a time series model, as opposed to an econometric model is that no attempt is made to formulate a behavioral relationship between the time series and other explanatory variables considered. The movements of the series are explained only in terms of its own past, or its position in relation to time or its structure. Prediction, if the problem we are trying to solve requires it, is made through extrapolation.

We also see that many time series, specially the ones we study in practice, do not have a constant mean and in most cases there are phases where there is relative calmness followed by periods of significant changes.

There is a characteristic of a time series of this type known with the name of *volatility*, that is not directly observable. To operate under the presence of volatility in time series it is necessary to use certain kind of models such as the *conditional heteroscedastic models*. In these models, the variance, or the volatility, of a series at any given point in time depends on their past and other information available up to that moment of time, so one must define a conditional variance, which is not constant and does not match the overall variance (unconditional or conditional) of the observed series.

## Research Model

### The state space approach

The basic state space model (or *MBEE* according to its acronym) used to describe the behavior in general terms of any series can be expressed mathematically as follows

$$\begin{aligned} \mathbf{y}_t &= \mathbf{Z}_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t & \boldsymbol{\varepsilon}_t &\sim N(\mathbf{0}, \mathbf{H}_t) \\ \boldsymbol{\alpha}_t &= \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \mathbf{R}_t \boldsymbol{\eta}_t & \boldsymbol{\eta}_t &\sim N(\mathbf{0}, \mathbf{Q}_t) \end{aligned} \quad (2)$$

where  $\mathbf{y}_t$  is a vector of  $p \times 1$  observations,  $\boldsymbol{\alpha}_t$  is a  $m \times 1$  vector of unobservable components called *state vector*, and  $\mathbf{Z}_t$ ,  $\mathbf{T}_t$  and  $\mathbf{R}_t$  are known matrices. The first equation in (2) is known as the *measure equation* while the second has the name of *transition equation*.

Formula (2) can be thought as a regression model where  $\mathbf{Z}_t$  can be seen as the regressor matrix and  $\boldsymbol{\alpha}_t$  can take the part of the coefficient vector which in this case is stochastic. This representation is a very general one and popular models such as the *ARIMA* are often seen as particular cases of this one.

The underlying idea behind this is that the different components of the series are modeled separately and these submodels are later put together in a single linear model. The relevant estimates (2) are done by applying elements such as the *filter* and the *Kalman smoother*.

In very brief terms we can say that the purpose of the Kalman filter is to update our knowledge of the system each time a new observation is obtained.

On the other hand, the smoother is used to estimate unobservable components, such as trend or the cycle, and to calculate statistics used to detect outliers and structural changes and is based on the entire sample.

### Volatility

Volatility in a time series can be defined as the variance of random variable, conditional on all past information. As volatility cannot be measured directly, it can manifest itself in several ways in a time series, especially for those series related to economic activity as this case.

Lets define our series by  $y_t$  whose dimension is given by  $p = 1$ . We can define

$$\mu_t = E(y_t | F_{t-1}) = E_{t-1}(y_t), \quad (3)$$

and

$$\begin{aligned} h_t &= \text{var}(y_t | F_{t-1}) = E\{(y_t - \mu_t)^2 | F_{t-1}\} \\ &= E_{t-1}(y_t - \mu_t)^2 = \text{var}_{t-1}(y_t), \end{aligned} \quad (4)$$

as the conditional mean and the conditional variance of  $y_t$  given the information up to moment  $t-1$  contained in  $F_{t-1}$ .

A volatility model has the following form

$$y_t = \mu_t + \sqrt{h_t} \varepsilon_t, \quad (5)$$

where  $E_{t-1}(\varepsilon_t) = 0$  and  $\text{var}_{t-1}(\varepsilon_t) = 1$ . Typically, the  $\varepsilon_t$  are independent and identically distributed (or *IID*) with an  $F$  distribution. The unconditional mean and variance will be denoted by  $\mu = E(y_t)$  and  $\sigma^2 = \text{var}(y_t)$ , respectively. The distribution of  $y_t$  will be denoted by  $G$ . It is clear that (3), (4) and  $F$  determine  $\mu, \sigma^2$  and  $G$  although the reverse it's not always true. Further details can be found at Abril (2014). Depending on how we define  $h_t$  we'll have *ARCH* or *GARCH* models or any other variation of them.

### The stochastic volatility

In the analysis of any economic series as in this case it is often seen that the variance of observational errors is subject to substantial variability over time. This phenomenon is known with the name *clustering volatility*. The variability of the models for this type of series can be obtained through *stochastic volatility models* (or *SV* as they are also known).

*SV* models also have a strong connection with the state space approach.

Lets denote the series under study as  $y_t$ . A basic model for  $y_t$  with *SV* is given by the following formula

$$y_t = a + \sigma \exp\left(\frac{1}{2}\theta_t\right) \varepsilon_t, \quad \varepsilon_t \sim N(0,1), \quad (6)$$

where the mean, given by  $a$  and the standard deviation, given by  $\sigma$  are supposed to be fixed and unknown. The signal  $\theta_t$  can be defined as the unobserved logarithm of the volatility and can be modeled as  $\theta_t = \mathbf{Z}_t \boldsymbol{\alpha}_t$  where  $\boldsymbol{\alpha}_t$  is generated by the second line in (2), that is, a basic state space form. In the standard cases,  $\theta_t$  is modeled by an *AR*(1) model with Gaussian disturbances, that is

$$\theta_t = \rho\theta_{t-1} + \eta_t \quad \eta_t \sim N(0, \sigma_\eta^2), \quad 0 < \rho < 1. \quad (7)$$

Using and adequate transformation we can express (6) together with (7) into a system that has a state space form, this is why we can use these two approaches together.

The model in (6) captures only the main features of the changing volatility in a time series. We can make it more precise when we model the mean in  $y_t$  adding explanatory variables.

Estimation of the parameters for the case of *SV* models can be a tricky problem that is not yet solved in an adequate way; we just have to read Shephard (1996) and Ghysels et al. (1996). Gaussian linear techniques only provide approximate maximum likelihood estimators of the parameters and can only be applied to the case of the basic model *SV* like the one in (6).

## Data Analysis and Discussion

### Statistical analysis of the North Central Railway line figures for this period.

The studied data was obtained from a statistical publication of this period called "*Estadísticas de los Ferrocarriles en Explotación*" surveyed by the Ministerio de Obras Públicas de la Nación Argentina.

To analyze the performance of the railway we have two sets of series. On one hand we have the series of rail lines tracks (expressed in kilometers), passengers, cargo(in tones) and the return series, the latter being a measure of the invested capital. Such series are shown in **Figure 1**.

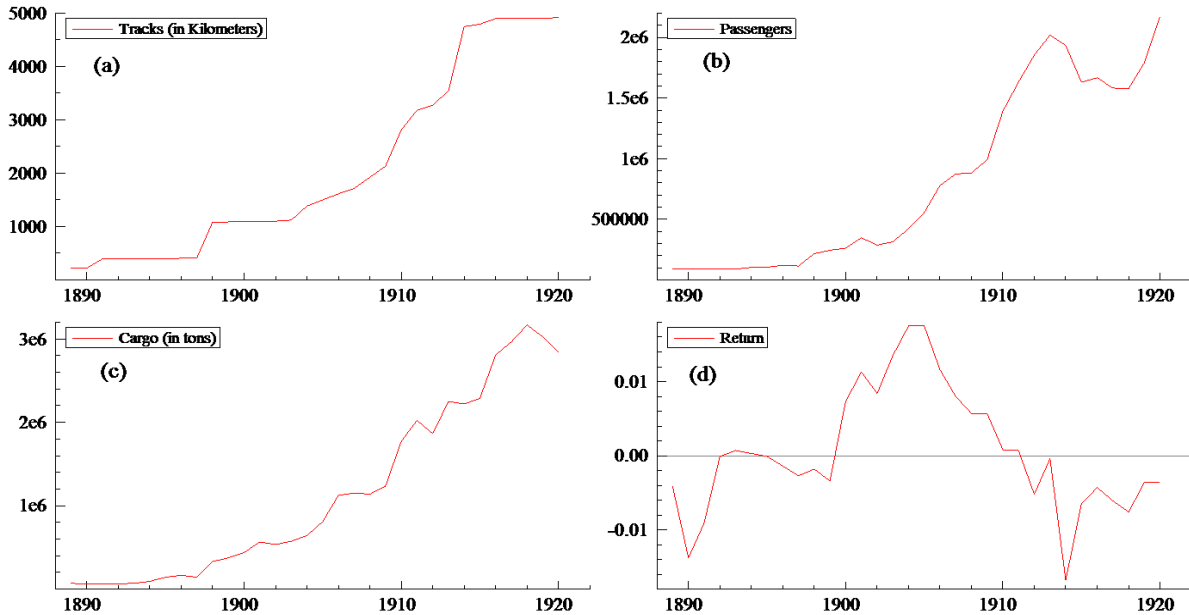


Figure 1. Annual series from 1889 to 1920. (a) Extension of rail tracks in kilometers; (b) number of passengers; (c) cargo transported (in tones) and (d) return on invested capital.

On the other we also have the following series, all of them presented on **Figure 2**. These are, passengers (divided by rail kilometers), total transported cargo (divided by rail kilometers) and again the measure of the invested capital that we decided to called *return* in our scheme.

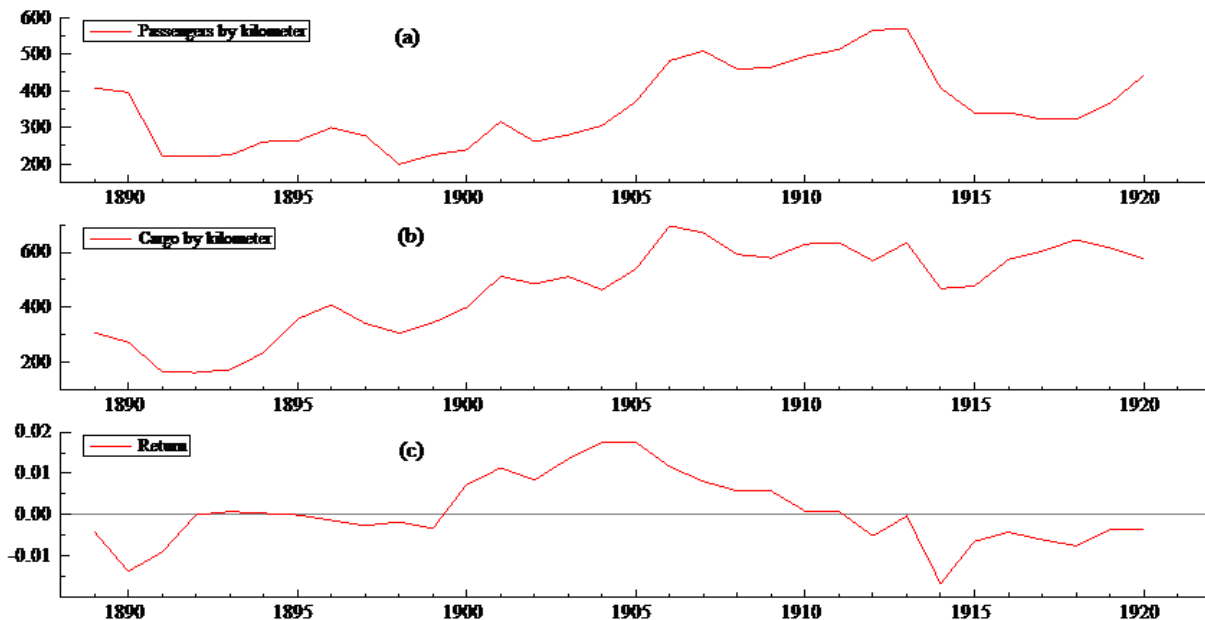


Figure 2. Annual series divided by line kilometers from 1889 to 1920. (a)Transported passengers by divided line kilometers. (b) Cargo divided by line kilometers and (c) return of the invested capital.

For the rail track series, that is, the first data set under study, we adjusted a basic structural model in state space form, with a fixed level, an stochastic slope and an intervention as an explanatory variable in 1913 to capture extraordinary value in the irregular movements, due to the Balkan Crisis, previous to the Great War. This time series is clearly shown in **Figure 3** where we can see the series along with the trend, the trend expressed individually, the slope and the estimated irregular component.

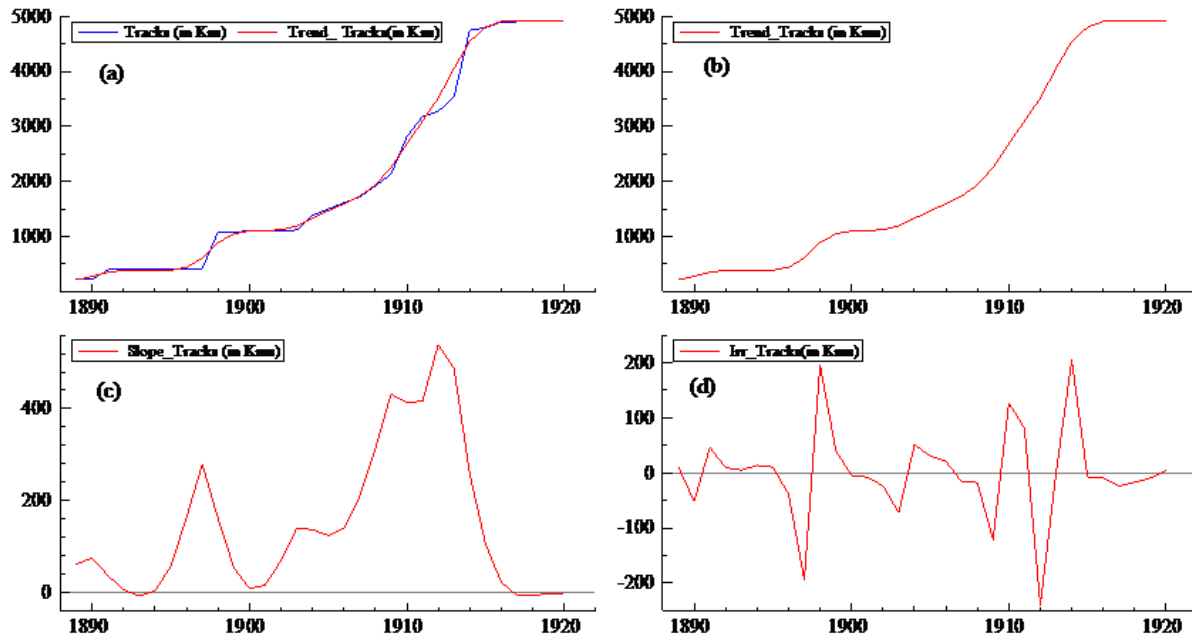


Figure 3. Annual series of rail tracks expressed in km from 1889 to 1920. (a) Original series (blue line) and its estimated trend (red line); (b) estimated trend; (c) estimated stochastic slope and (d) estimated irregular component.

In the case of the passengers series, cargo series and return series we tried to adjust a state space model as we did before but it turned out that it was not suitable. Therefore, we tried with an SV model with satisfactory results in all cases.

For the passenger series, we observe that volatility is a dominant feature since the disturbances turned out to have an extremely small variance. This situation is put in evidence in **Figure 4**, where the graphic describes a straight line that starts almost from the beginning and has a positive slope. We also have to note that in this case we needed to put an intervention in 1891 that corresponded to a payment balance crisis, due to an irresponsible behavior from the national government.

For the case of the return variable, expressed in (c) according to **Figure 1** and again in **Figure 2**, we needed to put two interventions. One in 1892, for the reason we just explained above, and also another in 1913, corresponding to Balkan Crisis, previous to the First World War. We see that volatility also plays a great role when this line had earnings as well as when it had losses.

For the cargo series we did not have to put any interventions. The estimated value for  $\rho$  was  $\hat{\rho} = 0,95282$ . Nevertheless we cannot come to the same conclusions as we did before. Even though there is an important relation between the volatility and the studied series it cannot be explained exclusively by the model since there are other forces involved in the behavior of the series. These forces are historical questions such as the ongoing struggle at that time to set a correct freight rate or by the fact that these lines were built with a

sense of promotion because its main task was to bring closer the production of those places that were far from the country capital and economic center. This situation is evidenced **Figure 6** below.

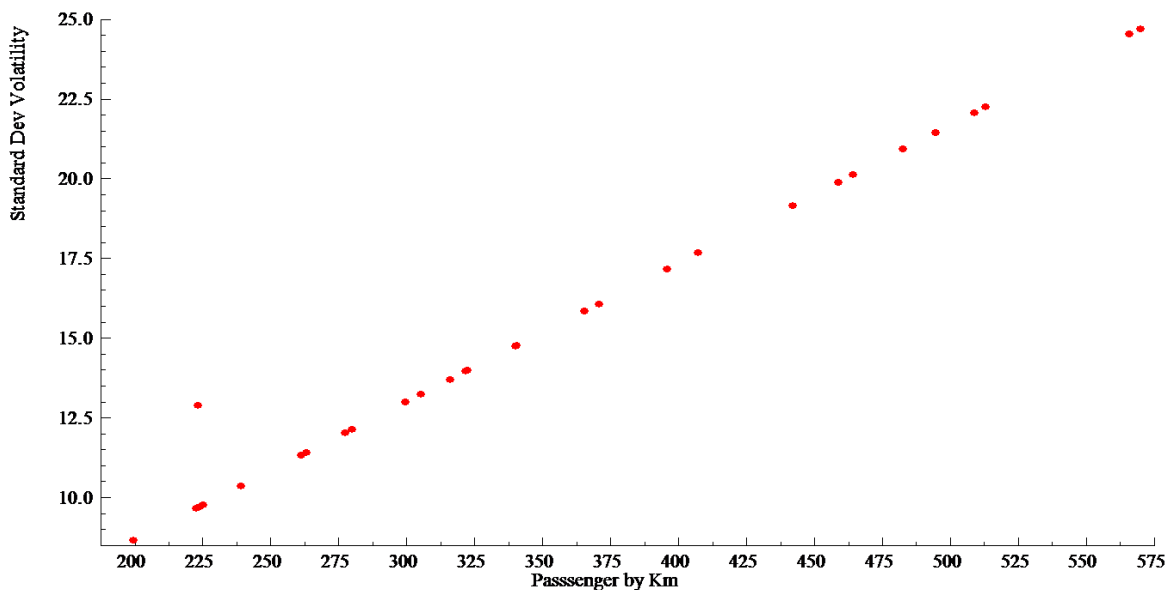


Figure 4. Estimated values of  $\sigma \exp\left(\frac{1}{2}\theta_t\right)$  (vertical axis) together with the observed values of the passengers by kilometers series for the period from 1889 to 1920.

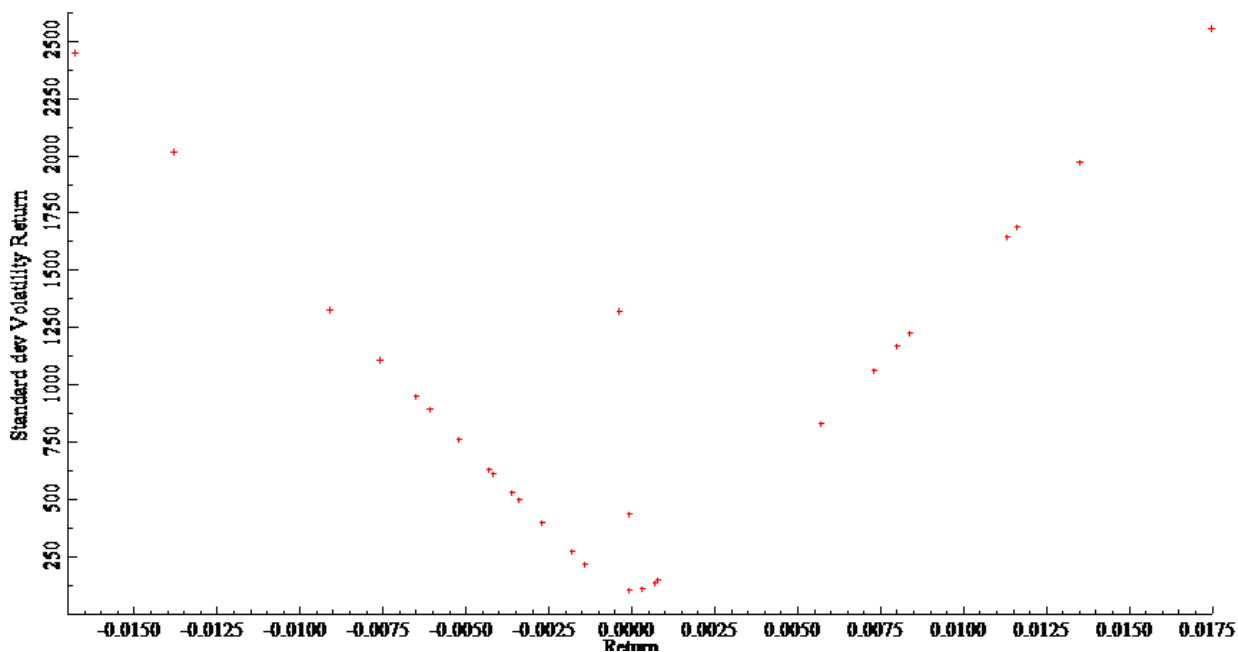


Figure 5. Estimated values of  $\sigma \exp\left(\frac{1}{2}\theta_t\right)$  (vertical axis) together with the observed values of the return series for the period from 1889 to 1920.

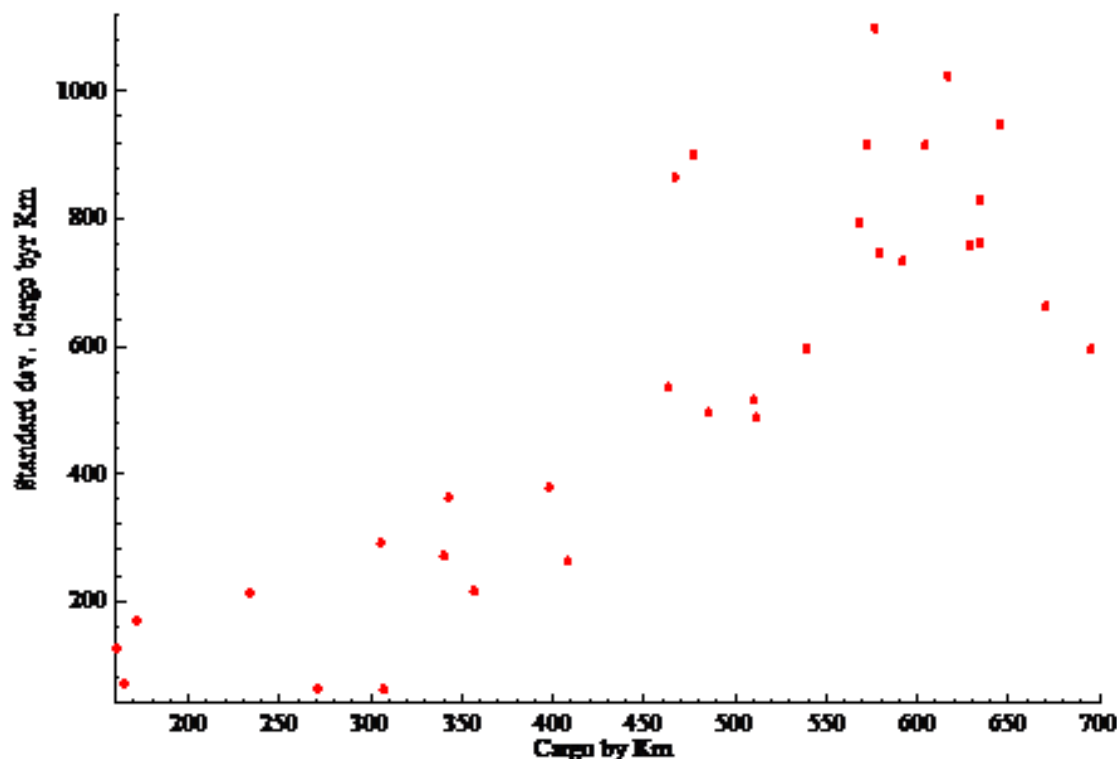


Figure 6. Estimated values of  $\sigma \exp\left(\frac{1}{2}\theta_t\right)$  (vertical axis) together with the observed values of the cargo by kilometer series for the period from 1889 to 1920.

### Discussion

The old proverb “A picture is worth a thousand words” is pretty accurate for the analysis of any information set. Before we apply any statistical method we need to observe the data set in hand in order to become familiar with them. This can have a number of benefits, because this process will serve as an indicator for better and more detailed study ideas.

In Argentina, the emergence of the rail was very different from the European nations, where it brought a new economic and social order as well as the improvement of the transport industry. The railways in our country can be better described as an element of political evolution. They were without the shadow of a doubt the sole factor that beat the large distances that isolated the population and made an illusion the idea of national unity.

Like most of the lines built in our Argentina, as well as the rest of Latin America, the North Central Railway was more the result of a previous demand than the generator of a new one. The state funded lines absorbed more than 30% of the public work budget without taking into account the fact that its own resources could had been reinvested. We can say that despite the positive impact that the railways had in our country, they suffered from bad management and chronic lack of financial resources that eventually affected its functioning. The initial investment to set this enterprise was huge. Nevertheless, the earnings were not as expected and the expenses derived from its maintenance were extremely high.

On the statistical point of view we see that volatility is the main feature in all the series we analyzed, that

is, there is a conditional variability that changes quickly in time. This is a confirmation of everything we said and analyzed from the historical point of view, that is, there was not a proper management policy of this line. Everything was done with a strong sense of promotion and it was tied with the financial availability as well as the existing credit in every period of time, according to the international situation. This explains why the benefits were very little and in many cases this line was forced to work at loss.

### References

- ABRIL, JUAN CARLOS. (1999), *Análisis de Series de Tiempo Basado en Modelos de Espacio de Estado*, EUDEBA: Buenos Aires.
- ABRIL, MARÍA DE LAS MERCEDES. (2014). *El Enfoque de Espacio de Estado de las Series de Tiempo para el Estudio de los Problemas de Volatilidad*. Doctoral Thesis. Facultad de Ciencias Económicas. Universidad Nacional de Tucumán. Argentina.
- BOLLESLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-27.
- CUCCORESE, HORACIO. (1984). *Historia de los Ferrocarriles en la Argentina*. Editorial Macchi: Buenos Aires.
- DIARIO "EL ORDEN" de la provincia de Tucumán, (1900 and subsequents).
- DURBIN, J. Y KOOPMAN, S. J. (2012). *Time Series Analysis by State Space Methods (2nd Edition)*. Oxford University Press: Oxford.
- ESTADÍSTICAS DE LOS FERROCARRILES EN EXPLOTACIÓN. (EJERCICIO 1936 - 1937). (1942). Tomo XLV, Ministerio de Obras Públicas de la Nación. Talleres Gráficos del Ministerio de Obras Públicas: Buenos Aires.
- FERMS, H.S. (1960). *Britain and Argentina in the nineteenth century*. Oxford University Press. London.
- GHYSELS, E., HARVEY, A. C Y RENAULT, E. (1996). Stochastic volatility. In C. R. Rao y G. S. Maddala (eds.), *Statistical Methods in Finance* (pages 119-191). North- Holland: Amsterdam.
- HARVEY, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press: Cambridge.
- HAUSCHILD, HENRY (2013). Joseph Telfener. *Handbook of Texas Online*. Texas State Historical Association. Texas.
- HOBSBAWM, ERIC J. (1982). *Industria e Imperio*. Ariel: Barcelona.
- KALMAN, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82 (1), 35 – 45.
- KOOPMAN, S. J., HARVEY, A. C., DOORNIK, J. A. Y SHEPHARD, N. (2000). *STAMP. Structural Time Series Analyser, Modeler and Predictor*. Timberlake: London.
- LÓPEZ, MARIO JUSTO; WADDELL, JORGE E. (2007). *Nueva Historia del Ferrocarril en la Argentina. 150 Años de Políticas Ferroviarias*. Lumiere: Buenos Aires.
- MEMORIA DE OBRAS PÚBLICAS. (July 1901-October 1904) (July 1916-May 1917) (June 1919-May 1920). Technical Library . Ministerio de Economía de la Nación. Buenos Aires.
- MINISTERIO DE OBRAS PÚBLICAS DE LA NACIÓN. (1910 and subsequents). *Estadísticas de los Ferrocarriles en Explotación*. Tomo XIX. Imprenta Guillermo Krieguer: Buenos Aires.
- SALERNO, ELENA. (2007). La evolución y los problemas de los Ferrocarriles del Estado durante la primera mitad del siglo XX. In *Nueva Historia del Ferrocarril en la Argentina. 150 Años de Políticas Ferroviarias*. Lumiere: Buenos Aires.
- SOARES, ERNESTO. (1937). *Ferrocarriles Argentinos. Sus orígenes, antecedentes legales que los rigen y reseñas estadísticas*. Compañía Impresora Argentina, S.A.. Buenos Aires.
- SCHICKENDANTZ, EMILIO Y REBUELTO, EMILIO. (1994). *Los Ferrocarriles en la Argentina 1857-1910*. Fundación Museo Ferroviario. Buenos Aires.
- SHEPHARD, N. (1996). Statistical aspects of ARCH and stochastic volatility. In D. R.Cox, D. V. Hinkley, and O. E. Barndorff-Nielsen (eds.), *Time Series Models in Econometrics, Finance and Other Fields*, (pages 1-67). Chapman and Hall: London.
- WRIGHT, WINTHROP R. (1980). *Los Ferrocarriles ingleses en la Argentina*. Emecé. Buenos Aires.