

01-introduction

October 27, 2021

1 Introduction

1.1 COVID-19 Effects and Mobility Habits in Italy

The COVID-19 pandemic set the world back in its pursuit of the seventeen Sustainable Development Goals (SDGs) (The 17 Goals, n.d.). This is also the case for Italy, as we can see from the 92 statistical measures collected by the Italian national institute for statistics (ISTAT, Istituto Nazionale di Statistica). When comparing 2019 with data from ten years prior, sixty percent of the measurements register an improvement, while twenty percent do not change and the remainder worsens. However, when comparing 2019 and 2020, only forty percent of the measurements improve, and almost the same amount displays a statistically significant drop (Dall’andamento degli obiettivi di sviluppo sostenibile alla mobilità, 2021).

To say that the pandemic brought about a lot of change would be a redundant understatement. Coming out of 2020, the Italian outlook is worse than the average of the (Western) European countries - as one can notice upon looking at macroeconomic variables such as the Gross Domestic Product (GDP). Part of the negative consequences could be offset, thanks to the government’s stimulus packages, the European Union’s (EU) funds and the European Central Bank’s (ECB) Pandemic Emergency Purchase Program, or PEPP - a massive asset purchase program of 1,850 billion euros. From the beginning, it was clear that the pandemic would accelerate digitalisation (Härmand, 2021) - and, indeed, it has: see (Anthony Jnr & Abbas Petersen, 2021) for a meta-review or (Truant et al., 2021) for Italy. The generalised lockdowns reduced the greenhouse gases (GHG) emissions (see (Jephcote et al., 2021), (Singh & Chauhan, 2020), (Lian et al., 2020) and (Gualtieri et al., 2020)) but only temporarily and relatively to the severity of the restrictions put in place: the pandemic altered electricity consumption habits, but in countries like Sweden “the consumption even increased” compared to 2019 (Bahmanyar et al., 2020). Besides, this also came at the expense of our transportation habits and the mobility sector in general.

According to ISTAT, private cars are now used for half of all travels, up from 44 percent in the pre-pandemic period. Furthermore, the so-called “mobilità dolce” (the Italian translation of micro-mobility) “does not seem to take off” (Dall’andamento degli obiettivi di sviluppo sostenibile alla mobilità, 2021). In 2020, 30 percent of interviewed families stated that they had “some or a lot of difficulties in connecting with public transportation in the area where they live” - an improvement from the 33 percent of the previous year. However, it is enough to look at the regional level to realise that improvements are not spread evenly, and some areas of the country are actually worse off with respect to 2019. There is a great deal of heterogeneity: the share of families with troubles in accessing public transportation services is lower in the North (26 percent) compared to the South (36 percent), while in Campania more than half of the families are affected (51 percent). Furthermore, only 27 percent of students reach the places where they study with public means of

transportation (and it's declining), while 75 percent reach the workplace with private means only (and the share is increasing) (Dall'andamento degli obiettivi di sviluppo sostenibile alla mobilità, 2021).

The only measurement that has been improving throughout the decade is that of air quality. However, these values remain significantly higher than the guidelines from the World Health Organisation (WHO) (Dall'andamento degli obiettivi di sviluppo sostenibile alla mobilità, 2021). Besides, on September 22, 2021 the WHO revised downwards their recommendations on air pollutants (the latest update dated back to 2005), and the European Commission (EC) declared that it will take this update into account when revising the guidelines for the upcoming year (Standards - Air Quality - Environment - European Commission, n.d.). Italy has always struggled to respect the European Directive on air pollution: on November 2020, the European Court of Justice (ECJ) judged that, from 2008 to 2017, Italy “systematically and continuously” violated the standards set by the Directive and failed in enacting countermeasures to avoid it (“Che Aria Respiriamo,” 2021). On July 2021, the European Environmental Agency (EEA) ranked European cities according to the average level of $PM_{2.5}$ in the two previous years: out of 323 cities, Cremona (Lombardy) was second to last, Brescia and Pavia (both in Lombardy) were 315th and 314th respectively, while Venezia (Veneto) was 311st, Bergamo 306th and Milano 303rd (all in Lombardy) (European City Air Quality Viewer — European Environment Agency, 2021). Air pollution caused more than 52 *thousands* premature deaths: almost 14 percent of the total premature death toll in Europe (Italy - Air Pollution Country Fact Sheet — European Environment Agency, 2020).

Unequal access to infrastructure (and, to a minor degree, air pollution) hinder the popularity of more sustainable mode of transport, such as biking and sharing services. The ISTAT report mentioned above estimated that in 2019 there were 30 million commuters in Italy (Gli spostamenti per motivi di studio o lavoro nel 2019 secondo il Censimento permanente della popolazione, 2021): more than two thirds (more than 20 million) need to reach their workplace, while the rest is made up of students (with a moderate variance: the share of students is higher in those regions where unemployment is higher, for example in Campania, where they make up 40% of commuters). However, the report did not present any data on biking habits or the use of sharing services. These can only be found in an earlier report, published in 2019 and referring to two years prior.

The 2019 report mentions two interesting statistics about commuting habits: “almost one in five [commuters] choose an ‘active’ mode of transport” - that is, walking or biking. However, most of the “active” commuters are actually walking to work (17,4%), whereas the bikers are only some 1,7%. In general, it is women, young and more educated people who use more public transportation means and bicycles, while private vehicles (the exclusive means of commuting for more than 73 percent of the employed) are spread among men between 25 to 44 and with an average level of education (Spostamenti quotidiani e nuove forme di mobilità, 2019). Car pooling is chosen by some 12 percent of the employed and 14,5 percent of students aged 18-24, while only less than half a million used bike sharing services at least once during the year (i.e., less than 2 percent). Such services are more popular across the young and more educated people, while the incidence is almost double the national average in metropolitan cities (Spostamenti quotidiani e nuove forme di mobilità, 2019).

1.2 Bike-Sharing Demand Determinants and the State of Public Transport Infrastructure in Italy

The picture drawn by the two ISTAT reports feels like a pool of untapped (if not wasted) potential. According to the most recent survey, in 2019 some 57,5% percent of commuters moved within the

same municipality of residence. This value is driven up by the students, who make up almost 71 percent of commuters within the same municipality. However, even after taking them out of the computation, we still end up with an even figure: more than 51 percent of workers move within the same municipality (Gli spostamenti per motivi di studio o lavoro nel 2019 secondo il Censimento permanente della popolazione, 2021).

Sure, it would be naive to argue that commuters who work in the same municipality where they live could all bike to reach their destinations: after all, there is a great deal of heterogeneity across municipalities under several dimensions - like their sheer extension, morphology and, of course, infrastructure. There are many factors that affect bike sharing demand: the first one that comes to mind are the weather conditions: precipitations, humidity and seasonal patterns; the one that arguably plays the biggest role is the so-called “built environment” (Eren & Uz, 2020), i.e. infrastructure such as the availability of isolated or dedicated bike lanes instead of mixed ones, but also safe parking areas and bike racks for private bikes. The terrain clearly plays a role: slopes have a negative effect on bike usage (as one of the many examples, see (Bordagaray et al., 2016)), but incentive schemes can be devised to promote returning bikes to up-hill stations and even the least loaded ones (which also improves the overall efficiency of the system) (Fricker & Gast, 2016). Furthermore, in this scenario there is a positive effect of e-bikes. Besides, the literature also outlines the role of land use: pick-ups are more frequent in commercial areas and parks, compared to residential ones and, more broadly “the proximity to green spaces and recreation areas, schools, universities, museums, shopping centers, sports areas, restaurants, hotels, bus/subway/train/suburban/ ferry transit hubs has a positive effect on the use of BSP [Bike Sharing Programs]” (Eren & Uz, 2020).

The degree of integration with the public transportation is also important, as bike-sharing systems are found to be complementary means for “[bridging] the gap between multiple transit hubs” (Eren & Uz, 2020). But bike sharing can also be a substitute “especially when public transport is not available, between 22:00 pm [and] 06:00 a.m., they can encourage users to use BSP” (Eren & Uz, 2020). This implies that there is no single channel to promote bike-sharing services and that coordination across institutional players is crucial. Despite this, investments in bike-sharing services shall not fall in the background: their complementary role as “first/last mile solution” is recognised in the literature and enhance public transport as a whole.

Improving the public transport infrastructure is a priority for Italy. As always, there are territorial imbalances on two different dimensions: on the macro level, there is a clear divide between North and South, but there continue to be striking differences even within the wealthiest regions. Bike lanes have been increasing steadily: the total number of kilometres has grown by 15,5 percent since 2015, totalling approximately 4700 kilometres. However, the infrastructure is still far from adequate in most cities (Ambiente urbano, 2021).

The report from ISTAT outlines that public transport (or TPL, “trasporto pubblico locale”) suffers both from lack of infrastructure and outdated fleets. As a starter, the TPL is over reliant on buses, which offer more than 55 percent of the number of seats per kilometre. However, once metropolitan cities are factored out, this figure skyrockets to well beyond 90-95 percent (Ambiente urbano, 2021). Only 32 percent of the bus fleets is in line with Euro 6 standard and some 34 percent belongs to the Euro 4 class - i.e., was deployed before 2008. Low emission buses make up some 28 percent of the total, but only slightly more than 3 percent are electric: the rest (almost 25 percent) is fuelled by natural gas. Unsurprisingly, the share of low-emission vehicles is higher in metropolitan cities.

Trolley buses are available in only 13 municipalities, trams in 11 and metropolitan trains in 7. However, there is a remarkable divide between the Italian champion, Milan, and the other cities.

Tram network density in Milan is measured as 122 km per 100 squared kilometres; the silver medal is awarded to Turin, which has almost half the kilometres than Milan: 66. The average of the other cities is a mere 16 km. In general, while the supply of TPL (measured in seats per kilometre, per inhabitant) is on the rise, we are still far from the levels before the Great Recession (-7,3 percent compared to 2008). However, the supply in the North is 25 percent greater compared to the Centre and almost three times bigger than the South. Public demand for TPL is increasing in the North, stationary in the Centre and even declining in the South.

A minor and uneven push is provided by sharing services. Car sharing is available in 37 out of 107 “comuni capoluogo” (i.e., the “capital” of a province, corresponding to the NUTS3 classification), of which only 8 are in the South. Besides, only 26 percent of the fleets is composed of electric cars. Bike sharing services are present in 53 *capoluoghi*, registering an overall decline from 2015. Luckily, the number of bikes has more than tripled: from 6 to 19 bicycles per ten thousand inhabitants. The divide, as always, is quite stark: the number of bikes is 29 in metropolis compared to provinces, and these services are much more common in the North (32 bikes per ten thousand citizens) than in the Centre (17) and the South (just 2). Much of this success is to be attributed to the appearance of free-float systems, which require greater fleets (Ambiente urbano, 2021). The report does not provide information on electric scooters.

Restructuring the public transport will require extensive coordination between national, regional and municipal administrations, across multiple channels simultaneously. It is a widespread hope that many of this results can (only) be achieved via the Next Generation EU (NGEU), the 750 billion euros stimulus that will be financed by bonds from the European Commission. The NGEU is a bold and unprecedented move from the EU: the fund will be made up with up to 390 billion euros in subsidies, while up to 360 billions will be given out as loans with low interest rates. Italy is the first beneficiary in absolute terms for the main facilities of the NGEU: the country will receive more than 190 billion euros, to which the government will add 30 billions of its own. According to the so-called PNRR (*Piano Nazionale di Recupero e Resilienza*, i.e. National Plan of Recovery and Resilience), almost 25 billion euros will be invested in railways. However, less than one billion will be used to improve on the regional railways - the Achilles’ heel of public transportation and the bane of commuters. In addition, more than 8,5 billion euros will be invested in TPL. But there’s a catch: the NGEU grants will only be available for projects to be completed within the year 2026. Interviewed by *Il Post*, prof. Gabriele Grea from Bocconi University stated that these funds will mostly be awarded to projects in an already “advanced state”: on one side, this will provide stronger guarantees about their completion, but will likely increase the inequalities across municipalities (Un po’ di cose notevoli dentro il PNRR, 2021).

1.3 The Case for Promoting Bike Sharing

Reforming the public transport will be crucial to reach carbon neutrality and possibly promote economic growth. After all, transport accounts for as much as 27 percent of emissions in the EU (Bergantino et al., 2021) and while the overall greenhouse gases (GHG) emissions has been declining since the 1990s, the emissions from road transportation has nonetheless been increasing ever since (Annual European Union Greenhouse Gas Inventory 1990–2018 and Inventory Report 2020 — European Environment Agency, 2020). Besides, it is well-known that Italy chose to privilege rubber over railways to transport goods: Eurostat estimated that from 2000 to 2016 less than 10 percent of goods travelled by train, while the EU average was almost 18 percent (Milano ha un’occasione storica, 2017), so there seems to be much room to increase productivity. And, besides, despite the fact that emissions have been decreasing since the 1990s, Italy is not on the side of the achievers:

since 1990, emissions in the country were reduced by 17,2 percent, compared to the EU28's 25,2 percent.

The Next Generation EU provides Italy with the perfect chance to narrow the divide with Western economies, while curbing emissions and finally improving the air quality. To promote more sustainable modes of transport, measures are needed on both the supply side (for example, by improving vehicle and fuel performance) and on the demand side by reducing demand for private transport, or at least increasing the demand for greener modes of transport (Bergantino et al., 2021). Investments in sharing services can and should play a role in this transition. Indeed, capital might be limited: after all, municipalities will receive a smaller share of the NGEU funds and upgrading their bus fleets seems more urgent. Besides, there are also the time constraints that need to be taken into account. However, biking infrastructure projects can be relatively cheaper compared to other TPL investments - especially if factoring in the presence of private entrepreneurs. Once infrastructure is in place, the costs of the service “only” amount to the human and technical cost to reallocate bikes to be at the right place at the right time.

This dissertation stems from the idea that bike sharing systems can be promoted with cheap measures. One of the most widely discussed problems in the literature is improving customer satisfaction through repositioning, i.e. forecasting the demand for bikes and “design efficient bike repositioning solutions” (Ghosh et al., 2019). The problem is ever more important since the introduction of free-float bike-sharing systems (FFBBS), as the bikes can be dropped anywhere and end up in sub-optimal places for the next customers. On the positive side, FFBBS do not require the upfront investments for building docking stations - which is necessary for station-based BSS (also known as SBSS). Furthermore, FFBSS “prevents bike theft”, “by tracking bikes in real-time with built-in GPS”, and “offers significant opportunities for smart management”(Pal & Zhang, 2017). This implies a greater satisfaction level for customers, “because obtaining and returning the bikes becomes much more convenient” (Pal & Zhang, 2017). However, all of this comes at the increased costs for bike rebalancing, because of how inefficient bike redistribution becomes and the and higher operating costs in terms of human and financial resources (Pal & Zhang, 2017). This has been recognised as “one of the main reasons why many FFBS enterprises lose money or even withdrawn from market.”(Tian et al., 2021).

Most of the new approaches involve deep learning (DL) techniques, such as long-term short-memory (LSTM) neural networks, which usually outperform other statistical and machine learning approaches (Xu et al., 2018). Some are incredibly sophisticated: Convolutional LSTM are deep learning models “stacked and fused by multiple convolutional long short-term memory (LSTM) layers, standard LSTM layers, and convolutional layers [...] to better capture the spatio-temporal characteristics and correlations of explanatory variables” (Ke et al., 2017). This, combined with external data such as “travel time rate, time-of-day, day-of-week, and weather conditions”, results in an improvement of error metrics (RMSE) by almost 50 percent (Ke et al., 2017).

There is no way we could best the performance of such sophisticated models. However, this may not be the ultimate goal for policymakers. Despite their impressive performances, deep learning methods are hard to implement. They require a considerable amount of resources and, unlike simpler models, are much more difficult to visualise and interpret. For one, they require senior data scientists and access to a local server cluster with access to Graphic Processing Units (GPUs) or cloud computing platforms such as Google Cloud, Microsoft Azure or Amazon Web Services. This infrastructure needs time to set up and delays are inevitable, since the data manipulated by public administration deserves a much greater degree of privacy. Besides, such models require a long time to train - which translates to more expensive models.

Given the policymaker constraints, DL techniques may well be out of time, and budget. Our goal is to satisfy the constraints of a local planner with tight budget and more pressing issues, or the limited options of a private BSS company who needs to sustain high operational costs. We will develop two classes of models, univariate and multivariate, using both statistical and machine-learning methods. We will also attempt to evaluate the usefulness of external data - which might be hard to require and process - and the performances of libraries such as Facebook’s Prophet, which have been specifically designed for ‘forecasting at scale’, i.e. forecasting multiple time-series (in the order of the thousands) with little to no pre-processing, feature engineering and simple models.

This experiment has many drawbacks and limitations: as a starter, it does not include a proper spatial analysis, and does not explore Vector Auto Regressive (VAR) models, nor Bayesian models. Our goal is to prove that feature engineering and better data can do a better job at improving a model compared to more advanced techniques, especially under (hypothetically) tight budget constraints. In other words, we might want to get the feel for the marginally decreasing utility of accuracy improvements, which come at progressively greater computational costs.

1.4 The Open Source Stack, Transparency and Reproducibility

This is one of the many (millions?) projects to benefit from the existence of the open source community. This project was developed end-to-end using open source tools, starting from PostgreSQL to store the data and many Python libraries to train the models. Jupyter Notebooks have been the main developing tool (Perkel, 2018), and Jupyter Books (Executable Books Community, 2020) to convert the code into a \LaTeX publication, thanks to `pandoc` working in the background (Pandoc - About Pandoc, n.d.).

Zotero was used as a bibliography manager (Zotero | Your Personal Research Assistant, n.d.), and merits are due to several extensions for JupyterLab that made writing on JupyterLab possible: in particular, `jupyterlab-citation-manager` was used to insert citations inside notebooks via the official Zotero API, and `jupyterlab-spellchecker` helped in spotting typographic mistakes. Of course, version control with `git` and hosting on GitHub played a crucial role in project management. The code is free to see on the dedicated GitHub repository; however, data cannot be accessed due to the terms of the partnership between the provider and the University of Milan. We will try to find the space to introduce and contextualise all the other open source libraries that have been actively used.

There would be much to say about why using open source software, even (and especially) in economics. When it comes down to open source against proprietary software, the differences are not merely technical:

There is an independent social dimension, where the metrics assess the interactions between people. Does it increase trust? Does it increase the importance that people attach to a reputation for integrity?

This is the Economics Nobel Prize Paul Romer (Romer, 2018), now 65, when comparing his experience with Jupyter and Mathematica notebooks. It goes on to make a bold claim:

Jupyter exemplifies the social systems that emerged from the Scientific Revolution and the Enlightenment, systems that make it possible for people to cooperate by committing to objective truth; Mathematica exemplifies the horde of new Vandals whose pursuit of

private gain threatens a far greater public loss – the collapse of social systems that took centuries to build.

When Romer tried to work with Mathematica notebooks and share their results, it became clear that “Wolfram made it hard to share a readable PDF version of a [Mathematica] notebook because it wanted someone like me to distribute content in its proprietary file format, the CDF”. The conclusion of his articles are quite dramatic:

The tie-breaker [between Wolfram and Jupyter, as well as proprietary and open source] is social, not technical. The more I learn about the open source community, the more I trust its members. The more I learn about proprietary software, the more I worry that objective truth might perish from the earth.

Jupyter Notebooks are being developed since 2001 and they are ever more popular. Perhaps they might even replace the scientific paper [NO_PRINTED_FORM]; what’s sure is that they are ever more present. Indeed, Romer is not the only Nobel prize using open source software: Thomas Sargent, currently 78, uses Julia for his scientific research and launched a website, [QuantEcon](#) built with Jupyter Books, to teach computational economics in Python and Julia.