Università Ca'Foscari Venezia

Master's Degree in

Economics and Finance

Second cycle (D.M. 270/2004)

Final Thesis

Real-Time and Latest Available data:

Do Revisions affect Unemployment Forecasting?

Supervisor
Ch. Prof. Irene Mammi

Graduand
Oscar Magnabosco
Matriculation Number 862328

Academic Year
2020 / 2021

Ai miei nonni.

ABSTRACT

Forecasting with macroeconomic variables brings about several challenges, one of them being the presence of data revisions. When new sample information is available to National Agencies, economic estimates are revised because macroeconomic variables are characterised by hard-to-process samples that need to be often updated as new data is gathered by governments. The thesis focuses on the relationship between revised (i.e., latest available) and unrevised (i.e., real-time) data, the role of revisions is assessed with respect to unemployment forecasting. The estimating sample consists of 28 countries, and unemployment data has been gathered from the OECD's Economic Outlook issues from 1996 to 2019. Regarding the work's structure, a description of the basic introductory concepts is followed by a preliminary analysis, aiming at shedding light on the impact of data revisions and whether such revisions translate into lower time-series volatility. The analysis carries on with an empirical exercise, developed using STATA, in which unemployment data is forecasted for the years 2019, 2020, and 2021 and then compared to the OECD's forecasts. Ultimately, the thesis aims to examine if, and if so, how unemployment forecasting is affected by data revisions.

I. FORECASTING WITH MACROECONOMIC VARIABLES

Forecasting with macroeconomic variables brings about several challenges which, generally, are not faced by financial variables' forecasters. This introductory chapter aims at shredding light on what is worth considering before starting any macroeconomic analysis.

1.1. Introductory concepts on data revisions

Before carrying out any forecasting, forecasters need to choose a proper data sample which will serve as a milestone for their subsequent analysis. In most cases, such sample selection entails gathering data from current databases, meaning data containers providing all information available up to the point in time the download is processed: this is what usually happens with financial variables (e.g., stock prices). Economic variables, such as gross domestic product and inflation among the others, present nevertheless a particular feature which in most cases may lead to misleading forecasts: the presence of revisions. Revisions deeply affect samples by giving rise to real-time and latest available data; for this reason, they need to be carefully accounted for by macroeconomic forecasters.

1.1.1. Issues with macroeconomic data samples

Revisions take place when governments' national agencies update previously released economic data. There are two main reasons which may compel national agencies to revise their numbers:

- First, governments initially release economic data shortly after the period related to such data, usually after one or two months. In an ideal national statistical system, such freshly released data would be complete and correct: high-quality interaction and coordination among the data management parties, incorporation of only reliable and relevant information and the development of efficient feedback systems to suppliers of data would ensure excellent economic data quality, which would not need any subsequent adjustment. Unfortunately, it is reasonable to assume that real-world governments' national agencies are not that efficient, thus causing longer processing times of huge quantities of samples. For this reason, national agencies update (or, better, revise) previously released data and keep on doing so month after month and year after year as samples become more and more complete, enabling countries to get increasingly precise estimates of the variables they are trying to measure [Fajingbesi (2001)];

- Second, many economic variables (i.e., the real ones) are measured according to a base year which is usually changed every 5 or 10 years; when such benchmark is modified, all formerly released data must be adapted to the new base year accordingly. Such benchmark adjustments most often lead to major data alterations which sensibly hinder data comparability [Bassanetti, Caivano and Locarno (2010)].

When revisions do take place and data are revised, an economist will find herself dealing with two different numbers for the same variable: an original value of a variable of, say, 2019 and a revised value of the same variable of 2019, published a month or a year later. What value will she choose to carry on with her forecasting?

The answer to the question is not straightforward. First of all, the economist would have to decide if she were trying to forecast a government's data first-release or, instead, some kind of "final" value which would only appear after many years of revisions. After doing so, she would have to compare the forecasted value to some benchmark, which would entail deciding which data vintage to use and set it as the benchmark: an arbitrary choice. Naïve analysis usually translates into using revised data available at that moment; such practice, while being simple and handy, could also potentially be misleading if there have been alterations in calculations' methodology which are being ignored by the forecaster. Better practice would rely on the usage of a mixture of data vintages: by combining first-release data, data which have been revised a year or so later (e.g., in order to address national agencies' short-term measurement inefficiencies) and data just prior to benchmark changes (e.g., allowing the economist to get the very best estimate before the benchmark alteration) a forecaster would certainly get a much more complete picture [Clements and Galvao (2017)].

1.1.2. Explaining real-time and latest available data: an example

Let us make an example: shortly after the end of the year 2013, the statistic agency of the Italian government (i.e., Istat) releases its brand-new macroeconomic data (e.g., gross domestic product, unemployment rate etc.) regarding the year that just passed by. Samples underlying such first data release could not be further from being final, though: processing huge quantities of information does take time, companies' annual reports have not been released yet and as previously seen, there can always be measurement errors and inaccuracies. For this reason, one year later the Istat revises its 2013 numbers and keeps on doing so each following year until 2019. Similarly, the agency performs the exact same revision practice for its 2014 macroeconomic data, for its 2015 one and so on, up until 2019.

If someone were to place such numbers on a table, it would look like the table below, which shows each unemployment rate value's release date for Italy from 2013 up until 2019, and how such data evolves over time as it gets revised by Istat. Table columns show data the economist would observe if she used the database in the year pointed out in the column header. Rows, on the other hand, represent the years in which the economic activity is measured by the national agency (e.g., for year 2013 the first data released took place in 2013 and was equal to 12.09%, one year later, in 2014, as new information has been gathered such number has been revised upward to 12.21%, and so on until 2019, when the latest revised value of 2013 can be found with a slightly higher value compared to the initial release).

By looking at the table, the last column represents the set of latest available data in 2019, meaning unemployment data from 2013 to 2019, revised up until 2019. On the contrary, data on the main diagonal (underlined numbers on the table) represent real-time data, meaning data referring to the same year the estimate was first released, without any revision (e.g., the 2019 unemployment rate value is both real-time and latest available as (1) it has been issued for the first time in 2019, meaning it has not been revised yet and (2) it is also in the latest available set of data at the economist's hand).

Table 1 - Unemployment rate in Italy from 2013 to 2019

Vintage Date (June):	2013	2014	2015	2016	2017	2018	2019
Year							
2013	***12.09***	12.21	12.16	12.11	12.12	12.13	*12.13*
2014	12.45	***12.44***	12.67	12.64	12.64	12.63	*12.62*
2015	12.09	12.30	***12.28***	11.91	11.91	11.91	*11.91*
2016		12.12	11.68	***11.50***	11.65	11.66	*11.67*
2017			11.04	11.04	***11.18***	11.26	*11.26*
2018				10.69	10.48	***10.43***	*10.63*
2019					10.11	9.72	***10.03***

Data from OECD (2013-2019), Economic Outlook Annual reports, https://www.oecd.org/economic-outlook/, personal reworking

It is now worth pointing out that the above-mentioned data has not been subject to any benchmark revision. The differences in the numbers from one data vintage to the other stem solely from incorporation of new information and/or noise reduction.

1.1.3. Real-time databases availability

The previous discussion exploited a sub-sample stemming directly from the real-time database (which is going to be thoroughly described in the next chapter) specifically put-together for the thesis work. If the economist were to decide, as previously shown, which data to choose and did not have any database at hand, she would have to either retrieve it from the internet or create it by herself. There exists indeed a matter of availability around such databases, as real-time data uses different data vintages which have to be put together and re-managed in order to be ready: this makes such databases quite rare; on the other hand, latest available data use only one vintage which, by construction, do not need further work.

Without real-time databases being readily available, it is unlikely that an analyst would even considering taking them into account and would simply end up resorting to latest available ones (i.e., using the latest column of Table 1). This, in turn, would mean neglecting the impact of data revisions in the analysis, with all the risks associated with it, which have been formerly outlined in *section 1.1.1*.

1.2. Modelling data revisions

One may think that data revisions are simply negligible adjustments devised to purify numbers from measurement mistakes and inaccuracies occasionally made by national agencies. Some studies [Holden and Peel (1982)] on the nature of revisions based on United Kingdom's macroeconomic data, though, prove them to be much more than that: it has been shown that time-series properties of revised data are different compared to unrevised one (i.e., first-release data); as a consequence of it, the data generating process underlying the two seems to be different as well, leaving room for further investigation. This sub-chapter delves into data revisions and attempts at grasping their nature.

1.2.1. A simple model

Revisions, to be modelled, do not necessarily need complex equations. Consider the following:

$$y_t^v = y_t^* + \varepsilon_t^v$$

Where y_t^* indicates the true value of variable y at time t, y_t^v is the value of the national agency's released data vintage in year v and ε_t^v is the error term of the equation which is dependent on statistical agencies' measurement and reporting assumptions. Now, it is apparent that the very true value of variable y is unknown (e.g., one can only have the first released value and the revised ones in the successive data vintages); as such, the previous equation could be rewritten as follows:

$$y_t^v = y_t^{v-1} + r_t^{v,v-1}$$

Here, y_t^v is the revised value in vintage v, y_t^{v-1} is the previously revised value in the vintage $v-1$ and $r_t^{v,v-1}$ is the data revision from vintage $v-1$ to vintage v. The latter represents the very fulcrum of the model that, if properly discerned in all its properties, can give useful insights which can turn out to be primary and fundamental in macroeconomic analysis and forecasting [Cimadomo (2008)].

1.2.2. Going deeper

In order to see how much revisions impact forecasts, one might carry out an analysis using two different models, with the first relying on real-time data and the second on latest available one. The discrepancy in the resulting forecasts would be proof of the significance of revisions.

One good method [Stark and Croushore (2002)], which is used in reality, stems directly from one of the most comprehensive studies on the subject and provides for many forecasts, each with the usage of a different vintage of data, to see how much different jumping-off points affect forecasted results. This method, called "Repeated Observation Forecasting" has provided researchers with remarkable findings: forecast results strongly differ from each other, with differences that go beyond standard forecasting uncertainty; further, not only it is argued that revisions do matter, but also that revisions might be the primary source of forecast uncertainty.

As an example, let us take the following model:

$$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t$$

Here, it is assumed for the sake of simplicity that the model follows an Auto Regressive Process of Order 1 (i.e., AR(1)). Just like it was argued in section 1.2.1., y_t represents the true value of the variable we are trying to measure which, unfortunately, is unknown; instead, y_t^v is used with v being the release vintage date, while $r_t^{v,v-1}$ is the revision from one vintage to the next. Now, let us assume that the economist wants to start forecasting using (1) the AR(1) process previously shown, (2) vintage data v and (3) a dataset having data dating back to $t-1$. The one-step ahead forecast would be the following:

$$y_{t|t-1,v} = \widehat{\mu_v} + \widehat{\phi_v} y_{t-1,v}$$

On the right-hand side of the equation the estimated coefficients can be found, which belong to the vintage of data v. Similarly, let us now consider a subsequent vintage of data, w, which generates the following one-point ahead forecast:

$$y_{t|t-1,w} = \widehat{\mu_w} + \widehat{\phi_v} y_{t-1,w}$$

Interesting insights can be gained by differencing the previous two equations, the result is as follows:

$$y_{t|t-1,w} - y_{t|t-1,v} = (\widehat{\mu_w} - \widehat{\mu_v}) + (\widehat{\phi_v} y_{t-1,w} - \widehat{\phi_v} y_{t-1,v})$$

This rearrangement shows what happens when we difference revised data in two different data vintages, v and w; thanks to it, the user is able to spot the two sources of data revision impacting on forecasts: changes within the data (i.e., they make $y_{t|t-1,w}$ differ from $y_{t|t-1,v}$), and changes within the coefficients.

Summarising, if revisions are small and negligible, the difference in the equation above will tend to zero as well; if, on the other hand, revisions are large and significant meaning that they impact heavily on the coefficient estimates and/or on the variables within the equation, then such difference will also be large (e.g., forecasts, being affected by revisions, will change and differ from each other) [Croushore (2010)].

A last point worth of consideration before carrying on with the next section relates to the length of the lag in the chosen model: in the immediately preceding discussion an AR(1) process has been used, with a lag length equal to one. It is important that the forecaster adopts a lag length that best suits the data and, in order to do so, usually resorts to information criteria such as AIC (i.e., Akaike's Information Criterion) and SIC (i.e., Schwarz Information Criterion); of course, this arbitrary choice adds another source of potential forecast error. The forecasting analysis of Chapter 3 provides for a thorough outlining and description of the chosen models, together with the rationale behind the choice of the lag lengths.

1.2.3. Noise reduction versus information addition

In the preceding discussion, little has been said about the nature of the revisions (e.g, the error term in the equations); this section lays down assumptions about its structure.

It has formerly been argued that national agencies, when they release macroeconomic estimates, do not have complete samples at hand; when samples are not complete, agencies need to take on assumptions in order to fill such information gaps. Two major paths can be spotted here as to how to proceed: first, the released estimates may just come from the sample at hand, meaning that the statistical agency releases data based solely on the numbers available at the moment, being them incomplete or not; second, the institute may combine the incomplete sample with other useful information available through other channels, with the aim of generating an optimal estimate [Sargent (1989)].

The difference in the two available paths is as follows: measurement errors of data released by a country complying with the former will not be correlated with the variable's true value; consequently, subsequent data revisions will solely reduce estimates' noise (i.e., revisions have a noise reduction purpose). In this case, the equation from the simple model in *section 1.2.1* would be modelled as follows [Croushore (2010)]:

$$y_t^v = y_t^* + \varepsilon_t^v \;\; \rightarrow \;\; y_t^v = y_t^* + u_t^v$$

Where the variable's true value is orthogonal to the error term:

$$y_t^* \perp u_t^v$$

Meaning that revisions are not correlated with the true variable the agency is attempting to measure (i.e., y_t^*); instead, the error will be correlated with earlier data estimates thus making revisions predictable.

If a country were to comply with the latter path (i.e., sample combined with other useful information), on the other hand, national agencies' data releases would be optimal forecasts of subsequent variables' estimates, and revisions would therefore add news. In this case the simple model of *section 1.2.1* would look like this:

$$y_t^v = y_t^* + \varepsilon_t^v \;\; \rightarrow \;\; y_t^* = y_t^v + e_t^v$$

With the error term being correlated to the variable's true value, and uncorrelated with any forecasted estimate:

$$y_t^v \perp e_t^v$$

The main consequence stemming from the forecasted estimate being orthogonal to the error term, is that data revisions will not be predictable.

To sum up, let us argue some of the reasons which might lead a government to take either one of the two paths. First, adding new information to a forecasting model requires exercising judgment which, in turn, might result into potential errors. On the other hand, the other possibility is reasonably cheaper and less subjective, two qualities which may come handy to national agencies: straightforward protocols providing for estimates based on incomplete samples, where gaps are filled with naïve projections, are less expensive and less exposed to political adjustments aiming at making numbers look better [Diebold and Rudebusch (1991)].

1.3. Empirical literature on some macroeconomic variables

First instances in which real-time data relevance has been investigated date back to the sixties, when forecasting models based on Canadian data provided for completely different results when using real-time data compared to latest available one [Denton and Kuiper (1965)]; a few years later, a new study [Cole (1969)] using consumption data proved that other than different results, the use of preliminary data led also to doubled measurement errors and made forecasts biased and inefficient, pressing for accuracy improvement. This section unveils empirical findings on revisions significance on some selected macroeconomic variables.

1.3.1. Stock returns, income, and consumption forecasting

A large bunch of literature argues that the consumption-wealth ratio can reliably predict stock-market returns: this statement does not seem to hold for real-time data [Guo (2003)]. Similarly, many studies argued that the saving rate could explain movements in consumption spending and income: this, just like before, proves true only for latest-available data. Real-time savings rate lead indeed to worse income forecasts and are found to be utterly uncorrelated with consumption spending forecasts [Nakamura and Stark (2007)].

1.3.2. Inflation forecasting

When it comes to inflation, common forecasting practice calls for the mark-up as main prediction ingredient (i.e., difference between costs and prices); this, though, proves to be true only when it comes to revised data [Koenig (2003)]. If one were to forecast inflation with real-time data, on the other hand, she would find herself with completely wrong results as the mark-up is unable to predict inflation when revised data is not used [Orphanides and Van Norden (2005)].

1.3.3. Recession prediction

Countless efforts have been put into the development of models for recession forecasting; all of them, though, provide for the use of revised data. Some researchers tried to test such models' reliability by using real-time data, instead. Results show how such models turn out to be unreliable and useless for forecasting when revised data is not used [Filardo (1999)].

1.3.4. Output growth forecasting

Unusual evidence comes from industrial production and real GDP forecasted through the usage of leading indicators: it has been found that such forecasting turns out to be reliable even when real-time data is used [Robertson and Tallman].

1.3.5. Exchange rates forecasting

When it comes to exchange rates, forecasting cannot be said to be easy. Some researchers did try to model them though and discovered them to be much more sensitive to the chosen data vintage compared to other variables. Contrary to what has been previously shown, studies argue that exchange rates prove to be more predictable when real-time data is used, instead of revised one [Molodtsova and Papell (2008)].

1.3.6. How about unemployment forecasting?

It has been shown that real-time and latest available data impact forecasting results in different forms, depending on the variable that is being modelled. The empirical literature on the subject covers many areas; nevertheless, much is yet to be discovered: existing empirical literature (1) mainly focuses on mainstream macroeconomic variables such as GDP, industrial production and inflation among the others and (2) most databases relate to Anglo-Saxon countries, such as the USA, the United Kingdom and Canada. This thesis work aims at filling such gap by investigating (1) if, and if so, how real-time and latest available data affect the forecasting of an out of the ordinary variable, this being the unemployment rate, (2) in a reasonably large sample of 28 countries.

II. FOUNDATIONS OF THE ANALYSIS

The thesis work deals with a forecasting analysis based on unemployment, applied to a sample of 28 countries. This preparatory chapter provides for all relevant information concerning the analysis' milestones: it attempts at explaining the choice and the features of the chosen variable, together with a description as to where the database stems from and how it has been built.

2.1. Unemployment

As previously pointed out, relevant studies concerning data revisions impact on forecasting and the resulting relationship between real-time and latest available data have prominently been based on a handful of conventional variables, such as GDP, inflation and consumption among the others. When dealing with unemployment, however, even though relevant literature concerning its forecasting can easily be found, none of it addresses the problem raised by data revisions.

2.1.1. Some clarifications

When it comes to unemployment, the first thing worth considering is the true meaning of "unemployed people". One might think that a person would be labelled as unemployed if she did not have a job; this statement, despite being logical, is only partially true. Indeed, a person who wants to call herself "unemployed" can only do so if (1) she is jobless, (2) she is *willing* to work and (3) has already taken all the necessary steps to find an occupation. All the people falling within that category represent the unemployed portion of a country's population that, if combined to the employed one, constitute the labour force. This strict meaning of unemployment may potentially be misleading if not properly accounted for; let us make an example: the Spanish unemployment rate in 2019 is equal to 14.24%, whereas the Italian one in the same year is equal to 10.03% [OECD (2013-2019)]. Such big difference, which mainly stems from the Spanish economy relying too much on tourism while lacking of a structured industrial apparatus, could also be partially explained by the fact that there are more people in Spain who would like to have a job but cannot find any; as a consequence of it, an uninformed reader looking at those numbers would go to the conclusion that Italy's unemployment rate is way lower than Spain's, because the former country has a higher number of available job positions per capita. Before issuing such a statement, a more experienced reader would instead also consider all the people seeking work but not immediately and, more importantly, people available to work but not seeking any: after taking into account such information, the resulting difference between the two countries' unemployment rates may not be that large anymore [Blanchard, Amighini and Giavazzi (2017)]. To sum up, before carrying on with any analysis providing for the usage of unemployment data, the correct definition of the word "unemployment" has to be clear in the reader's mind, in order not to reach misleading conclusions.

Other issues worth of consideration are instead more related to national statistical agencies' particular measurement methodologies, and to incomparability in unemployment results due to countries' heterogenous legislation. The former relates to assumptions made by national agencies which may, for example, count as civilian labour force career members of the armed forces living in private households; similarly, other national agencies may assume that people with no residence may not be considered part of the labour force, thus causing large discrepancies which are nevertheless necessary when dealing with emerging countries, characterised by large numbers of unbanked adults (e.g., Brazil). The latter point worth considering regards heterogeneous legislation negatively impacting on unemployment comparability among countries: one clear example is the working age as set by law, which may cause inconsistencies in labour force count when comparing a country setting a working age starting at 15 years old (e.g., Italy), with another one setting it at 16 (e.g., United Kingdom). Appendix A.1 provides for a thorough outlining of the measurement methodologies and legislation details concerning the 28 countries belonging to the sample.

2.1.2. Insights on unemployment forecasting

Literature concerning unemployment forecasting focuses on two main approaches. The first approach aims at predicting unemployment by exploiting its time series nature: given that national agencies measure macroeconomic variables every month, every year and so on, these can be forecasted by using time series models (e.g., ARMA(1,1)) based on their past data [Montgomery, Zarnowitz, Tsay and Tiao (1998)]; this is the forecasting method employed in the thesis work. The second method is instead based on the macroeconomic nature of the unemployment variable: forecasting analysis can indeed be performed by exploiting the relationship between output growth and unemployment itself, also known as Okun's law [Okun (1962)].

A recent attempt [Barnichon and Nekarda (2012)] at forecasting unemployment has instead been envisaged in 2012, according to which the variable can be forecasted through the input of work force flows; this turns out to be useful as unemployment is characterised by asymmetric movements which can be better captured with the usage of individual work force flows whose contributions, by construction, change over time. Despite the fact that such method outperforms the other two in some instances, it has been marked as unreliable and not ready for final usage, yet.

To sum up, forecasting models and literature concerning unemployment can easily be found; none of it, though, deals with the contrasting results that *may* stem from the usage of real-time compared to latest available data.

2.2. The database

The analysis that will follow in the next chapters is based on a database that has been rearranged with the usage of general-purpose statistical software STATA. Data has been gathered from OECD (i.e., Organisation for Economic Cooperation and Development), an international organisation which focuses on economic studies concerning its member states, these being developed countries featuring market economies. This sub-chapter lays down the features of the chosen sample, together with an outline as to where it stems from: the OECD's Economic Outlook reports.

2.2.1. OECD's Economic Outlook

The Economic Outlook is a twice-yearly analysis issued by the OECD which aims at describing the major economic trends and projections of some chosen countries: economists and analysts behind the Economic Outlook, by reviewing member states' statistical data and national economic policies, lay down projections concerning key macroeconomic variables such as output, government spending, inflation and unemployment among the others for each individual country. More precisely, projections are made for the same year in which the issue is released (i.e., also known as "nowcast") and for the subsequent two years (i.e., "forecasts") [Organisation for Economic Cooperation and Development].

As previously pointed out in the chapter, the analysis is based on one key variable: unemployment; OECD's economists, when gathering historical information about labour data, take into account national and international sources in order to end up with high-quality, recent and relevant data. In the view of getting reliable and consistent results, smoothing techniques such as splicing are used to fill data breaks; because of this, the resulting database may differ compared to other macroeconomic datasets. Going into more detail, unemployment data from all countries is gathered from two main types of sources, which do not always provide for equal results:

- Labour force household surveys; and
- National account databases.

The former are used by OECD's economists to compute unemployment rate, whereas the latter are used for labour productivity and costs' instead [OECD Economic Outlook Statistical Sources].

Concerning each country's unemployment data sources and their particular features, Appendix A.1 provides for a thorough description as to which national agency gathers macroeconomic data, together with assumptions based on each country's legislation (e.g., working age) and related breaks when such legislation has been subject to amendments (e.g., changes in working age).

2.2.2. Database description

Sample unemployment data comes from the twice-yearly OECD's Economic Outlook reports [OECD Statistics], starting from the second issue of 1996 (i.e., Economic Outlook No. 60, December 1996) until the second issue of 2019 (i.e., Economic Outlook No. 106, December 2019), for a total number of 47 issues, more than 20 thousand observations, covering 26 years of data (i.e., let us not forget that the latest 2019 issue provides for 2 years of forecasts, meaning that time-series do not end in year 2019, but in year 2021 instead). Issue 60 represents a break compared to previous ones, as OECD deeply changed its databases organisation (e.g., from Issue 60 on each volume is organised by variable, whereas previous ones were organised by country), for this reason it has been considered a proper sample starting point.

The choice of the 28 countries is straightforward, too: in order to guarantee sample consistency, only countries showing up in all 47 issues have been selected, thereby discarding countries which have started being accounted for by OECD in later years (e.g., South Korea, accounted for by OECD for the first time in Economic Outlook No. 62, December 1997). The following table shows the time-series length for each country belonging to the sample, the time-series starting year can instead be found between brackets.

Table 2 - Years of unemployment data by country (time-series starting year[1])

Australia	61 (1960)	Japan	61 (1960)
Austria	67 (1954)	Luxembourg	47 (1974)
Belgium	65 (1956)	Mexico	41 (1980)
Canada	65 (1956)	Netherlands	52 (1969)
Czech Republic	28 (1993)	New Zealand	61 (1960)
Denmark	71 (1950)	Norway	65 (1956)
Finland	62 (1959)	Poland	31 (1990)
France	56 (1965)	Portugal	65 (1956)
Germany[2]	28 (1993)	Spain	61 (1960)
Greece	65 (1956)	Sweden	65 (1956)
Hungary	29 (1992)	Switzerland	65 (1956)
Iceland	61 (1960)	Turkey	51 (1970)
Ireland	65 (1956)	United Kingdom	61 (1960)
Italy	62 (1959)	United States	65 (1956)

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

[1] The ending year, as previously stated, is 2021

[2] Given the fall of the Berlin wall and the disruption in German macroeconomic data that followed, it has been deemed appropriate to cut the time-series in 1993. Other countries' time-series presenting the same issue (e.g., Czech Republic) have been subject to no adjustments as such time-series already have the nineties as starting years.

To make things clear, each Economic Outlook report provides for a nowcast and two years of forecasts meaning that, for example, the Australian time-series in the issue of December 1996 starts in 1960 (as it can be seen from the table above), has the 1996 unemployment value as nowcast and 1997 and 1998 values as forecasts; on the other hand, the same time-series in the issue of December 2000 still starts in 1960, has the 2000 unemployment value as nowcast and 2001 and 2002 values as forecasts. Older unemployment data (i.e., values that are not nowcast nor forecasts) represent the portion of the sample which has been subject to data revisions in that particular issue (e.g. in the case of Issue December 2000, all Australian data from 1960 to 1999 have been subject to revisions).

Now that the sample has been presented and thoroughly described in its structure, the analysis may continue with some preliminary findings shedding light on if, and if so, how revisions impact unemployment data in the chosen countries.

III. PRELIMINARY FINDINGS

This chapter provides for a description of some preliminary results of the analysis, which are deemed to be propaedeutic and helpful for a thorough understanding of what will come later. First, real-time and latest available time series gathered from the previously mentioned database are here described with some basic statistics, useful to grasp their structure; second, national agencies' data revisions will be investigated as well, in order to let the reader get an idea of the relevance of data revisions when it comes to unemployment.

3.1. Real-time and latest available series: some basic statistics

Through the usage of general-purpose statistical software STATA, the database has been rearranged in two ways: first, real-time data series have been gathered by taking only values' first releases, in order to get time-series for each country characterised by only unrevised data (e.g., values on the main diagonal of Table 1, Chapter 1); second, latest available data has been gathered by picking all the values present in the latest issue (i.e., Issue no. 106, December 2019), with the aim of getting only the newest, revised unemployment numbers (e.g., values on the far-right column of Table 1, Chapter 1); let us remind that the 2019 latest available value, being a first-release, is a real-time one, too. As a result, there will be two time-series per country (i.e., one real time and one latest available), for a total number of 56 series.

3.1.1. Volatility analysis with full sample

As a first step of the analysis, let us question how revisions impact time-series volatility. The following table provides for some basic statistics concerning the 56 series gathered from the sample. For a more complete outlining of such time-series' statistics, refer to Appendix A.2 and A.3.

Table 3 - Real-time and Latest available descriptive statistics (full)

	Real-time		Latest available	
	Mean (%)	SD (%)	Mean (%)	SD (%)
Australia	5.982	1.186	5.731	2.435
Austria	5.348	0.668	3.697	1.374
Belgium	8.393	1.889	6.277	3.074
Canada	7.247	1.002	7.367	1.972
Czech Republic	6.104	2.087	5.615	2.147
Denmark	5.814	1.415	5.755	2.159
Finland	9.073	2.289	6.562	4.083
France	9.751	1.238	6.728	3.283
Germany	7.055	2.394	7.037	2.425
Greece	14.680	6.546	14.752	6.221
Hungary	7.465	2.359	7.565	2.652
Iceland	3.751	1.742	2.788	1.697
Ireland	8.033	3.755	9.261	4.375
Italy	9.898	2.028	7.566	2.910
Japan	4.094	0.893	2.770	1.264
Luxembourg	4.794	1.476	3.699	1.905
Mexico	3.929	1.087	4.326	0.989
Netherlands	4.563	1.424	5.007	2.938
New Zealand	5.447	1.273	3.928	2.913
Norway	3.666	0.599	3.303	1.185
Poland	11.184	4.885	10.979	5.088
Portugal	8.519	3.532	6.668	3.180
Spain	16.645	5.560	14.843	4.968
Sweden	6.453	1.400	5.215	2.896
Switzerland	3.936	0.867	3.003	1.875
Turkey	9.624	2.073	8.634	1.909
United Kingdom	5.946	1.345	6.133	2.589
United States	5.710	1.732	5.890	1.619

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

By first looking at the table, one can argue that the real-time and latest available series are not that different compared to each other. Still, differences in mean and standard deviation can be found, suggesting that data revisions do play a role when dealing with unemployment. In order to get a better view of how significant such revisions are, let us consider the following table, showing the mean of the means and of the standard deviations:

Table 4 - Mean of means and standard deviations (full)

Mean (SD)				Mean (Mean)			
Real-time	2.098	Latest available	2.719	Real-time	7.254	Latest Available	6.468

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

Considering all real time and latest available series belonging to the sample, latest available data's standard deviations are, on average, roughly 30% higher than real-time ones. Real-time series' means, on the other hand, are about 12% higher than latest available's. Now, in order to get rid of potential outliers, the medians are considered instead:

Table 5 - Median of means and standard deviations (full)

Median (SD)				Median (Mean)			
Real-time	1.604	Latest available	2.512	Real-time	6.279	Latest Avaiable	6.012

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

Despite the usage of the median, standard deviations in latest available time series are sensibly higher (56%) compared to real-time's; means, on the other hand, look more alike compared to before, with real time data being on average 4,5% higher than latest available's.

Such large differences, especially concerning volatility, may be due to a combination of the following two reasons:

- Data revisions impact heavily on time-series, thus making real-time and latest available data largely differ from each other; and
- Real-time time series are significantly shorter than latest available's.

Special attention must be paid to the last point, which is directly linked to the real-time databases' availability issues raised in Chapter 1: despite the database at hand being quite large (such large databases allowing for building both real-time and latest available time series are quite uncommon, too), real-time data series are quite short compared to the latest available ones, as they all go back to 1996 which is the year of the oldest issue at hand. Latest available series are instead, on average, way longer as they start depending on the oldest revised values released by national agencies in the latest year of issue which, in this case, is 2019. Appendix A.2 and A.3 outline real-time and latest available time series lengths, respectively.

3.1.2. *Volatility analysis with truncated latest-available series*

As an attempt to improve real-time and latest available time series comparability, while at the same time making volatility comparisons entirely based on revisions' impact, let us replicate the analysis carried out in sub-chapter 3.1.1 by using, this time, truncated latest available time series which will now start off in 1996 and end in 2019; in this way, real-time and latest available series have equal length and the latter is not altered by (1) the presence of data older than 1996 and (2) the 2020-2021 OECD projections which, by construction, are not present in the real-time series.

A summary of the basic statistics of the truncated latest available time series, together with the previously seen real-time ones, can be found in the table below. For more detailed statistics about the truncated latest available time-series, refer to Appendix A.4.

Table 6 - Real-time and Latest available descriptive statistics (truncated)

	Real-time		Latest available	
	Mean (%)	SD (%)	Mean (%)	SD (%)
Australia	5.982	1.186	5.884	1.107
Austria	5.348	0.668	4.767	0.722
Belgium	8.393	1.889	7.836	1.010
Canada	7.247	1.002	7.216	0.980
Czech Republic	6.104	2.087	6.078	2.021
Denmark	5.814	1.415	5.790	1.169
Finland	9.073	2.289	8.957	2.019
France	9.751	1.238	9.197	0.969
Germany	7.055	2.394	7.211	2.394

Greece	14.680	6.546	14.896	6.522
Hungary	7.465	2.359	7.405	2.295
Iceland	3.751	1.742	7.405	2.295
Ireland	8.033	3.755	8.330	3.937
Italy	9.898	2.028	9.637	1.912
Japan	4.094	0.893	4.072	0.868
Luxembourg	4.794	1.476	4.535	1.589
Mexico	3.929	1.087	4.277	0.805
Netherlands	4.563	1.424	5.096	1.334
New Zealand	5.447	1.273	5.327	1.133
Norway	3.666	0.599	3.607	0.594
Poland	11.184	4.885	11.288	4.967
Portugal	8.519	3.532	8.532	3.520
Spain	16.645	5.560	15.768	5.417
Sweden	6.453	1.400	7.628	1.527
Switzerland	3.936	0.867	4.382	0.622
Turkey	9.624	2.073	9.339	2.000
United Kingdom	5.946	1.345	5.919	1.340
United States	5.710	1.732	5.696	1.723

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

By looking at the revised table, it is now apparent that the two series look much more alike and that truncated latest available series' standard deviations are way lower than before. Let us now have a look at the means:

Table 7 - Mean of means and standard deviations (truncated)

Mean (SD)				Mean (Mean)			
Real-time	2.098	Latest available	2.028	Real-time	7.254	Latest Available	7.360

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

And medians:

Table 8 - Median of means and standard deviations (truncated)

Median (SD)				Median (Mean)			
Real-time	1.604	Latest available	1.558	Real-time	6.279	Latest Avaiable	7.213

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

Given that all time-series have now equal length, any difference in standard deviations will be entirely due to revisions impact. By looking at means and medians, one can argue that standard deviations look a lot more similar compared to before, signalling that the widely different time-series lengths were the main reason behind volatility differences in the previous section. Still, latest available's standard deviations are not equal to real-time ones and such differences can give interesting insights on revisions impact; indeed, the latter's time series are, on average, 3.5% more volatile than the former. Medians behave similarly, with real-time series' standard deviations being 3% higher than latest available's. This is the result one would expect when computing volatility on revised and unrevised data: in Chapter 1 it has been argued that first-release data suffer from many problems (e.g., incomplete samples and incomplete information among the others); when such provisional data are revised, such problems even out and estimates become increasingly precise and reliable. For this reason, it can be reasonable to expect latest available time series to be more precise (i.e., less volatile) than real-time ones, and now it can be argued that this is exactly what happens, on average, also when dealing with unemployment data samples.

3.2. Investigating revisions' impact

By taking 1996 and 2006 as benchmarks, this sub-chapter aims at shedding light on how often data is revised and on how much revisions impact numbers through all data vintages at hand. Given the issues national agencies face when gathering sample data, which have been outlined in Chapter 1, an economist would expect revisions to take place in the early years following the first release, after which samples should be complete and data should not need to be revised anymore. In this sense, 1996 represents a proper benchmark as it is the first year at hand providing for first releases. Another reason leading national agencies to revise their data, as previously pointed out, may be linked to changes in measurement assumptions. 2006, in this sense, is a reasonable half-way through the sample of data vintages available which, if compared to 1996 data, can shed light on governments' measurement changes.

3.2.1. 1996 Italian unemployment data

Let us begin by considering only one country: Italy. This will be useful in order to get some basic insights concerning how revisions work; such primary results will come handy for the next section's aggregate analysis, which will exploit the whole sample of countries.

In Table 9 the Italian unemployment rate in 1996 can be found, with the 1996 value being the first release, and the following ones being the yearly revisions released by the Italian statistical agency, ISTAT. The first release of 1996 marks an unemployment rate equal to 12.16%, this value has been revised downward in 1997, in 1999 and so on until 2019, with the most up-to-date value at hand being equal to 11.18%.

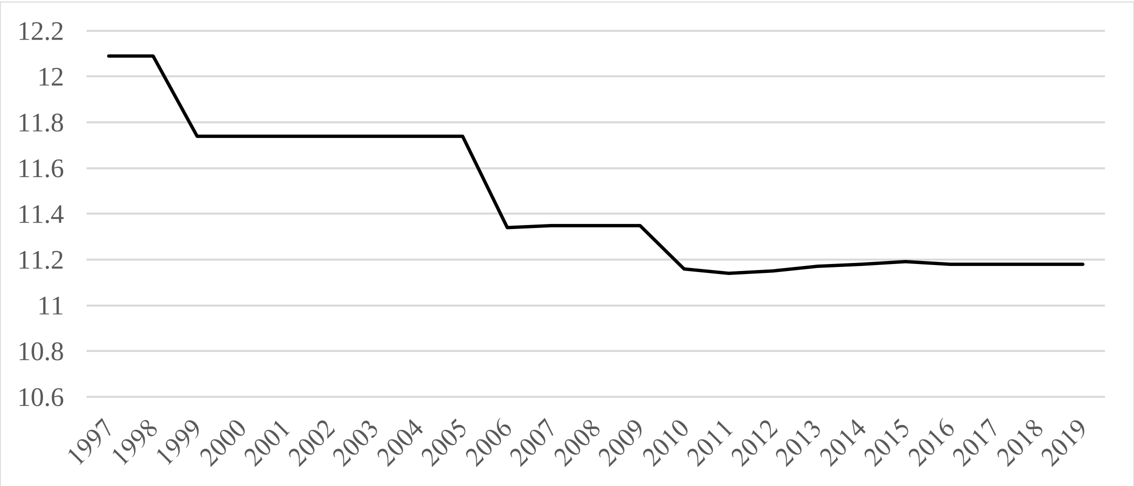
Table 9 - Italian unemployment rate in 1996 (data vintages 1996 - 2019)

1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
12.16	12.09	12.09	11.74	11.74	11.74	11.74	11.74	11.74	11.74	11.34	11.35
2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
11.35	11.35	11.16	11.14	11.15	11.17	11.18	11.19	11.18	11.18	11.18	11.18

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

The following figure, on the other hand, provides for a graphical representation of how revisions impact the 1996 Italian unemployment rate:

Figure 1 – Italian Unemployment rate in 1996 (data vintages 1996 – 2019)



Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

With the aim of enhancing comprehension, the following table shows instead only the net effects of the revisions:

Table 10 - Net effect of revisions, Italian UNR in 1996 (data vintages 1996 - 2019)

1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
-	-0.07	0.00	-0.35	0.00	0.00	0.00	0.00	0.00	0.00	-0.40	0.01
2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
0.00	0.00	-0.19	-0.02	0.01	0.02	0.01	0.01	-0.01	0.00	0.00	0.00

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

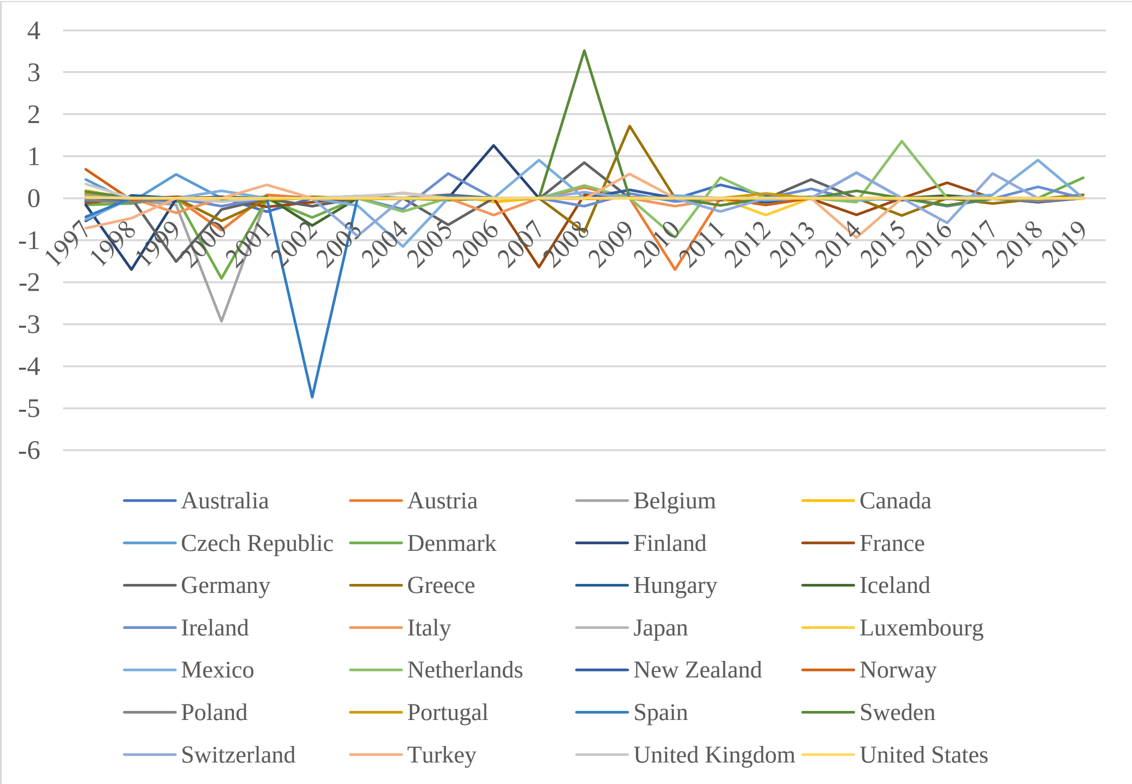
It is now apparent that Italy's 1996 unemployment data, contrarily to what was originally thought, have not been revised only in the early subsequent years that followed the first release; indeed, numbers have been revised by the ISTAT even 20 years after the first release. If the 2016's revision effect can be argued to be negligible, the same cannot be said for the 2006 and 2010 one, when the unemployment rate has been revised downwards by 0.4% and almost 0.2%, respectively. Before getting to any conclusion, let us now consider the aggregate sample of 28 countries, in order to see if revisions do follow the same Italian pattern, or if they instead support the theory that data is revised only in the early years following the first release.

3.2.2. 1996 unemployment data: whole sample

Given the huge quantity of data to manage when dealing with the whole sample of 28 countries, in what is going to follow graphs will be used to describe data. Appendix A.5 and A.6, though, provide for tables showing 1996 unemployment data for all sample countries in all data vintages, and the net effects of data revisions on such unemployment data, respectively.

The following figure plots net revision effects of the sample countries on 1996 unemployment rates. A very interesting insight can be gained here, as one can argue that Italy is not an isolated case and that data in all countries are *continuously* revised even 20 years after the first release, signalling that (1) data is revised long after complete samples are gathered by national agencies and (2) measurement changes (considering also very small ones) seem to take place quite often, for all countries.

Figure 2 - Sample countries' net revision effects on UNR in 1996 (data vintages 1996 – 2019)



Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

Some outliers can be found here, such as Spain's 4.7% downward revision in 2002, Sweden's 3.5% upward revision in 2008 and Belgium's 2.9% downward revision in 2000. Despite such outliers, though, the yearly means are all quite balanced with an overall mean equal to -0.020%:

Table 11 - Net revision effects on UNR in 1996, yearly means (data vintages 1996 - 2019)

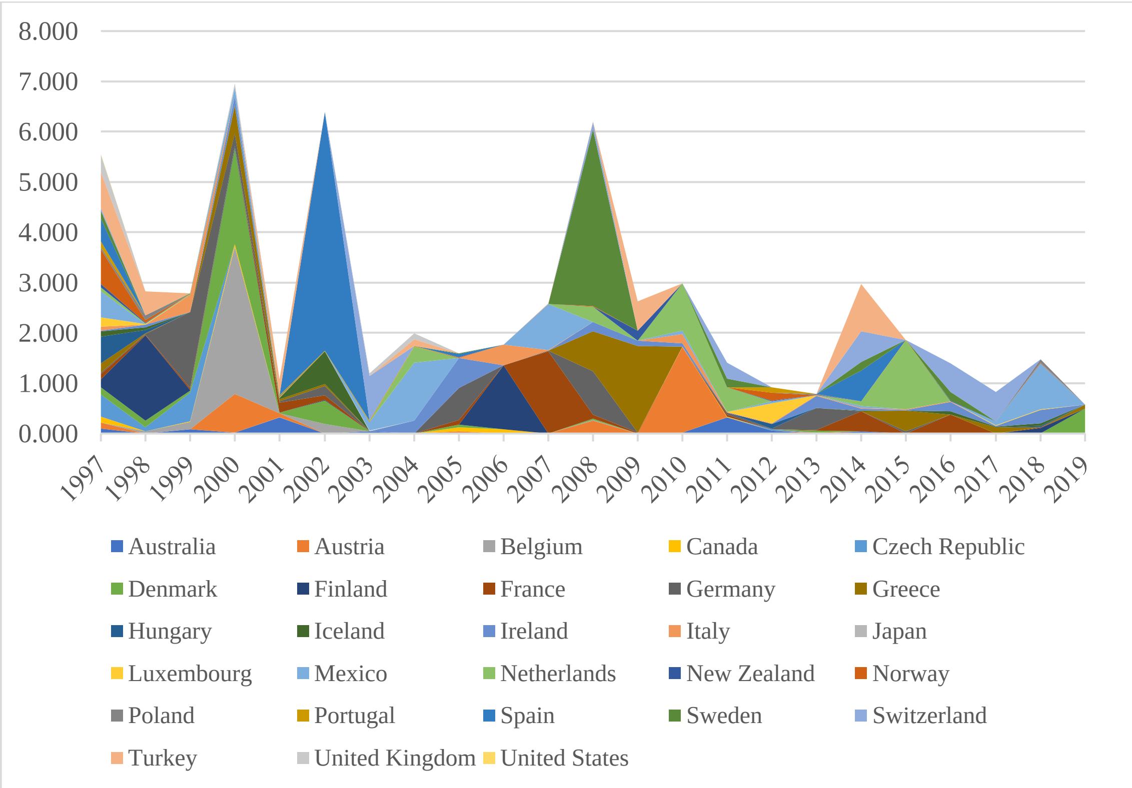
1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
-	-0.006	-0.094	-0.053	-0.233	-0.008	-0.223	-0.035	-0.053	0.004	0.028	-0.025
2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
0.150	0.094	-0.101	0.008	-0.019	0.027	-0.006	0.032	-0.019	0.018	0.034	0.020

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

Yearly medians are balanced too, as they all equal 0.000%.

In order to see whether there are some years in which revisions take place more often than others, the absolute values of the revisions' net effects have been computed and plotted in a cumulative graph, which can be found below:

Figure 3 - Cumulative net revision effects on UNR in 1996 (data vintages 1996 – 2019)



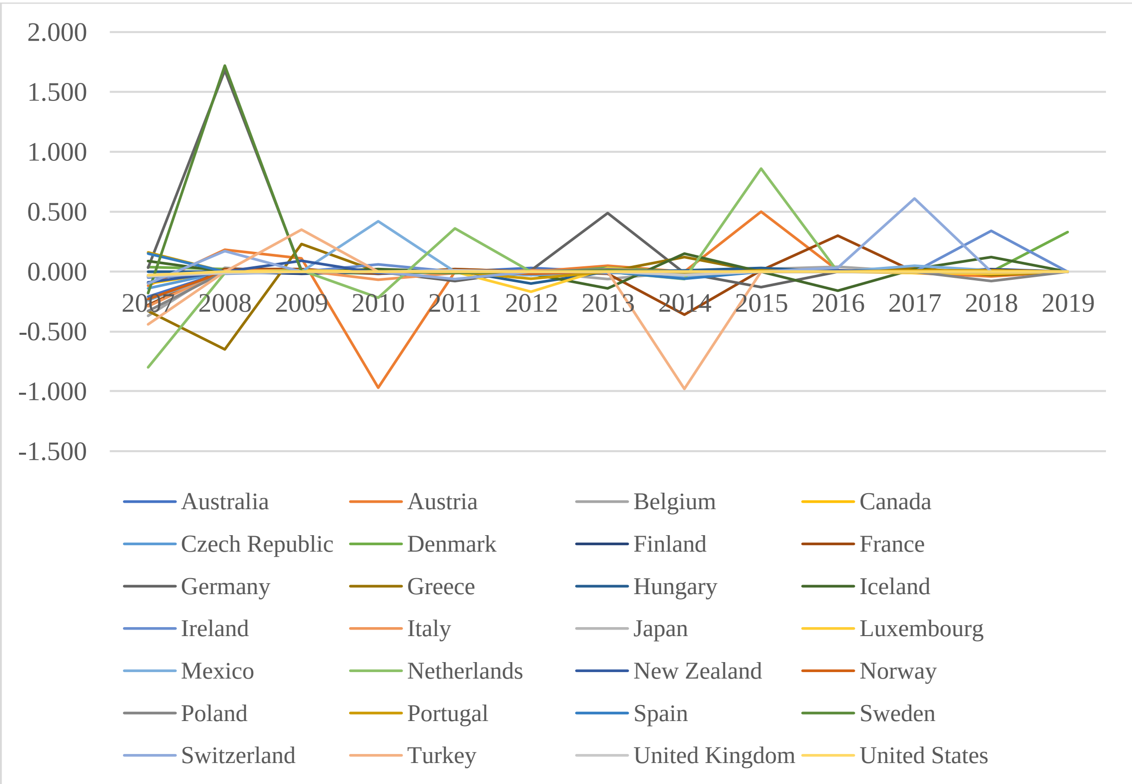
Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

As expected, almost all countries revise their data in the early years following the first release (e.g., the total amount of cumulative revisions one year after the first release, 1997, equals roughly 5%) after which, though, revisions do not stabilise but instead keep on taking place: even without considering the previously mentioned outliers, data revisions seem to reach a balance in the early 2000s, after which almost all national agencies resume revision practices reaching a peak between 2008 and 2009. Years from 2011 to 2013 are characterised by rare data revisions which, nevertheless, seem to partially resume in the following years (e.g., total cumulative revisions from 2013 to 2015 roughly equal to 2%). To sum up, the expected pattern featuring revisions only in the early years following the first release seems to be replaced by a cyclical pattern, where periods characterised by rare data revisions are followed by others in which revision practices are often resumed.

3.2.3. 2006 unemployment data: whole sample

As a double check, let us see whether revisions in 2006 unemployment data behave similarly to the 1996's ones. The following figures plot net revision impacts of the sample countries on 2006 unemployment rates, for data vintages spanning from 2006 to 2019:

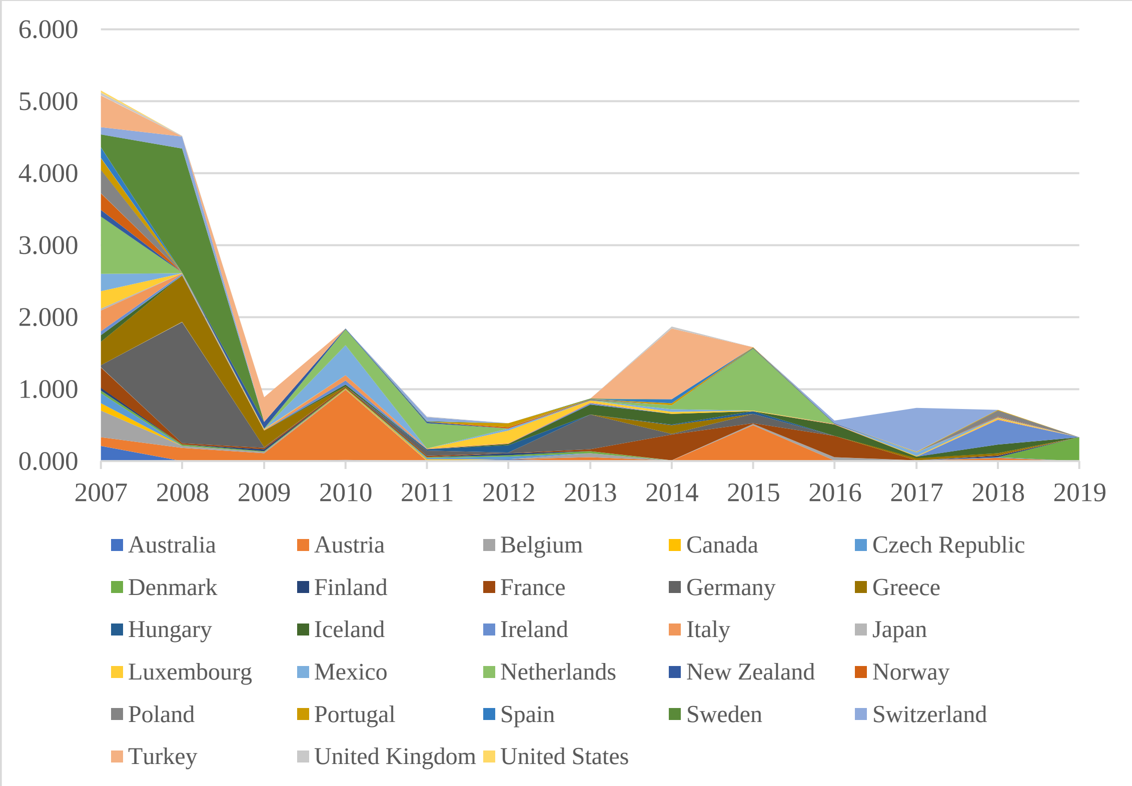
Figure 4 - Sample countries' net revision effects on UNR in 2006 (data vintages 2006 – 2019)



Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

And the cumulative graph featuring absolute values of revisions' net effects:

Figure 5 - Cumulative net revision effects on UNR in 2006 (data vintages 2006 – 2019)



Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

Just like before, early years after the 2006's initial data release are characterised by many revisions, amounting to a total cumulative value in 2007 roughly equal to 5%, just like 1997's data revisions following 1996's first data release. Revisions seem to stabilise in the following years until 2013, when national agencies seem to sporadically resume their revision practices.

To conclude, national agencies do revise their data in the early years following the first release and, contrarily to what one would expect, they keep on doing it also 10, even 20 years after real time data releases, with a seemingly decreasing cyclical pattern which seems to be more pronounced during periods of economic crisis. Given this thesis' work purposes and this chapter's preparatory nature for what comes next, at this point the reader has to bear in mind that revisions do affect unemployment data, and they continuously do so throughout the life of any particular data release.

IV. UNEMPLOYMENT FORECASTING

Given what has been shown about revisions when it comes to unemployment, the analysis will now carry on with an empirical exercise involving unemployment forecasting with real-time and latest available time series. OECD's annual reports, by construction, report estimates computed (1) for the same year in which the issue is released (i.e., nowcast) and (2) for the subsequent two years (i.e., forecasts). Given the latest available issue at hand, then, the thesis work's forecasting results will be provided for years 2019, 2020 and 2021 and will here be compared to OECD's estimates which will work as benchmarks, in order to see whether forecasts based on latest available time series are more precise than real-time ones.

4.1. Assumptions

Unemployment forecasting attempts have been discussed in section 2.1.2., in which it has been argued that, among the available prediction models, there exists an approach that aims at predicting unemployment by exploiting its time series nature: given that national agencies measure macroeconomic variables every month, every year and so on, these can be forecasted by using time series models (e.g., ARMA(1,1)) based on their past data [Montgomery, A. L., Zarnowitz, Tsay and Tiao (1998)]; this is the forecasting method employed in the thesis work. This section describes and argues all assumptions underlying the forecasting exercise which is going to follow.

4.1.1. Stationarity checking

Stationarity has been checked for each time series through the Augmented Dickey Fueller Test. As a first step, the 56-time series at hand (i.e., 28 real time and 28 latest available time series) have been graphically analysed, in order to understand whether each of them presents (1) a plain random walk with no drift nor trend, (2) a random walk with a drift, (3) a random walk with both drift and trend. Given that the unemployment rate cannot, by construction, be negative, drifts will be expected in all time series; trends, on the other hand, will depend only on the presence of linear trends as non-linear ones are not accounted for in STATA's forecasting tool and will therefore be regarded as drift.

The following table shows graphical analysis' results for all time series:

Table 12 - Dickey Fueller Test's options, results

Latest available series				Real-time series			
AU	Drift	*JP*	Drift	*AU*	Drift	*JP*	Drift
AT	Drift	*LU*	Drift + Trend	*AT*	Drift	*LU*	Drift + Trend
BE	Drift	*MX*	Drift	*BE*	Drift + Trend	*MX*	Drift
CA	Drift	*NL*	Drift	*CA*	Drift	*NL*	Drift
CZ	Drift	*NZ*	Drift	*CZ*	Drift	*NZ*	Drift
DK	Drift	*NO*	Drift	*DK*	Drift	*NO*	Drift
FI	Drift	*PL*	Drift + Trend	*FI*	Drift	*PL*	Drift + Trend
FR	Drift	*PT*	Drift	*FR*	Drift	*PT*	Drift
DE	Drift	*ES*	Drift	*DE*	Drift	*ES*	Drift
GR	Drift	*SE*	Drift	*GR*	Drift	*SE*	Drift
HU	Drift	*CH*	Drift	*HU*	Drift	*CH*	Drift
IS	Drift	*TR*	Drift	*IS*	Drift	*TR*	Drift
IE	Drift	*GB*	Drift	*IE*	Drift	*GB*	Drift
IT	Drift	*US*	Drift	*IT*	Drift	*US*	Drift

As expected, all time-series are characterised by drifts and few of them also present trends; interestingly enough, almost all countries presenting latest available time series with trend have their real-time counterparts with the same feature.

Once Dickey Fueller's options have been chosen for each time-series, the tests have been carried out twice: first with two, then with three lags (i.e., Augmented Dickey Fueller), which, with the aim of enrichening the test's dynamics, are deemed to be adequate for annual data. A p-value's cut-off of 5% has been chosen to decide for stationarity. The following table provides for latest available time series' ADF Test results:

Table 13 - Latest Available series, ADF Test results

Country	P-value 1 Lag	P-value 2 Lags	Country	P-value 1 Lag	P-value 2 Lags
AU	0.0272	0.0267	*JP*	0.0701	0.0354
AT	0.0357	0.0331	*LU*	*0.1548*	*0.0897*
BE	0.0383	0.0157	*MX*	0.0106	0.0092
CA	0.0165	0.0438	*NL*	0.0267	0.0194
CZ	*0.1697*	*0.1046*	*NZ*	0.0377	0.0369
DK	0.0024	0.0009	*NO*	0.0124	0.0238
FI	0.0211	0.0248	*PL*	*0.2267*	*0.2018*
FR	0.0414	0.0367	*PT*	0.0053	0.0044
DE	*0.3817*	*0.3776*	*ES*	0.0015	0.0008
GR	0.0593	0.0078	*SE*	0.0386	0.0198
HU	0.0041	0.0012	*CH*	*0.0898*	*0.0745*
IS	0.0134	0,0001	*TR*	0.0478	0.1265
IE	0.0013	0.0112	*GB*	0.0346	0.0253
IT	0.0473	0.0323	*US*	0.0039	0.0067

Whereas real-time time series' ADF Test results can be found in the table below:

Table 14 - Real Time series, ADF Test results

Country	P-value 1 Lag	P-value 2 Lags	Country	P-value 1 Lag	P-value 2 Lags
AU	0.0047	0.0177	*JP*	*0.2992*	*0.1235*
AT	0.0064	0.0019	*LU*	*0.0813*	*0.0965*
BE	0.0052	0.0008	*MX*	0.0578	0.0427
CA	0.0179	0.0266	*NL*	0.0121	0.0147
CZ	*0.2599*	*0.5345*	*NZ*	0.0033	0.0034
DK	0.0144	0.0214	*NO*	0.0086	0.0146
FI	0.0382	0.0144	*PL*	0.0035	0.0006
FR	0.0049	0.0033	*PT*	0.0288	0.0320
DE	*0.2158*	*0.5876*	*ES*	0.0227	0.0061
GR	0.0502	0.1134	*SE*	*0.0992*	*0.0975*
HU	0.0161	0.0084	*CH*	0.0479	0.0627
IS	0.0312	0.0323	*TR*	*0.0903*	*0.0874*
IE	0.0138	0.0128	*GB*	0.0407	0.0052
IT	0.0147	0.0169	*US*	0.0395	0.0439

By looking at the two tables, it can be argued that almost all time-series are stationary (i.e., with a 5% p-value cut-off), meaning that they are ready for model fitting. Series which, on the other hand, are not stationary (i.e., series whose p-values are higher than 5% are shown between asterixis) need first to be first-differenced, then stationarity can be checked again. By construction, drifts-featuring series will see their drifts disappear and trends-featuring ones will turn into drifts after first-differencing; further, differenced ADF Test will be characterised by a unit increase in lags (e.g., from two to three lags).

As it can be seen from the table below, all 11 non-stationary time-series (i.e., 5 for latest available and 6 for real-time series) become stationary after first-differencing, meaning that they are all integrated of order one (i.e., I(1)):

Table 15 - ADF Test results after first-differencing

Latest available series		Real-time series	
Country	P-value 1 Lag	Country	P-value 1 Lag
CZ	0.0004	CZ	0.0032
DE	0.0017	DE	0.0002
LU	0.0007	JP	0.0078
PL	0.0036	LU	0.0423
CH	0.0002	SE	0.0035
		TR	0.0032

With all time series at hand being stationary, being them either integrated of order zero (i.e., I(0)) or integrated of order one (i.e., I(1)), model fitting can now be performed.

4.1.2. Model fitting

Model fitting is a crucial step preceding any time series forecasting practice. Given the challenges an economist faces when attempting at fitting a model to a certain time series, things have been kept simple and Auto Regressive Integrated Moving Average (i.e., ARIMA) models have been chosen, as deemed appropriate for tackling such fitting issues. Given time series lengths (especially real-time ones which, as already pointed out, are on average way shorter than latest available's), the following set of possible ARIMA models has been chosen for the time-series to fit as:

Table 16 - Available models for ARIMA fitting

AR	1	1	2	2	0	0	1	2
MA	1	2	1	2	1	2	0	0

Such set of 8 possible models guarantees a reasonable amount of fitting flexibility, and at same time respects the nature of real-time time series which, by construction, are quite short and would not allow for any fitting with higher AR and/or MA grades. Schwarz Information and Akaike Information Criteria (i.e., AIC and SIC, respectively) have been chosen in order to pick best fitters. Fitting results are reported in the table below for both latest available and real-time time series:

Table 17 - Fitting results

Latest available series				Real-time series			
AU	ARMA(1,1)	*JP*	ARMA(1,1)	*AU*	AR(1)	*JP*	MA(1)
AT	AR(1)	*LU*	MA(1)	*AT*	MA(1)	*LU*	ARMA(1,1)
BE	ARMA(1,1)	*MX*	ARMA(2,2)	*BE*	ARMA(1,1)	*MX*	ARMA(2,2)
CA	AR(2)	*NL*	ARMA(2,1)	*CA*	ARMA(1,1)	*NL*	AR(2)
CZ	MA(1)	*NZ*	ARMA(1,1)	*CZ*	MA(1)	*NZ*	n.a.
DK	ARMA(2,2)	*NO*	ARMA(1,1)	*DK*	ARMA(2,2)	*NO*	ARMA(2,2)
FI	ARMA(2,1)	*PL*	AR(2)	*FI*	AR(2)	*PL*	ARMA(2,1)
FR	ARMA(1,1)	*PT*	AR(2)	*FR*	AR(2)	*PT*	ARMA(2,2)
DE	MA(1)	*ES*	AR(2)	*DE*	ARMA(1,2)	*ES*	ARMA(2,2)
GR	ARMA(2,1)	*SE*	ARMA(1,1)	*GR*	AR(2)	*SE*	MA(1)
HU	ARMA(2,1)	*CH*	MA(1)	*HU*	ARMA(2,1)	*CH*	MA(2)
IS	ARMA(1,1)	*TR*	AR(1)	*IS*	AR(2)	*TR*	MA(2)
IE	ARMA(2,2)	*GB*	ARMA(2,1)	*IE*	ARMA(2,2)	*GB*	ARMA(2,2)
IT	AR(2)	*US*	ARMA(2,2)	*IT*	ARMA(2,2)	*US*	ARMA(2,2)

All time series' fittings converged, apart from the one of New Zealand where a flat log pseudolikelihood has been encountered.

4.2. Results

As previously pointed out, the envisaged forecasting model generates estimates for the same year in which the latest available issue at hand has been released (i.e., 2019), and for the two subsequent years (i.e., 2020 and 2021). Before presenting any result, though, it is deemed appropriate to first deal with how OECD annual reports' estimates are computed and what assumptions are used; once the reader has clear in mind how both forecasting practices are constructed (i.e., the OECD's forecasting model and the thesis work's one), OECD's estimates will be compared to the ones generated with real-time and latest available time series.

4.2.1. OECD's forecasting process

When OECD's economists carry out macroeconomic assessments which are later going to be published (e.g., the Economic Outlook), the National Institute Global Econometric Model (i.e., NIGEM) is used. NIGEM has been developed by the British National Institute of Economic and Social Research and is currently adopted by many central banks and policymakers around the world for carrying out economic forecasting, scenario building and stress-testing practices [National Institute of Economic and Social Research]. The model falls within a so called "New-Keynesian" framework by assuming that agents are forward looking and adjustment processes are slow: nominal rigidities impact model balancing after external events; flexibility in the model is ensured as behaviour and policy assumptions can be changed. Structure-wise, on the other hand, the model is based on the national income identity and, despite presenting some resemblances with the Dynamic Stochastic General Equilibrium (i.e., DSGE) model, estimation is entirely based on historical data; this latter fact is particularly important as it aligns NIGEM to the model envisaged in the thesis work, with the only difference being that the former is able to balance both theory and actual data, thus providing the reader with state-of-the-art forecasting estimates. Given NIGEM's flexibility, the OECD uses it both for policy analysis and macroeconomic forecasting [Organisation for Economic Cooperation and Development].

Going into more detail, countries belonging to the OECD are modelled separately by Economic Outlook's economists; countries which, instead, do not belong to such group are modelled through regional blocks, such as Latin America, Developing Europe, East Asia, Africa and OPEC. All countries not falling within the previously listed blocks, to conclude, are modelled as a miscellaneous group (e.g., this mainly addresses some countries in West Asia). Predictably, modelling practices carried out for OECD countries are way more complex than non-OECD ones; all of them, though, are based on production functions supported by dynamic error-correction structures on the estimated equations, allowing models to gradually adjust towards equilibrium following economic shocks [Organisation for Economic Cooperation and Development].

4.2.2. Unemployment forecasting: results

Once stationarity has been checked and ARIMA models have been fitted to the sample's time-series, unemployment forecasting practices can finally begin: through the usage of general-purpose statistical software STATA's forecasting tool, time-series have been refitted, the forecasting models (i.e., one for each time-series, 28 for latest available and 28 for real-time data) have been estimated and unemployment has been forecasted for the years 2019 (i.e., nowcast), 2020 and 2021 (i.e., forecasts). When carrying out such prediction exercise, dynamic forecasting has been chosen according to which forecasted values are themselves used in order to compute subsequent periods' forecasts [StataCorp LLC].

Forecasting results are going to follow in groups of four countries per table (e.g., in alphabetical order), together with benchmark (i.e. OECD) data. In order to enhance data comparison, the analysis will be supplemented by some graphical details.

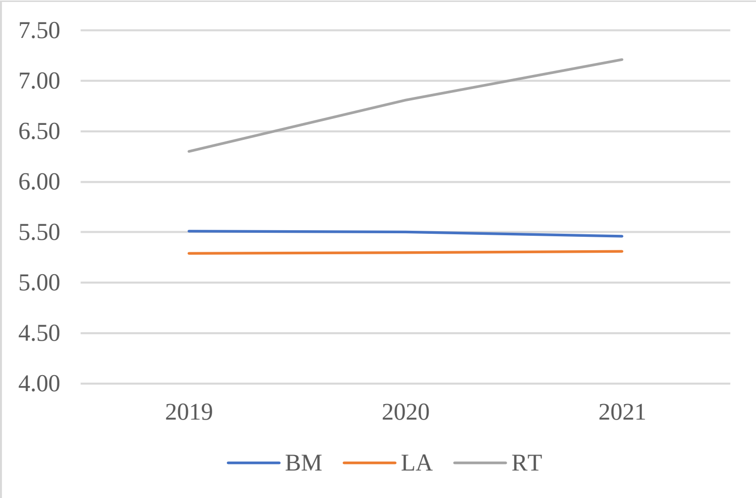
Table 18 shows forecasting results for the first four countries analysed. Here, the BM column shows benchmark data (i.e., nowcast and forecasts produced by the OECD), the LA column represents instead forecasts based on latest available time series; the RT column, finally, shows forecasted data using real-time time series.

Table 18 - Unemployment Forecasts, countries 1 - 4

	Australia			Austria			Belgium			Canada		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2019	5,20	5,16	5,44	4,60	4,78	5,13	5,51	5,29	6,30	5,65	6,60	6,15
2020	5,28	5,16	5,51	4,55	4,71	5,38	5,50	5,30	6,81	5,77	6,24	6,37
2021	5,21	5,15	5,58	4,59	4,65	5,38	5,46	5,31	7,21	5,76	6,79	6,56

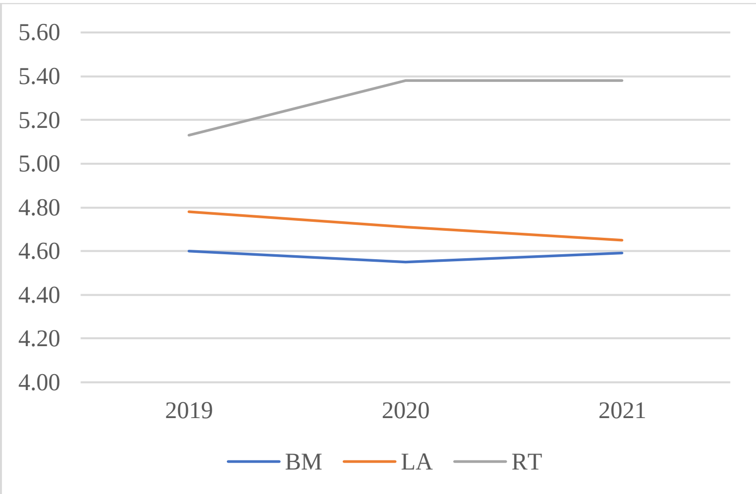
In this sub-sample of four countries, forecasts based on latest available time-series for Australia, Austria and Belgium are all much closer to benchmark than real-time ones; further, LA forecasts do seem to follow the OECD estimates, whereas RT ones seem to have their own pattern. These two facts can be appreciated in the following two figures, featuring Belgium's forecasts:

Figure 6 - Unemployment Forecasts, Belgium



And Austria's:

Figure 7 - Unemployment Forecasts, Austria



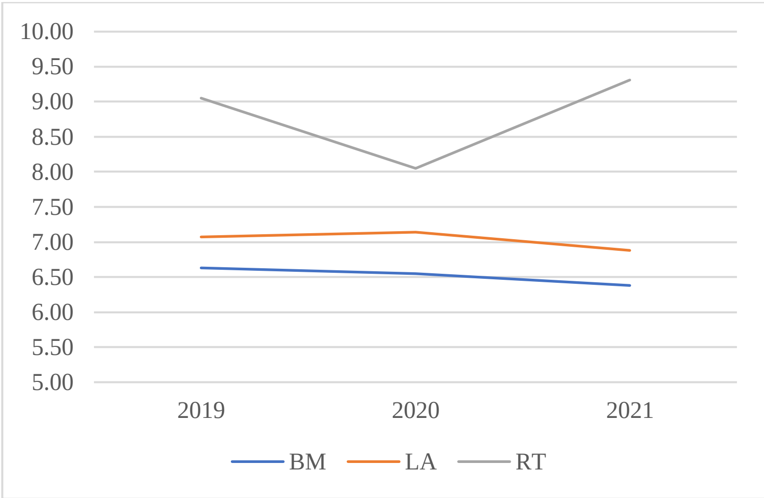
Canada represents the only instance in which real-time forecasts are closer to benchmark for years 2019 and 2021. Let us now see the subsequent group of four countries:

Table 19 - Unemployment Forecasts, countries 5 - 8

	Czech Republic			Denmark			Finland			France		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2019	1,99	1,78	1,88	4,99	5,75	5,88	6,63	7,07	9,05	8,52	8,96	9,69
2020	2,11	1,78	1,88	4,97	5,11	5,66	6,55	7,14	8,05	8,25	8,89	9,48
2021	2,16	1,78	1,88	5,04	5,64	5,93	6,38	6,88	9,31	8,11	8,83	9,82

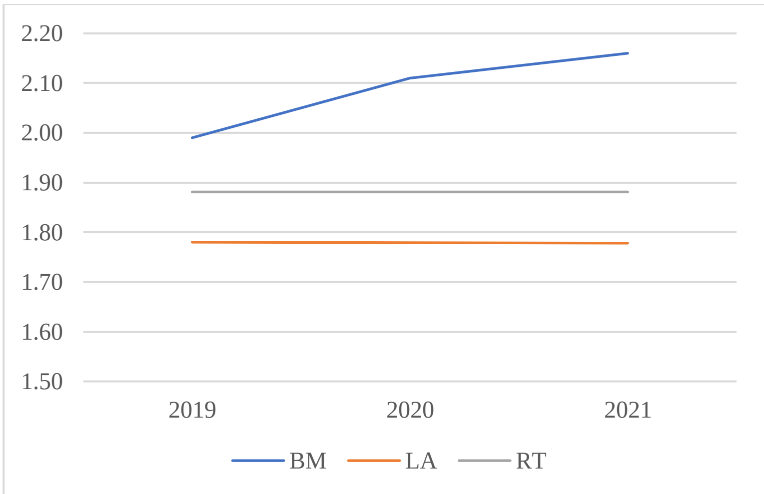
Similarly to what has been previously seen, unemployment forecasts based on latest available series are closer to benchmark for Denmark, Finland and France and seem to better follow the OECD's estimates pattern, as can be seen from Finnish data:

Figure 8 - Unemployment Forecasts, Finland



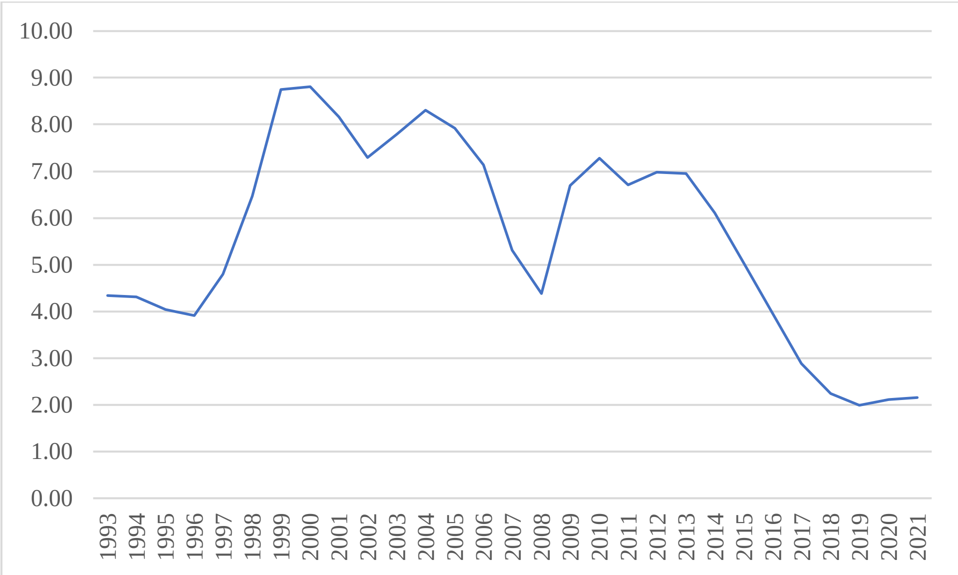
Czech Republic's benchmark pattern, on the other hand, seems to be entirely different compared to both LA and RT results, as can also be appreciated from the following graph:

Figure 9 - Unemployment Forecasts, Czech Republic



Given also that RT forecasts are all equal to 1.88% and LT ones equal 1.78%, the forecasting power of the fitted models seems to be impaired. The solution, which can be retrieved by looking at the Czech's unemployment time-series and at Table 2 in the second chapter, is straightforward: because of the split of the URSS in the early nineties and the turbulent period that followed, the series starts in 1993 and, as can be seen from Figure 10, it is very unstable. As a consequence of it, any forecasting exercise based solely on this kind of historical data would thus return misleading results.

Figure 10 – Unemployment rate in Czech Republic, 1993 – 2021, Latest Available data



Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

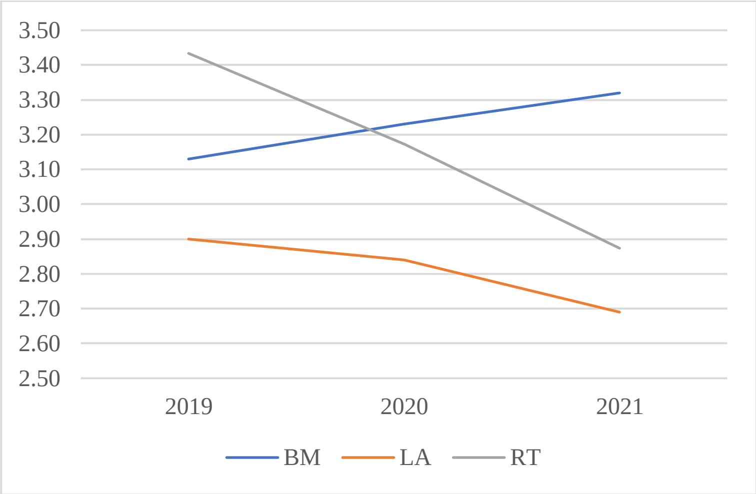
Germany, Greece, Hungary and Iceland forecasts now follow in Table 20:

Table 20 - Unemployment Forecasts, countries 9 - 12

	Germany			Greece			Hungary			Iceland		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2019	3,13	2,90	3,43	17,52	18,77	20,14	3,36	3,93	3,88	3,65	2,69	3,32
2020	3,23	2,84	3,17	16,25	18,35	18,41	3,20	4,54	4,52	4,11	2,67	3,25
2021	3,32	2,69	2,87	14,83	17,93	18,94	3,07	4,72	4,76	4,11	2,66	3,53

Given that, just like the Czech Republic, also Germany and Hungary have been impacted by the split of the URSS, predictably their forecasts seem to be quite unprecise, too: despite RT and LA forecasts being close to each other, they do not follow the benchmark at all:

Figure 11 - Unemployment Forecasts, Germany



Greek forecasted data, on the other hand, seem to properly follow BM both for LA and RT time series, with the former being closer to benchmark than the latter. Iceland's forecasts, to conclude, seem instead to be better represented by the real-time time series than the latest available one:

Figure 12 - Unemployment Forecasts, Iceland

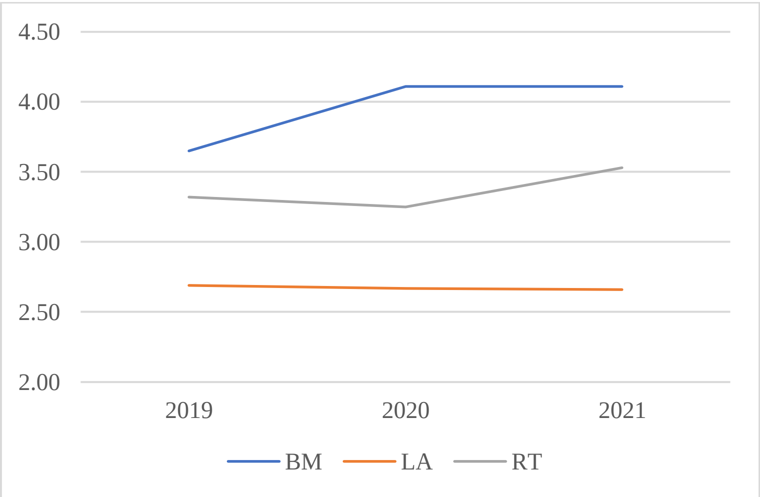


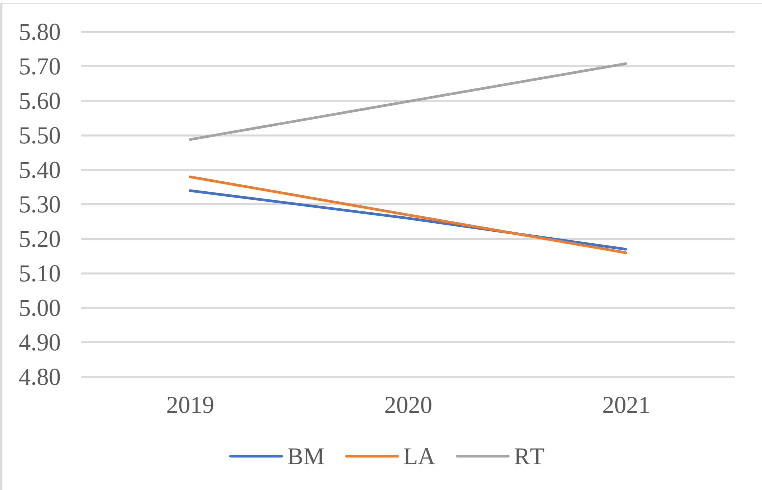
Table 21 shows forecasted data for Ireland, Italy, Japan and Luxemburg:

Table 21 - Unemployment Forecasts, countries 13 - 16

	Ireland			Italy			Japan			Luxemburg		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2019	5,27	7,03	7,06	10,03	10,99	10,75	2,38	2,26	2,11	5,34	5,38	5,49
2020	4,85	6,31	6,40	10,02	10,41	10,23	2,39	2,27	2,05	5,26	5,27	5,60
2021	4,56	8,17	7,78	10,15	10,75	10,54	2,29	2,29	1,99	5,17	5,16	5,71

Here, Irish unemployment forecasts do not seem to follow the benchmark properly, especially for 2021. Italy's data resembles instead the case of Iceland, in which the RT time-series is more precise than the LA's one. Japan and Luxemburg, to conclude, represent outstanding examples of successful forecasting and model fitting, as their latest available-based forecasts, despite being based only on historical data, are very close to benchmark, with the Luxembourger one being very close to it; forecasts based on RT time-series, on the other hand, present entirely different patterns which are far away from BM and LA data.

Figure 13 - Unemployment Forecasts, Luxemburg



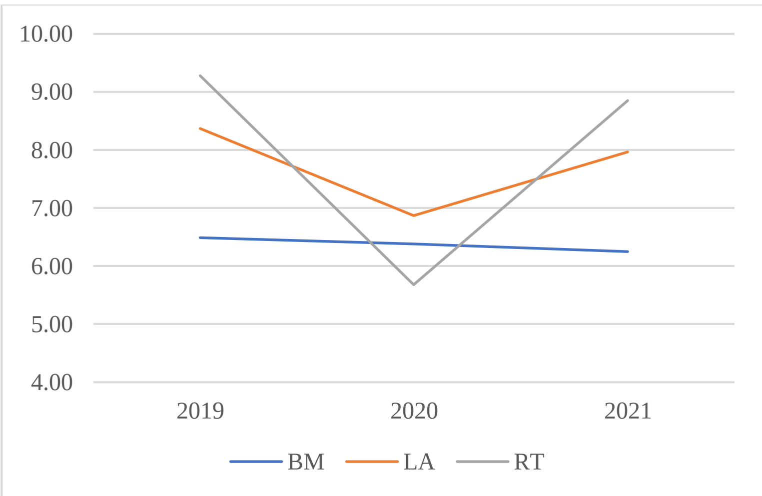
Forecasting results for the other 12 countries belonging to the sample resemble cases which have already been formerly seen. Netherlands, Poland, Sweden and Spain represent instances in which LA time-series tend to follow decently the benchmark, whereas their RT counterparts seem to be more unprecise. Mexico, Turkey and Switzerland seem to behave contrarily, as their own RT time-series are closer to BM than LA's ones.

Table 22 - Unemployment Forecasts, countries 17 - 28

	Mexico			Netherlands			New Zealand			Norway		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2019	3,47	3,75	3,63	3,41	3,83	4,73	4,13	4,11	n.a.	3,38	3,68	3,38
2020	3,45	4,01	3,59	3,46	3,93	4,43	4,21	4,06	n.a.	3,19	3,60	3,56
2021	3,30	4,16	3,87	3,61	3,92	4,68	4,33	4,01	n.a.	3,16	3,53	3,51
	Poland			Portugal			Spain			Sweden		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2019	3,42	3,27	4,45	6,49	8,37	9,28	14,24	16,39	16,75	6,76	6,36	5,90
2020	3,05	2,73	5,35	6,38	6,87	5,68	14,11	14,93	15,19	7,05	6,05	5,83
2021	2,82	2,28	5,92	6,25	7,97	8,85	13,55	15,77	17,00	6,98	6,11	5,76
	Switzerland			Turkey			United Kingdom			United States		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2019	4,50	4,81	4,41	13,53	10,46	10,90	3,85	4,04	5,10	3,68	4,20	5,18
2020	4,51	4,91	4,37	13,23	10,04	11,15	4,03	4,25	4,75	3,53	4,84	4,84
2021	4,47	5,10	3,98	13,04	9,71	11,32	4,13	4,21	5,70	3,70	5,00	5,64

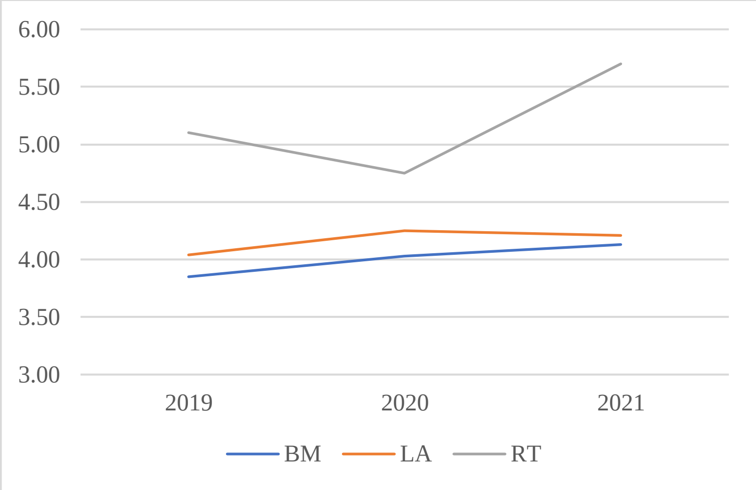
All other 6 countries lie in the middle, with cases of wrong predicted pattern for both latest available and real-time time series, such as Mexico:

Table 23 - Unemployment Forecasts, Mexico



With others, such as the United Kingdom, in which LA predictive power is very good, and estimates are very close to benchmark.

Table 24 - Unemployment Forecasts, United Kingdom



Now, given that forecasting results have been thoroughly argued and discussed for each country, let us present some aggregate numbers and face once again the very first question asked at the end of Chapter one, which can finally be answered: how do real-time and latest available data impact unemployment forecasting? Well, deviation of real-time unemployment forecasts from benchmark is, on average, equal to -0.77; latest available's is, on the other hand, equal to -0.29. If one were to consider absolute values, instead, real-time forecasts would, on average, deviate from benchmark by 1.13, whereas latest available ones would deviate only by 0.78. To conclude, considering absolute values, *unemployment forecasting carried out with the latter time-series is, on average, 46% more precise than the forecasts carried out with the formers*.

4.2.3. Economic shocks' impact on unemployment forecasting: Covid-19

Now that light has been shed on the impact of data revisions on unemployment forecasting, the analysis can be further deepened by looking at how economic shocks impact forecasting practices. Instead of taking as latest available OECD's Economic Outlook issue the one of December 2019, this insight will deal with the one of December 2020, featuring Covid-19 impact on latest available data. Pandemic impact will be here investigated on OECD's predictions and, most importantly, on the forecasting power of the model envisaged in the thesis work.

As everyone can imagine, Covid-19 had a major impact on countries' macroeconomic standing, and the chosen sample of 28 countries makes no exception. Unemployment-wise, thanks to job retention schemes job occupation in the short term has not been impacted as much as it would have been expected; nevertheless, dynamics for longer periods depend on countries' macroeconomic policies which, as a consequence of the pandemic, have been deeply altered compared to the ones of 2019. Given that the Economic Outlook's forecasting model considers both historical data and theory (i.e., as it has also been seen in Section 4.2.3.), such major alterations have been appraised and incorporated by OECD's economists in their forecasts, making up new benchmarks for 2020's nowcast and 2021 and 2022's forecasts (e.g., the nowcast is no more the 2019's unemployment value, as now the latest issue considered is the one of 2020 and, as such, all time-series are one period longer and forecasts are shifted one period ahead) [OECD (2019)].

In order to enhance data comparability, forecasting results stemming from the benchmark and the model envisaged in the thesis work will now be compared among the two different data vintages (i.e., 2019 and 2020). For the sake of simplicity, only forecasts pertaining to years 2020 and 2021 will be shown in what follows, as those are the only years in which prediction data can be found for both data vintages (i.e., Economic Outlook of December 2019 provides the nowcast and 2020 and 2021's forecasts, whereas the 2020 one provides the reader with the nowcast and forecast for 2021 and 2022). Two things can be noticed by looking at the table below: first, OECD estimates for 2020 and 2021 before and after Covid-19 are completely different, signalling that the major economic disruptions due to the pandemic have been taken into account by OECD's economists, which resulted in forecasts being deeply altered from one year to the other. Second, forecasting results based on LA and RT time-series are influenced by 2020's outliers and do not seem to follow the benchmark very much anymore; this, as it is going to be shown in the analysis that follows, is a totally predictable result due to the thesis work's model being solely based on time-series historical data, not allowing for any other incorporation of information.

Table 25 - Forecasting results comparison

		No Covid-19 Influence		Covid-19 affected data	
		2020	2021	2020	2021
Australia	BM	5.28	5.21	6.76	7.93
	LA	5.16	5.15	5.15	5.15
	RT	5.51	5.58	5.28	5.36
Austria	BM	4.55	4.59	5.6	5.63
	LA	4.71	4.65	3.89	3.56
	RT	5.38	5.38	n.a.	n.a.
Belgium	BM	5.50	5.46	5.71	7.94
	LA	5.30	5.31	5.47	5.47
	RT	6.81	7.21	5.72	6.23
Canada	BM	5.77	5.76	9.58	8.75
	LA	6.24	6.79	6.22	6.1
	RT	6.37	6.56	6.51	7.25
Czech Rep.	BM	2.11	2.16	2.63	3.62
	LA	1.78	1.78	4.15	5.8
	RT	1.88	1.88	2.1	2.1
Denmark	BM	4.97	5.04	5.71	6.16
	LA	5.11	5.64	5.12	5.08
	RT	5.66	5.93	5.62	5.56
Finland	BM	6.55	6.38	7.93	8.31
	LA	7.14	6.88	6.76	6.56
	RT	8.05	9.31	n.a.	n.a.
France	BM	8.25	8.11	8.38	10.47
	LA	8.89	8.83	8.13	8.07
	RT	9.48	9.82	9.35	9.08
Germany	BM	3.23	3.32	4.23	4.8
	LA	2.84	2.69	2.96	2.8
	RT	3.17	2.87	n.a.	n.a.
Greece	BM	16.25	14.83	16.87	17.84
	LA	18.35	17.93	18.15	16.64
	RT	18.41	18.94	17.35	16.88
Hungary	BM	3.20	3.07	5.03	6.41
	LA	4.54	4.72	6.07	5.5
	RT	4.52	4.76	4.36	4.33
Iceland	BM	4.11	4.11	5.42	7.54
	LA	2.67	2.66	3.74	3.51
	RT	3.25	3.53	3.26	3.7
Ireland	BM	4.85	4.56	5.3	7.96
	LA	6.31	8.17	6.19	5.64
	RT	6.40	7.78	5.53	6.05

Italy	BM	10.02	10.15	9.35	11
	LA	10.41	10.75	10.37	9.69
	RT	10.23	10.54	9.96	10.1
Japan	BM	2.39	2.29	2.8	2.94
	LA	2.27	2.29	2.93	2.78
	RT	2.05	1.99	2.71	2.94
Luxembourg	BM	5.26	5.17	6.44	7.01
	LA	5.27	5.16	5.46	5.57
	RT	5.60	5.71	5.15	5.51
Mexico	BM	3.45	3.30	5.33	5.01
	LA	4.01	4.16	4.17	4.37
	RT	3.59	3.87	3.6	3.7
Netherlands	BM	3.46	3.61	4.11	6.13
	LA	3.93	3.92	3.49	3.55
	RT	4.43	4.68	4.39	4.26
New Zealand	BM	4.21	4.33	4.87	5.82
	LA	4.06	4.01	3.88	3.86
	RT	n.a.	n.a.	5.58	5.62
Norway	BM	3.19	3.16	4.47	4.95
	LA	3.60	3.53	3.52	3.3
	RT	3.56	3.51	3.67	3.66
Poland	BM	3.05	2.82	3.76	5.5
	LA	2.73	2.28	2.92	2.53
	RT	5.35	5.92	8.19	7.16
Portugal	BM	6.38	6.25	7.32	9.51
	LA	6.87	7.97	6.84	6.42
	RT	5.68	8.85	7.99	8.45
Spain	BM	14.11	13.55	15.76	17.43
	LA	14.93	15.77	14.89	14.04
	RT	15.19	17.00	15.1	14.42
Sweden	BM	7.05	6.98	8.6	8.96
	LA	6.05	6.11	7.06	6.86
	RT	5.83	5.76	7.15	7.14
Switzerland	BM	4.51	4.47	4.92	5.21
	LA	4.91	5.10	4.23	4.32
	RT	4.37	3.98	4.21	4.33
Turkey	BM	13.23	13.04	12.55	14.78
	LA	10.04	9.71	n.a.	n.a.
	RT	11.15	11.32	n.a.	n.a.
UK	BM	4.03	4.13	4.64	7.4
	LA	4.25	4.21	n.a.	n.a.
	RT	4.75	5.70	4.58	4.45

USA	BM	3.53	3.70	8.09	6.36
	LA	4.84	5.00	4.88	4.76
	RT	4.84	5.64	4.72	4.61

Let us now analyse only forecasting results generated by real-time and latest available time-series featuring Covid-19 shocks (i.e., results based on 2020 as latest available issue at hand), in order to see how much forecasting power is lost due to the pandemic's economic shock. The following table, which features a structure very similar to the ones of the previous section, shows 2020's unemployment nowcast and 2021 and 2022's unemployment forecasts for the benchmark (i.e., BM, from OECD predictions) and for the forecasting predictions generated by the thesis work's model based on latest available (i.e., LA) and real-time (i.e., RT) time series.

By looking at the table, two things can be noticed: first, some forecasting results are not available, which means that flat log pseudolikelihoods have been encountered during model fitting, signalling time-series instability. Second, prediction based on real-time time-series seems to be slightly closer to benchmark compared to forecasts based on latest available series, which is exactly the opposite compared to what has been found in the previous section. To be precise, deviation of real-time based unemployment forecasts from benchmark is, on average, equal to 0.83 (in the previous section it was equal to -0.77); latest available's is, on the other hand, equal to 1.17 (previously it was equal to -0.29). If one were to consider absolute values, results confirm the opposite pattern compared to last section, which is even more pronounced as real-time forecasts, on average, deviate from benchmark by 1.28 (1.13 with 2019 latest available data), whereas latest available forecasts deviate even more, with a value equal to 1.49 (0.78 with 2019 latest available data). To sum up, considering absolute values, unemployment forecasting carried out with LA time-series is, on average, 16% less precise than the forecasts carried out with RTs (in the previous analysis, forecasts based on latest available time-series were more precise by 46%).

Table 26 - Unemployment Forecasts with 2020 Data

	Australia			Austria			Belgium			Canada		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2020	*6.76*	*5.15*	*5.28*	*5.60*	*3.89*	*n.a.*	*5.71*	*5.47*	*5.72*	*9.58*	*6.22*	*6.51*
2021	*7.93*	*5.15*	*5.36*	*5.63*	*3.56*	*n.a.*	*7.94*	*5.47*	*6.23*	*8.75*	*6.10*	*7.25*
2022	*7.35*	*5.15*	*5.43*	*5.09*	*3.63*	*n.a.*	*6.80*	*5.47*	*6.65*	*7.74*	*6.50*	*7.25*
	Czech Republic			Denmark			Finland			France		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2020	*2.63*	*4.15*	*2.1*	*5.71*	*5.12*	*5.62*	*7.93*	*6.76*	*n.a.*	*8.38*	*8.13*	*9.35*
2021	*3.62*	*5.80*	*2.1*	*6.16*	*5.08*	*5.56*	*8.31*	*6.56*	*n.a.*	*10.47*	*8.07*	*9.08*
2022	*3.60*	*5.80*	*2.11*	*5.70*	*5.14*	*5.8*	*7.66*	*6.62*	*n.a.*	*10.21*	*8.02*	*9.78*
	Germany			Greece			Hungary			Iceland		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2020	*4.23*	*2.96*	*n.a.*	*16.87*	*18.15*	*17.35*	*5.03*	*6.07*	*4.36*	*5.42*	*3.74*	*3.26*
2021	*4.80*	*2.8*	*n.a.*	*17.84*	*16.64*	*16.88*	*6.41*	*5.50*	*4.33*	*7.54*	*3.51*	*3.7*
2022	*4.33*	*2.64*	*n.a.*	*17.22*	*17.28*	*16.75*	*5.73*	*7.84*	*5.71*	*7.27*	*3.34*	*3.52*
	Ireland			Italy			Japan			Luxemburg		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2020	*5.30*	*6.19*	*5.53*	*9.35*	*10.37*	*9.96*	*2.80*	*2.93*	*2.71*	*6.44*	*5.46*	*5.15*
2021	*7.96*	*5.64*	*6.05*	*11.00*	*9.69*	*10.1*	*2.94*	*2.78*	*2.94*	*7.01*	*5.57*	*5.51*
2022	*7.76*	*7.54*	*6.47*	*10.92*	*10.14*	*10.05*	*2.85*	*2.78*	*3.12*	*6.43*	*5.67*	*5.61*
	Mexico			Netherlands			New Zealand			Norway		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2020	*5.33*	*4.17*	*3.6*	*4.11*	*3.49*	*4.39*	*4.87*	*3.88*	*5.58*	*4.47*	*3.52*	*3.67*
2021	*5.01*	*4.37*	*3.7*	*6.13*	*3.55*	*4.26*	*5.82*	*3.86*	*5.62*	*4.95*	*3.30*	*3.66*
2022	*4.80*	*4.54*	*3.78*	*6.34*	*3.63*	*4.53*	*5.41*	*3.80*	*5.53*	*4.42*	*3.30*	*3.67*
	Poland			Portugal			Spain			Sweden		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2020	*3.76*	*2.92*	*8.19*	*7.32*	*6.84*	*7.99*	*15.76*	*14.89*	*15.1*	*8.60*	*7.06*	*7.15*
2021	*5.50*	*2.53*	*7.16*	*9.51*	*6.42*	*8.45*	*17.43*	*14.04*	*14.42*	*8.96*	*6.86*	*7.14*
2022	*4.25*	*2.14*	*11.05*	*8.17*	*6.72*	*8.45*	*16.94*	*14.61*	*16.01*	*7.98*	*6.69*	*7.12*
	Switzerland			Turkey			United Kingdom			United States		
	BM	LA	RT	BM	LA	RT	BM	LA	RT	BM	LA	RT
2020	*4.92*	*4.23*	*4.21*	*12.55*	*n.a.*	*n.a.*	*4.64*	*n.a.*	*4.58*	*8.09*	*4.88*	*4.72*
2021	*5.21*	*4.32*	*4.33*	*14.78*	*n.a.*	*n.a.*	*7.40*	*n.a.*	*4.45*	*6.36*	*4.76*	*4.61*
2022	*4.81*	*4.41*	*4.46*	*15.34*	*n.a.*	*n.a.*	*6.25*	*n.a.*	*5.3*	*5.61*	*5.38*	*5.13*

Given these results, it can be argued that the pandemic not only impaired the forecasting ability of the thesis work's model (e.g., deviations from benchmark are much higher, especially for LA), but it also made forecasts based on real-time time series more precise than the ones based on latest availables', which is the opposite result compared to what has been found in the previous section and, more in general, to what one would expect when carrying out such an analysis, given the information-addition role of data revisions and the empirical literature discussed in Chapter 1.

CONCLUSIONS

Forecasting with macroeconomic variables is no easy matter, and the analysis carried out in this work of thesis has proved it. The national agencies' revision practices, necessary in order to maximise data quality due to data collection inefficiencies, complicate prediction models by creating two versions of data, revised and unrevised ones that, if not properly accounted for, might lead to potentially misleading conclusions. Needless to say, predictions vary in structure and features depending on the chosen macroeconomic variable, thereby making them respond to revisions differently. It has indeed been seen how, for example, consumption-wealth ratios based on latest available time series successfully predict stock-market returns, whereas predictions based on their real-time counterparts fail to do so. Exchange rates, on the other hand, behave contrarily as predictions based on real-time data prove more efficient compared to forecasting with latest available series; further, such kind of variable is found to be more sensitive to the chosen data vintage compared to others. Such empirical literature on the subject, despite covering many areas, focuses nevertheless on mainstream macroeconomic variables and most analyses are based on Anglo-Saxon databases. This is where the thesis work sneaks in: by choosing an out of the ordinary variable, this being the unemployment rate, and a large sample of 28 countries, revisions and the resulting real-time and latest available series have first been thoroughly analysed, in order to grasp their relevance when dealing with unemployment rates; such preliminary analysis has then been followed by a forecasting exercise aiming to shed light on which kind of data, between latest available and real-time, works best at predicting unemployment.

The preliminary analysis showed that national agencies do revise unemployment data in the early years following the first release and, contrarily to what one would expect, they keep on doing it also 10, even 20 years after first releases, with a seemingly decreasing cyclical pattern which seems to be more pronounced during periods of economic crisis. Expectedly, such persisting revisioning does affect latest available time-series, which prove to be less volatile than their unrevised counterparts. As a second step, unemployment has been forecasted through the usage of time-series-based forecasting models, with results that align the variable's behaviour to the majority of the available empirical literature on other macroeconomic variables: unemployment forecasting based on latest available time-series has indeed turned out to be, on average, 46% closer to the chosen benchmark than the predictions based on real-time series. In order to get successful results when predicting unemployment rates, therefore, this work of thesis proves that the former time-series should be used.

To conclude, an economist trying to forecast any macroeconomic variable should consider one final warning-bell: the presence of economic shocks. It has indeed been argued how the Covid-19 pandemic impaired the forecasting power of the work's envisaged model which, by making forecasts based on real-time series more precise than the ones based on latest availables', resulted into utterly different outcomes. Economic shocks translate indeed into outliers deeply altering time-series' structures that, if not properly accounted for by a forecaster with a time-series-based prediction model at hand, would make him end up with misleading results and a blind forecasting tool.

APPENDIX

A.1. Unemployment data sources and country notes

The following table outlines the sources of macroeconomic data of the 28 countries belonging to the analysis' sample, together with some remarks regarding assumptions and adjustments made by national agencies and/or the OECD.

Table A.1 - Unemployment data sources (last update: Dec 14th, 2020)

Country	Sources	Remarks
Australia	Australian Bureau of Statistics (ABS)	Data from monthly Household Labour Force Survey; working age: 15 y.o. Persons laid-off for less than four weeks (because of bad weather or plant breakdown) are included among the employed; all other layoffs are considered as unemployed or out of the labour force. Quarterly data are seasonally adjusted at source.
Austria	Österreichisches Institut für Wirtschaftsforschung (WIFO)	Data from the national Labour Force Survey; working age: 15 y.o. The unemployment series is corrected for statistical breaks. The labour force is derived as the sum of total employment and unemployment.
Belgium	National Bank of Belgium (NBB)	Data from Labour Force Survey (LFS); working age: 15 y.o. The series is monthly adjusted by using the administrative national unemployment figures (e.g., it complies with Eurostat methodology).
Canada	Statistics Canada	Data from monthly Labour Force Survey; working age: 15 y.o. The sample also includes people who, while not actively looking for work in the preceding four weeks, were available for work, but were on temporary layoff or had a new job to start in four weeks or less. Quarterly data are seasonally adjusted at source.
Czech Republic	Czech Statistical Office	Data from quarterly Labour Force Survey; working age: 15 y.o. Labour force is derived from total employment and unemployment. Quarterly data are seasonally adjusted by the OECD.
Denmark	Statistics Denmark	Data harmonised by Eurostat. Seasonal adjustment performed by the OECD.
Finland	Statistics Finland	Data from quarterly Labour Force Survey; working age: 15 y.o. Unemployment is derived from labour force and total employment. Quarterly data are seasonally adjusted by the OECD.

France	Institut National de la Statistique et des Etudes Economiques (INSEE)	Data from "quarterly unemployment rate (ILO definition) by gender and age (%)"; working age: 15y.o.; the sample relates to all people living in France including oversea departments.
Germany	Statistisches Bundesamt	Data gathering based on ILO concept; working age: 15 to 74 y.o.; recalculation according to the results of Population and Housing Census 2011.
Greece	Hellenic Statistical Authority (ELSTAT)	Data from quarterly Labour Force Statistics; working age: 15 y.o. Labour force is derived from employment and unemployment. Quarterly data are seasonally adjusted by the OECD.
Hungary	Hungarian Central Statistical Office	Data from Labour Force Survey; working age: 15 y.o.
Iceland	Statistics Iceland	Data from quarterly Labour Force Survey, though up to 2003Q1 the figures are from the annual Labour Force Survey; working age: 16 to 74 y.o. Unemployment is derived from total employment and labour force.
Ireland	Central Statistics Office (CSO)	Data from Quarterly National Household Survey, though up to 1997 the figures are from the annual Labour Force Survey and relate to mid-April of each year; working age: 15 y.o.
Italy	Istituto Nazionale di Statistica (ISTAT)	Number of unemployed people who are 15 years old and over.
Japan	Economic and Social Research Institute (ESRI), Cabinet Office (CAO)	Working age: 15 y.o. From March to September 2011, official data excluded Tohoku area. However, in April 2012, the statistics bureau provided estimates for the area. The figures are not official series but referential ones for time-series comparison.[3]
Luxembourg	Service central de la Statistique et des Études Économiques (STATEC)	Seasonally adjusted STATEC data.
Mexico	National Institute of Statistics and Geography (INEGI)	Data taken from the National Survey of Employment (ENE) concerning labour force, total employment and unemployment as well as self-employment and dependent employment, covering the whole territory. Quarterly data are seasonally adjusted by the OECD.
Netherlands	Statistics Netherlands	Derived from Monthly data. Source CBS. Since 2015, labour data are fully consistent with ILO definitions.
New Zealand	Statistics New Zealand	Data from quarterly Household Labour Force Survey; working age: 16 y.o. Labour force is derived from employment and unemployment. Quarterly data are seasonally adjusted at source.

[3] Ministry of Internal Affairs and Communications, https://www.soumu.go.jp/english/

Norway	Statistics Norway	Data for labour force and unemployment are taken from the quarterly Labour Force Survey and refer to the population aged 15 and over. Total employment is derived from labour force and unemployment. Quarterly data are seasonally adjusted by the OECD.
Poland	Central Statistical Office (GUS)	Data from quarterly Household Labour Force Survey from 2010Q1. The survey covers members of randomly selected households. Data refer to the non-institutional population and cover all persons aged 15 years and over living in households continuously for at least two months.[4]
Portugal	Instituto Nacional de Estatistica (INE)	Data from quarterly Labour Force Survey; working age: 15 y.o. Unemployment is derived from labour force and total employment.
Spain	Instituto Nacional de Estadistica (INE)	Data from the Economically Active Population Survey; working age: 16 y.o.
Sweden	Statistiska centralbyran (SCB)	Seasonally adjusted values; working age: 15 to 74 y.o.
Switzerland	Federal Statistical Office (FSO)	Data gathering based on ILO definitions; working age: 15 y.o.
Turkey	Turkish Statistical Institute (TURKSTAT)	Data from monthly Labour Force Survey; working age: 15 y.o. Quarterly data are seasonally adjusted by the OECD.
United Kingdom	Office for National Statistics (ONS)	Data are taken from the quarterly Labour Force Survey and refer to the population aged 16 and over. Unemployment is derived from total employment (civilian employment plus the armed forces) and the labour force.
United states	Bureau of Labor Statistics	Data from monthly Current Population Survey; working age: 16 y.o. The data are seasonally adjusted by the Bureau of Labor Statistics (i.e., at source).

OECD's Economic Outlook Statistical Sources, https://www.oecd.org/economy/outlook/statistical-annex/

[4] The population not living in private households is excluded, such as enlisted soldiers in military barracks, persons in jail, and persons with no place of residence. Career members of the armed forces who live in private households are included in civilian labour force. The armed forces only include conscripts.

A.2. Real-time series descriptive statistics

Table A.2 shows basic descriptive statistics of the sample's real-time series (e.g., data on the main diagonal in Table 1). As it can be seen, all time-series' observations equal 24; this happens because earliest real-time values in the sample relate to year 1996, which have been first released in the same year's 60th Issue of the Economic Outlook. The latest real-time observations, on the other hand, relate to year 2019 which have been released in 2019's Issue No. 106, the newest issue at hand; the number of real-time observations is, therefore, fixed.

Table A.2 - Real-time series descriptive statistics

	Observations	Mean	Std. dev.	Min	Max
Australia	24	5.982	1.186	4.252	8.705
Austria	24	5.348	0.668	4.214	6.322
Belgium	24	8.393	1.889	5.513	12.858
Canada	24	7.247	1.002	5.650	9.609
Czech Republic	24	6.104	2.087	1.991	8.970
Denmark	24	5.814	1.415	3.119	8.937
Finland	24	9.073	2.289	6.210	16.444
France	24	9.751	1.238	7.345	12.444
Germany	24	7.055	2.394	3.125	11.400
Greece	24	14.680	6.546	7.639	27.213
Hungary	24	7.465	2.359	3.358	11.320
Iceland	24	3.751	1.742	1.311	7.495
Ireland	24	8.033	3.755	4.210	14.792
Italy	24	9.898	2.028	5.936	12.440
Japan	24	4.094	0.893	2.380	5.450
Luxembourg	24	4.794	1.476	2.503	7.117
Mexico	24	3.929	1.087	2.394	5.995
Netherlands	24	4.563	1.424	2.474	6.899
New Zealand	24	5.447	1.273	3.587	8.311
Norway	24	3.666	0.599	2.530	4.748
Poland	24	11.184	4.885	3.422	19.723
Portugal	24	8.519	3.532	4.122	16.670

Spain	24	16.645	5.560	8.067	26.437
Sweden	24	6.453	1.400	3.955	8.447
Switzerland	24	3.936	0.867	1.781	5.280
Turkey	24	9.624	2.073	6.129	14.622
United Kingdom	24	5.946	1.345	3.848	8.085
United States	24	5.710	1.732	3.677	9.665

OECD's Economic Outlook Statistical Sources, https://www.oecd.org/economy/outlook/statistical-annex/

A.3. Latest available series descriptive statistics (full)

Table A.3 shows basic descriptive statistics of the sample's latest available series (e.g., data on the far-right in Table 1). In this case observations are not fixed, but instead change depending on the length of the time-series as provided by countries' national agencies. Let us consider Greece as an example: the total number of observations at hand, which can be found in Table 2, equals 65; by looking at the table below, though, latest available observations are only 27, this happens because in Issue 106, December 2019, Greece' ELSTAT released revised data that did not go all the way back to 1956, but stopped at 1995 instead. At this point, filling previous years of data with older issues' information would not be correct, as such information belongs to different data vintages and, as such, cannot be considered latest available.

Table A.3 - Latest available series descriptive statistics (full sample)

	Observations	Mean	Std. dev.	Min	Max
Australia	58	5.731	2.435	1.256	10.874
Austria	53	3.697	1.374	1.079	6.014
Belgium	62	6.277	3.074	0.916	10.825
Canada	62	7.367	1.972	3.336	11.994
Czech Republic	29	5.615	2.147	1.991	8.814
Denmark	53	5.755	2.159	0.865	10.223
Finland	62	6.562	4.083	1.095	17.764
France	62	6.728	3.283	1.047	10.679
Germany	29	7.037	2.425	3.125	11.021
Greece	27	14.752	6.221	7.760	27.466
Hungary	30	7.565	2.652	3.070	11.965

Iceland	58	2.788	1.697	0.893	7.552
Ireland	32	9.261	4.375	4.177	16.014
Italy	62	7.566	2.910	2.691	12.620
Japan	62	2.770	1.264	1.124	5.357
Luxembourg	37	3.699	1.905	1.028	7.075
Mexico	31	4.326	0.989	3.301	7.847
Netherlands	62	5.007	2.938	0.571	12.740
New Zealand	62	3.928	2.913	0.043	10.654
Norway	50	3.303	1.185	1.390	5.787
Poland	29	10.979	5.088	2.820	19.843
Portugal	62	6.668	3.180	2.423	16.183
Spain	45	14.843	4.968	4.319	26.094
Sweden	62	5.215	2.896	1.540	11.665
Switzerland	47	3.003	1.875	0.186	5.062
Turkey	62	8.634	1.909	5.606	13.529
United Kingdom	62	6.133	2.589	2.666	11.774
United States	62	5.890	1.619	3.507	9.719

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

A.4. Latest available series descriptive statistics (truncated)

Similarly to table A.3, table A.4 shows basic descriptive statistics of the sample's latest available series, with the difference that in this case it shows only data from 1996 to 2019; as a result of it, all series have the same length, equal to 24.

Table A.4 - Latest available series descriptive statistics (truncated)

	Observations	Mean	Std. dev.	Min	Max
Australia	24	5.884	1.107	4.234	8.506
Austria	24	4.767	0.722	3.535	6.014
Belgium	24	7.836	1.010	5.513	9.550
Canada	24	7.216	0.980	5.650	9.625
Czech Republic	24	6.078	2.021	1.991	8.814
Denmark	24	5.790	1.169	3.710	7.796
Finland	24	8.957	2.019	6.364	15.757

France	24	9.197	0.969	7.337	10.679
Germany	24	7.211	2.394	3.125	11.021
Greece	24	14.896	6.522	7.760	27.466
Hungary	24	7.405	2.295	3.358	11.178
Iceland	24	7.405	2.295	3.358	11.178
Ireland	24	8.330	3.937	4.177	15.449
Italy	24	9.637	1.912	6.130	12.620
Japan	24	4.072	0.868	2.380	5.357
Luxembourg	24	4.535	1.589	2.211	7.075
Mexico	24	4.277	0.805	3.301	6.228
Netherlands	24	5.096	1.334	2.974	7.472
New Zealand	24	5.327	1.133	3.590	7.735
Norway	24	3.607	0.594	2.493	4.679
Poland	24	11.288	4.967	3.422	19.843
Portugal	24	8.532	3.520	4.009	16.183
Spain	24	15.768	5.417	8.232	26.094
Sweden	24	7.628	1.527	5.834	11.665
Switzerland	24	4.382	0.622	2.773	5.062
Turkey	24	9.339	2.000	5.997	13.529
United Kingdom	24	5.919	1.340	3.848	8.111
United States	24	5.696	1.723	3.677	9.623

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

A.5. Unemployment rates in 1996

The following two tables show sample countries' unemployment rates in year 1996, with data vintages spanning from 1996 to 2019. This view is particularly useful as it enables the user to see yearly data revisions on 1996's unemployment rates.

Table A.5.1 - Unemployment rates in 1996 (data vintages 1996 -2007)

Data Vintage	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Australia	8.42	8.52	8.52	8.44	8.46	8.14	8.14	8.19	8.19	8.19	8.19	8.19
Austria	6.21	6.32	6.32	6.32	5.56	5.64	5.64	5.64	5.64	5.60	5.60	5.60
Belgium	12.86	12.85	12.81	12.66	9.73	9.73	9.54	9.54	9.54	9.54	9.54	9.54
Canada	9.61	9.72	9.72	9.71	9.65	9.65	9.65	9.65	9.65	9.74	9.65	9.65
Czech Rep.	3.02	3.46	3.37	3.94	3.94	3.94	3.94	3.94	3.94	3.94	3.94	3.94
Denmark	8.94	8.79	8.67	8.71	6.80	6.78	6.32	6.32	6.32	6.27	6.27	6.27
Finland	16.44	16.28	14.58	14.56	14.59	14.59	14.61	14.61	14.61	14.61	15.87	15.87
France	12.42	12.31	12.29	12.32	12.32	12.12	12.03	12.03	12.03	12.12	12.12	10.48
Germany	10.34	10.32	10.33	8.82	8.56	8.54	8.35	8.35	8.35	7.71	7.72	7.71
Greece	10.15	10.33	10.34	10.34	9.80	9.76	9.80	9.80	9.80	9.80	9.80	9.80
Hungary	10.58	10.04	10.11	10.11	10.11	10.11	10.11	10.11	10.11	10.11	10.11	10.11
Iceland	4.30	4.40	4.35	4.35	4.34	4.37	3.72	3.72	3.72	3.72	3.72	3.72
Ireland	11.96	11.94	11.88	11.88	11.69	11.69	11.69	11.69	11.43	12.02	12.02	12.02
Italy	12.16	12.09	12.09	11.74	11.74	11.74	11.74	11.74	11.74	11.74	11.34	11.35
Japan	3.35	3.35	3.35	3.35	3.35	3.35	3.35	3.36	3.36	3.36	3.35	3.35
Luxembourg	3.09	3.28	3.27	3.27	3.27	3.26	3.26	3.26	3.26	3.26	3.26	3.26
Mexico	6.00	5.49	5.49	5.49	5.66	5.66	5.66	5.49	4.34	4.34	4.34	5.25
Netherlands	6.57	6.65	6.64	6.65	6.65	6.65	6.65	6.65	6.33	6.33	6.33	6.33
New Zealand	6.19	6.12	6.11	6.11	6.11	6.10	6.10	6.10	6.10	6.10	6.10	6.10
Norway	4.19	4.88	4.83	4.83	4.83	4.83	4.83	4.83	4.83	4.82	4.82	4.83
Poland	12.48	12.44	12.34	12.34	12.34	12.34	12.34	12.34	12.34	12.34	12.34	12.34
Portugal	7.18	7.30	7.30	7.30	7.30	7.27	7.28	7.26	7.25	7.25	7.25	7.25
Spain	22.66	22.21	22.21	22.21	22.21	22.21	17.47	17.47	17.47	17.54	17.54	17.54
Sweden	7.89	8.05	8.05	8.04	8.04	8.04	8.04	8.04	8.04	8.05	8.05	8.05
Switzerland	4.61	4.66	4.66	4.66	4.66	4.66	4.66	3.76	3.76	3.76	3.76	3.76
Turkey	7.22	6.50	6.03	6.03	6.03	6.35	6.35	6.35	6.48	6.48	6.48	6.49
UK	7.65	7.99	7.99	7.99	7.91	7.91	7.92	7.97	8.09	8.09	8.09	8.09
USA	5.38	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

Table A.5.2 - Unemployment rates in 1996 (data vintages 2008 -2019)

Data Vintage	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Australia	8.19	8.19	8.17	8.49	8.55	8.55	8.51	8.51	8.51	8.51	8.51	8.51
Austria	5.86	5.86	4.16	4.16	4.16	4.18	4.18	4.18	4.18	4.18	4.18	4.18
Belgium	9.55	9.55	9.55	9.55	9.55	9.55	9.55	9.55	9.55	9.55	9.55	9.55
Canada	9.65	9.65	9.65	9.63	9.63	9.63	9.63	9.63	9.63	9.63	9.63	9.63
Czech Rep.	3.94	3.94	3.94	3.94	3.92	3.92	3.92	3.92	3.92	3.92	3.91	3.91
Denmark	6.29	6.29	6.29	6.29	6.29	6.33	6.33	6.33	6.33	6.33	6.33	6.82
Finland	15.87	15.87	15.87	15.87	15.86	15.86	15.86	15.86	15.86	15.86	15.76	15.76
France	10.57	10.59	10.59	10.56	10.56	10.55	10.15	10.15	10.52	10.52	10.53	10.53
Germany	8.56	8.56	8.56	8.52	8.51	8.95	8.95	8.90	8.90	8.90	8.90	8.90
Greece	9.00	10.72	10.72	10.72	10.72	10.72	10.72	10.31	10.31	10.19	10.18	10.26
Hungary	10.11	10.11	10.11	10.11	10.02	10.02	10.02	10.02	10.03	10.02	10.00	10.00
Iceland	3.72	3.72	3.72	3.72	3.72	3.72	3.70	3.70	3.76	3.75	3.70	3.70
Ireland	11.83	11.94	11.87	11.87	11.87	12.10	12.06	12.07	11.88	11.87	12.14	12.14
Italy	11.35	11.35	11.16	11.14	11.15	11.17	11.18	11.19	11.18	11.18	11.18	11.18
Japan	3.35	3.35	3.35	3.35	3.35	3.35	3.35	3.35	3.35	3.35	3.35	3.35
Luxembourg	3.26	3.26	3.26	3.26	2.86	2.86	2.86	2.86	2.86	2.85	2.84	2.83
Mexico	5.25	5.25	5.32	5.32	5.28	5.28	5.24	5.24	5.24	5.32	6.23	6.23
Netherlands	6.63	6.63	5.70	6.19	6.19	6.19	6.10	7.46	7.46	7.46	7.47	7.47
New Zealand	6.10	6.30	6.30	6.30	6.31	6.31	6.31	6.31	6.30	6.30	6.30	6.30
Norway	4.82	4.82	4.82	4.82	4.66	4.66	4.66	4.66	4.66	4.66	4.67	4.67
Poland	12.34	12.34	12.34	12.34	12.34	12.36	12.36	12.36	12.36	12.36	12.29	12.29
Portugal	7.25	7.25	7.25	7.25	7.36	7.36	7.36	7.34	7.34	7.34	7.34	7.34
Spain	17.54	17.54	17.54	17.54	17.54	17.54	18.15	18.15	18.15	18.15	18.15	18.15
Sweden	11.57	11.57	11.57	11.40	11.40	11.40	11.57	11.57	11.40	11.40	11.40	11.40
Switzerland	3.90	3.90	3.90	3.58	3.58	3.58	4.19	4.19	3.61	4.20	4.20	4.20
Turkey	6.48	7.06	7.06	7.06	7.06	7.06	6.12	6.12	6.12	6.12	6.12	6.12
UK	8.10	8.10	8.10	8.10	8.10	8.10	8.10	8.10	8.10	8.10	8.10	8.10
USA	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

A.6. Unemployment rates in 1996 – net revision impacts

The following two tables show the net impact of revisions for sample countries' unemployment rates in year 1996, with data vintages spanning from 1997 to 2019 (i.e., 1996 is not present as it cancels out when subtracting). This view is particularly useful as it enables the user to see the yearly net amount of revisions in 1996's unemployment rates.

Table A.6.1 - Net revision impacts, UNR in 1996 (data vintages 1997 -2008)

Data Vintage	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Australia	0.10	0.00	-0.08	0.02	-0.32	0.00	0.05	0.00	0.00	0.00	0.00	0.00
Austria	0.11	0.00	0.00	-0.76	0.08	0.00	0.00	0.00	-0.04	0.00	0.00	0.26
Belgium	-0.01	-0.04	-0.15	-2.93	0.00	-0.19	0.00	0.00	0.00	0.00	0.00	0.01
Canada	0.11	0.00	-0.01	-0.06	0.00	0.00	0.00	0.00	0.09	-0.09	0.00	0.00
Czech Rep.	0.44	-0.09	0.57	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Denmark	-0.15	-0.12	0.04	-1.91	-0.02	-0.46	0.00	0.00	-0.05	0.00	0.00	0.02
Finland	-0.16	-1.70	-0.02	0.03	0.00	0.02	0.00	0.00	0.00	1.26	0.00	0.00
France	-0.11	-0.02	0.03	0.00	-0.20	-0.09	0.00	0.00	0.09	0.00	-1.64	0.09
Germany	-0.02	0.01	-1.51	-0.26	-0.02	-0.19	0.00	0.00	-0.64	0.01	-0.01	0.85
Greece	0.18	0.01	0.00	-0.54	-0.04	0.04	0.00	0.00	0.00	0.00	0.00	-0.80
Hungary	-0.54	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Iceland	0.10	-0.05	0.00	-0.01	0.03	-0.65	0.00	0.00	0.00	0.00	0.00	0.00
Ireland	-0.02	-0.06	0.00	-0.19	0.00	0.00	0.00	-0.26	0.59	0.00	0.00	-0.19
Italy	-0.07	0.00	-0.35	0.00	0.00	0.00	0.00	0.00	0.00	-0.40	0.01	0.00
Japan	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	-0.01	0.00	0.00
Luxembourg	0.19	-0.01	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mexico	-0.51	0.00	0.00	0.17	0.00	0.00	-0.17	-1.15	0.00	0.00	0.91	0.00
Netherlands	0.08	-0.01	0.01	0.00	0.00	0.00	0.00	-0.32	0.00	0.00	0.00	0.30
New Zealand	-0.07	-0.01	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Norway	0.69	-0.05	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.01	-0.01
Poland	-0.04	-0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Portugal	0.12	0.00	0.00	0.00	-0.03	0.01	-0.02	-0.01	0.00	0.00	0.00	0.00
Spain	-0.45	0.00	0.00	0.00	0.00	-4.74	0.00	0.00	0.07	0.00	0.00	0.00
Sweden	0.16	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	3.52
Switzerland	0.05	0.00	0.00	0.00	0.00	0.00	-0.90	0.00	0.00	0.00	0.00	0.14
Turkey	-0.72	-0.47	0.00	0.00	0.32	0.00	0.00	0.13	0.00	0.00	0.01	-0.01
UK	0.34	0.00	0.00	-0.08	0.00	0.01	0.05	0.12	0.00	0.00	0.00	0.01
USA	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

Table A.6.2 - Net revision impacts, UNR in 1996 (data vintages 2009 -2019)

Data Vintage	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Australia	0.00	-0.02	0.32	0.06	0.00	-0.04	0.00	0.00	0.00	0.00	0.00
Austria	0.00	-1.70	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
Belgium	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Canada	0.00	0.00	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Czech Rep.	0.00	0.00	0.00	-0.02	0.00	0.00	0.00	0.00	0.00	-0.01	0.00
Denmark	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.49
Finland	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.10	0.00
France	0.02	0.00	-0.03	0.00	-0.01	-0.40	0.00	0.37	0.00	0.01	0.00
Germany	0.00	0.00	-0.04	-0.01	0.44	0.00	-0.05	0.00	0.00	0.00	0.00
Greece	1.72	0.00	0.00	0.00	0.00	0.00	-0.41	0.00	-0.12	-0.01	0.08
Hungary	0.00	0.00	0.00	-0.09	0.00	0.00	0.00	0.01	-0.01	-0.02	0.00
Iceland	0.00	0.00	0.00	0.00	0.00	-0.02	0.00	0.06	-0.01	-0.05	0.00
Ireland	0.11	-0.07	0.00	0.00	0.23	-0.04	0.01	-0.19	-0.01	0.27	0.00
Italy	0.00	-0.19	-0.02	0.01	0.02	0.01	0.01	-0.01	0.00	0.00	0.00
Japan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Luxembourg	0.00	0.00	0.00	-0.40	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01
Mexico	0.00	0.07	0.00	-0.04	0.00	-0.04	0.00	0.00	0.08	0.91	0.00
Netherlands	0.00	-0.93	0.49	0.00	0.00	-0.09	1.36	0.00	0.00	0.01	0.00
New Zealand	0.20	0.00	0.00	0.01	0.00	0.00	0.00	-0.01	0.00	0.00	0.00
Norway	0.00	0.00	0.00	-0.16	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Poland	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	-0.07	0.00
Portugal	0.00	0.00	0.00	0.11	0.00	0.00	-0.02	0.00	0.00	0.00	0.00
Spain	0.00	0.00	0.00	0.00	0.00	0.61	0.00	0.00	0.00	0.00	0.00
Sweden	0.00	0.00	-0.17	0.00	0.00	0.17	0.00	-0.17	0.00	0.00	0.00
Switzerland	0.00	0.00	-0.32	0.00	0.00	0.61	0.00	-0.58	0.59	0.00	0.00
Turkey	0.58	0.00	0.00	0.00	0.00	-0.94	0.00	0.00	0.00	0.00	0.00
UK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
USA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Data from OECD, Economic Outlook Annual reports (1996-2019), https://www.oecd.org/economic-outlook/, personal reworking

BIBLIOGRAPHY

Barnichon, R. and Nekarda, C. J., (2012), *The Ins and Outs of Forecasting Unemployment: Using Labour Force Flows to Forecast the Labour Market*, Brookings Papers on Economic Activity, Fall 2012

Bassanetti, A., Caivano, M., and Locarno, A., (2010), *Modelling Italian potential output and the output gap*, Banca D'Italia, Number 771

Blanchard, O., Amighini, A. and Giavazzi, F., (2017), *Macroeconomics: A European Perspective*, Pearson

Cimadomo, J., (2008), *Fiscal Policies in Real Time*, European Central Bank (ECB), Number 919

Clements, M. P., Galvao, A. B., (2017), *Data Revisions and Real-Time Probabilistic Forecasting of Macroeconomic Variables*, SSRN

Cole, R. J., (1969), *The Use of Preliminary Data in Econometric Forecasting*, The Review of Economics and Statistics

Croushore, D., (2010), *Forecasting with Real-Time Data Vintages*, University of Richmond

Denton, F. T., and Kuiper, J., (1965), *The Effect of Measurement Errors on Parameter Estimates and Forecasts: A Case Study Based on the Canadian Preliminary National Accounts*, Review of Economics and Statistics, volume XLVII(2)

Diebold, F. X. and Rudebusch, G. D., (1991), *Forecasting Output with the Composite Leading Index: a Real-Time Analysis*, Journal of the American Statistical Association, volume 86 number 415

Fajingbesi, A. A., (2001), *Data Management in Macroeconomic Policy Analysis*, National Centre for Economic Management and Administration (NCEMA)

Filardo, A. J., (1999), *How realiable are recession prediction models?*, Economic Review, Federal Reserve Bank of Kansas City, Volume 84

Guo, W., (2003), *Dynamic State-Space Models*, Journal of Time-Series Analysis, Volume 24 Issue 2

Holden, K., and Peel, D. A., (1982), *Data revisions and time-series models of GDP and its components*, Applied Economics, 1982b

Koenig, E. F., (2003), *Is the Markup a useful Real-Time Predictor of Inflation?, Federal Bank of Philadelphia,* Research Department

Molodtsova, T. and Papell, D. H., (2008), *Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals*

Montgomery, A. L., Zarnowitz, V., Tsay, R. S. and Tiao, G. C., (1998*), Forecasting the U.S. Unemployment Rate*, Journal of the American Statistical Association, Volume 93 Issue 442

OECD (2019b), Employment Outlook, OECD Publishing, Paris

Okun, A. M., (1962), *Potential GNP: its Measurement and Significance*, Statistical Association, Proceedings of the Business and Economics Statistics Section

Nakamura, L. I. and Stark, T., (2007), *Mismeasured personal saving and the permanent income hypothesis*, Federal Bank of Philadelphia, Working Papers 06-7

Orphanides, A. and Van Norden, S., (2005), *The Realibility of Inflation Forecasts Based on Output Gap Estimates in Real-Time*, Journal of Money, Credit and Banking, Volume 37 Issue 3

Robertson, J. C. and Tallman, E. W., *Vector Autoregressions: Forecasting and Reality, Economic Review*, Federal Reserve Bank of Atlanta

Sargent, T., (1989), *Two Models of Measurements and the Investment Accelerator*, Journal of Political Economy, volume 97, issue 2

Stark, T., and Croushore, D., (2002), *Forecasting with a real-time data set for macroeconomists*, Journal of Macroeconomics, volume 24

SITOGRAPHY

OECD (2013-2019), Economic Outlook Annual reports, https://www.oecd.org/economic-outlook/

OECD's Economic Outlook Statistical Sources, https://www.oecd.org/economy/outlook/statistical-annex/

Organisation for Economic Cooperation and Development, https://www.oecd-ilibrary.org/economics/oecd-economic-outlook_16097408

OECD Statistics, https://stats.oecd.org

National Institute Global Econometric Model (NiGEM), National Institute of Economic and Social Research, https://www.niesr.ac.uk/national-institute-global-econometric-model-nigem

Forecasting Methods and Analytical Tools, Organisation for Economic Cooperation and Development, https://www.oecd.org/economy/outlook/forecastingmethodsandanalyticaltools.htm

STATA Time-Series Reference Manual, StataCorp LLC, https://www.stata.com/manuals/ts.pdf

Ministry of Internal Affairs and Communications, https://www.soumu.go.jp/english/