The limits of statistical information: How important are GDP revisions in Italy?*

Claudio Lupi
University of Molise, Campobasso

Franco Peracchi
"Tor Vergata" University, Rome

March 28, 2003

Abstract

The use of Gross Domestic Product (GDP) as a summary measure of the level of economic activity is pervasive in empirical economics, policy analysis and forecasting. This pervasive role of GDP and its nature of "public good" raises obvious problems of timeliness and accuracy of the data.

A striking feature of GDP data (and, more generally, of all national accounts figures) is the presence of "data vintages". That is, the GDP estimate for a specific year or quarter is subject to several revisions after its first release. As a result, both the level and the profile of GDP over a given period may change, sometimes substantially, through time.

This paper presents some evidence on the extent of GDP revisions in Italy, with particular emphasis on revisions of the quarterly national accounts series, and compares the Italian evidence with the available evidence from other countries. After discussing some areas in which data revisions can have potentially important consequences, the paper concludes with some policy recommendations.

* We thank Sergio de Nardis, Enrico Giovannini, Andrea Ichino, Ugo Trivellato, Ignazio Visco, and participants to the ISAE Conference "Monitoring Italy II" (Rome, January 10, 2003) for helpful comments. Part of the research was carried out while Claudio Lupi was a Research Director at ISAE. The opinions expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of ISAE or its staff.

1 Introduction and motivations

The use of Gross Domestic Product (GDP) as a summary measure of the level of economic activity is pervasive in empirical economics, policy analysis and forecasting. We illustrate with a few examples.

In standard macroeconomic analysis, GDP is used to measure the dynamic behavior of an economy (through GDP growth), the output gap (the difference between actual and "potential" GDP), the turning points of the business cycle, etc. Per-capita GDP is often used to summarize the level of economic development of a country, and to compare it to others. Sometimes, cross-country comparisons of per-capita GDP levels have important political implications (as in the case of the Italian "surpass" of the UK in the mid-1980s).

Some GDP components carry information which is essential for the estimation of other macroeconomic quantities. For example, in Italy as well as other countries, the capital stock is constructed from the fixed investment series using the perpetual inventory method. Revisions in these data are therefore mirrored in the estimates of capital stock and total factor productivity (TFP). The extent to which these variables are influenced by revisions in the fixed investment series depends on the length of the sample that is subject to changes.

The use of GDP is not confined to macroeconomic analysis. For example, estimates of GDP levels are used by the European Union (EU) and other international organizations as the basis for levying contributions and distributing grants. The Stability and Growth Pact sets upper bounds on the ratio of European governments' budget deficit to GDP and to the ratio of public debt to GDP. Additional indicators involving GDP are increasingly required by the European Commission, the "cyclically adjusted" deficit being an example. Other important macroeconomic indicators involving GDP are the money supply to GDP ratio, the consumption to GDP ratio, the investment to GDP ratio, etc.

In finance, GDP growth is often utilized to predict asset prices or asset returns, along with other macroeconomic and financial variables. Recently, GDP growth also acquired an important role in the evolution of Italian public pension liabilities. In fact, the 1995 Social Security reform changed the Italian system of public pensions from defined-benefits to defined-contributions. In the new benefit formula, pension benefits at retirement are proportional to the capitalized value of lifetime contributions, with the proportionality factor depending on the age at retirement and with lifetime contributions capitalized using a 5-year moving average of past GDP growth rates.

The pervasive role of GDP and its nature of "public good" raise obvious problems of timeliness and accuracy of the data. There is a clear trade-off between these two goals (for an early discussion, see Stekler & Burch 1968), and the statistical agencies in the various countries differ depending on what point they select along this trade-off.[1]

GDP does not have a life of its own, but is the result of a system of conventions. The focus of this paper is not on GDP revisions that occur when this system of conventions is changed, for example after the introduction of a new system of national income and product accounts (as happened when the new European System of National Accounts, ESA 1995, substituted the previous one, ESA 1979). We focus instead on the revisions that occur, within a given definition of GDP, for a period of two-three years (sometimes more) after the first release of the data. This kind of revisions adds considerable uncertainty to any measure based on GDP and reveals the basic trade-off between timeliness and accuracy of the statistical information. However, in this paper we do not address the problem of the *accuracy* of national accounts estimates.[2] Rather, we focus on data revisions and their implications, taking for granted that revisions always increase data quality and reliability.

The paper is organized as follows. Section 2 defines GDP revisions and examines the reasons why they arise. Section 3 presents some evidence on the extent of GDP revisions in Italy, with particular emphasis on revisions of the quarterly national accounts series. Section 4 compares the Italian evidence with the available evidence from other countries. Section 5 discusses some areas in which data revisions can have potentially important consequences. Finally, Section 6 offers some policy recommendations.

2 GDP revisions

The measurement of GDP is a public good. Because producing GDP measurement would be too costly or even impossible for a single private firm, it is typically left to public statistical

[1] Sometimes data users may offer indications in this respect. In Italy, a recent survey highlighted that timeliness appears to be far more important than accuracy for the final data users (Drudi 2002).

[2] Interested readers may refer, among others, to Calzaroni and Puggioni (1998), Carson (2001), and Filippucci (2002).

agencies, often to a single one.

In Italy, national accounts are compiled by the National Accounts Department of the National Statistical Institute (Istat). Starting from 2002, the first estimates of the main quarterly national accounts items are released with a communication to the press around the 10th of March, June, September, and December, approximately 70 days after the end of the reference quarter.[3] The detailed set of quarterly accounts is usually disseminated some days later via floppy disk and the Web.[4] In the past, there used to be also a dedicated paper publication (*Conti Economici Nazionali Trimestrali*) that was lately suppressed mainly to avoid printing delays and extra costs. The main annual national accounts figures are released at the end of February/early March with a communication to the press. The complete set of annual accounts is disseminated later via floppy disk and the Web.[5]

A striking feature of GDP data (and, more generally, of all national accounts figures) is the presence of "data vintages". That is, the GDP estimate for a specific year or quarter is subject to several revisions after its first release. As a result, both the level and the profile of GDP over a given period may change, sometimes substantially, through time. For example, Figure 1 plots the quarterly and the annual GDP time series as estimated one year apart.

In this paper, we argue that perfect measurement of GDP is impossible and that measurement errors could be reduced as new information accrues, but cannot be completely eliminated. Data revisions are aimed at reducing the measurement errors in the initial estimates, and are therefore for the good of the users: series without revisions are not necessarily more accurate than series subject to periodic changes. Indeed, we believe that, from the user's standpoint, it might even be better to have data series embodying frequent small revisions rather than rare large ones. The issue is then what should be intended for "small". This, of course, partly depends on the particular uses made of the provisional estimates.

[3] A preliminary release of quarterly GDP at constant prices (estimated from the supply side) is disseminated about 45 days after the end of the reference quarter.

[4] The 2003 dissemination calendar confirm the same dates. Interested readers can download the Italian quarterly national accounts series from `http://con.istat.it` (free registration required) and from `http://www.istat.it/Economia/Conti-nazi/index.htm`.

[5] The set of annual national accounts (expressed in Euros) for the years 1970–2001 was released on June 12, 2002. See again `http://www.istat.it/Economia/Conti-nazi/index.htm`.

2.1 Sources of data revisions

Revisions in national accounts figures reflect the very nature of these data. While indicators stemming from sample surveys are subject to sample variability, but are otherwise the result of well-defined data collection processes, national accounts derive from many sources and virtually infinite data collection steps. The reason for this difference is that national accounts represent the functioning of the economy as a whole, and are therefore based on a large number of different, sparse, and often incomplete indicators. Information derives from a variety of sources and is seldom synchronized. According to the Istat, Italian annual national accounts are built upon data stemming from more than 80 different surveys, a number of administrative sources, and information coming directly from private organizations.[6]

In the current practice, there is a long delay in the collection and organization of this huge information set. For example, the final estimates of the survey on business firms are available to the Istat National Accounts Department only after about three years from the end of the reference year[7] (*e.g.*, the final estimates for the year 1999 were incorporated in the national accounts estimates for the first time in February 2003).[8] This means that the national accounts Department had to revise its annual and quarterly figures accordingly, changing the annual ones from 1999 onward, and the quarterly ones from 1997.[9] This delay is considered as acceptable by current Eurostat's regulations. Occasional changes to important surveys may also cause revisions in national accounts estimates. For example, this occurred just a few months after the major revision induced by the adoption of ESA 1995, when a change in the quarterly labor force survey forced the National Accounts Department to a new revision of its figures.

The timeliness of the quarterly figures is also influenced by the speed with which the information necessary to elaborate a first estimate becomes available. Since 1995, the first release of the quarterly national accounts has been anticipated by nearly one month, reducing the publication delay from about 95 to 70 days after the end of the reference quarter. This result is mainly due to a reduction of the time required to process a number

[6] Istat, Communication to the Press, March 8, 2001.
[7] See again Istat, Communication to the Press, March 8, 2001.
[8] The final estimates for the year 2000 were also made available in February 2003.
[9] See below for a brief description of Istat's official revision scheme.

of surveys conducted by various other Istat departments. According to Bruno *et al.* (2002), in 1995 about 50% of the information needed to estimate quarterly GDP from the supply side was available after 45 days from the end of the reference quarter, about 67% after 60 days, and about 70% after 75 days. We do not have enough elements to evaluate the current situation, but we conjecture an approximately proportional reduction of these delays, for given amounts of information, along with the abatement of the publication delay.

In any event, early national accounts vintages (both annual and quarterly) are based on smaller and less accurate information sets than later ones. As time passes, more information is collected and national accounts estimates are revised accordingly. On top of this, there are other potential sources of revisions, such as preliminary treatment of statistical indicators, seasonal adjustment, and the use of indirect estimation methods.

In our view, the main sources of errors, and successive data revisions, involve both annual and quarterly estimates and can be summarized as follows:[10]

1. Approximations to the conventional definition of GDP: GDP is defined using a complex set of conventions whose practical implementation is not always straightforward, so that approximations to the theoretical definitions are often needed (the measurement of underground economy is an obvious example).

2. Incomplete information: Even precisely defined quantities may be difficult to estimate when the necessary information is incomplete (for example, investment in software, intangible assets, imputed rents, own-production, etc.). More in general, basic data derived from very different sources must be adjusted and integrated, sometimes using *ad hoc* studies, to be reconciled with national accounts definitions and coverage.

3. Sampling errors due to the sampling nature of most elementary data: Most elementary micro data are derived from sample surveys and are therefore subject to sampling errors (*e.g.* employment rates, prices, etc.).

4. Nonsampling errors: Elementary items (*e.g.* industrial production, prices, employment) are themselves subject to nonsampling errors, such as measurement, reporting,

[10] The list is not necessarily exhaustive. A discussion from the producer's viewpoint can be found *e.g.* in Eurostat (2001).

coverage, or non-response errors.

5. Preliminary treatment of elementary items: Micro data and intermediate aggregates are subject to a number of preliminary statistical treatments such as outlier removal, seasonal adjustment, treatment of missing observations, adjustments for under-reporting of income and other variables, etc. All these steps require judgmental decisions, *ad hoc* solutions, and measurement problems.

6. Indirect estimation of quarterly variables: The use of model-based procedures to estimate quarterly national accounts implies the existence of a non-null variance of the estimated quantities and the need of periodic revisions of the published figures.[11] Furthermore, additional sources of revisions are related to changes in the indicators series and to the need of forecasting the indicators when they are not available at the time in which quarterly national accounts are estimated. Finally, when new or revised annual series are available, model-based procedures imply the re-specification and the re-estimation of the model equations, with revisions on the whole quarterly time series.

7. Adjustment of national accounts: The existence of measurement errors causes some accounting identities not to be satisfied. When this is the case, national accounts items must be adjusted in order to respect all the accounting constraints and identities. This is the very final step of the production process of the annual national account data. Though based on optimization criteria (Stone *et al.* 1942, Stone 1990), adjustment procedures nevertheless embody some *a priori* information on the part of the statistician and can introduce new measurement problems and data revisions. These are then mirrored in the quarterly national accounts series via the indirect estimation procedures.

8. Human errors: A huge number of different data sources and estimation methods are used to compile both annual and quarterly national accounts, and human errors (both conceptual and programming errors) cannot be excluded.

[11] A brief description of the methodology followed by Istat is offered in Appendix A.

Successive data revisions are also driven by feedback from users. Serious data users can provide a thorough check on data quality. This is simply because of the extensive and careful use they make of these data.

Given the public good nature of GDP measures, an adequate amount of resources should be devoted to reduce these errors, although they cannot be completely eliminated. Are the resources that Italy currently devotes to the production of national accounts "adequate"? Answering this question is not easy because of the lack of data. A recent document of Eurostat (presented to the CPS of 21/03/02) shows that, as of the end of year 2000, there are 42.3 employees of the National Statistical Institute per million inhabitants in Italy against a EU average (Italy excluded) of 102.9. The budget of the Istat represents 0.0128 percent of Italian GDP against a EU average (Italy excluded) of 0.025. According to these indicators, Italy would seem to devote to its national statistical institute about half the resources devoted on average by the other EU countries.

Of course, even if GDP revisions can be interpreted as "corrections" to initially "wrong" data, nevertheless, from the user's standpoint, they add further uncertainty to the basic one coming from the imperfect measurement of GDP. This gives rise to some important questions. How important is this extra uncertainty? Why does it arise? How should it be handled by the users? How could it be reduced? In the sequel of the paper, we will address some of these questions.

2.2 Notation and conventions

We now introduce some notation and conventions used throughout the paper. By "release" we mean the whole collection of national accounts time series. Time is measured in quarters, and we denote by release *yyyy*q*n* the national accounts time series whose most recent data point corresponds to year *yyyy*, quarter *n*. For example, release 1999q4 is made of all the national accounts time series ending in 1999q4.

Where not stated differently, Y_t denotes seasonally adjusted log GDP at time t, and $Y_{t,i}$ denotes the ith estimate (vintage) of Y_t. We assume that $Y_{t,i}$ is available after i quarters (at time $t+i$). This implies that $Y_{t,1}$, the first estimate of Y_t, is assumed to be available after one quarter (at time $t+1$).[12] We also assume that $Y_{t,d} = Y_t$, that is, Y_t is fully observed

[12] This approximation is very good for Italian national accounts data until 2002, when the delay has been

with a delay of $d \geq 1$ periods.

Under these two assumptions, the time-series released at time t (that is, release $t-1$) consists of

$$Y_{t-1,1}, Y_{t-2,2}, \ldots, Y_{t-d-1,d-1}, Y_{t-d}, Y_{t-d-1}, \ldots,$$

where $Y_{t-1,1}$ is the first estimate (vintage) of the most recent quarter, $Y_{t-2,2}$ is the second estimate (vintage) of the previous quarter, and so on, whereas the time-series released at time $t+1$ (that is, release t) consists of

$$Y_{t,1}, Y_{t-1,2}, \ldots, Y_{t-d,d-1}, Y_{t-d+1}, Y_{t-d}, \ldots.$$

The sequence $Y_{t-i,1}, \ldots, Y_{t-i,i+1}$ summarizes the history of the various estimates of Y_{t-i}, from the first vintage (available at time $t-i+1$) to the $(i+1)$th vintage (available at time $t+1$). In practice, convergence to the final data may take a long time, although later revisions are usually of lesser importance than earlier ones. The decay of the size of the revisions need not be monotonic, however. This is because quarter to quarter revisions are superimposed on the annual ones.

The various vintages are related in different ways, and a number of different comparisons may be of interest depending on the specific purposes of the analysis. Three main comparisons may be distinguished:

1. A comparison of $Y_{t,1}$ and $Y_{t-1,2}$ carries information on the dynamics of GDP within the same release, and is typically of interest to applied economists and policy makers.

2. A comparison of $Y_{t-1,2}$ and $Y_{t-1,1}$ carries information on the revision of the estimates of Y_{t-1} between successive releases, and is the key to the analysis of the revision process.

3. The difference $Y_{t,1} - Y_{t-1,1} = (Y_{t,1} - Y_{t-1,2}) + (Y_{t-1,2} - Y_{t-1,1})$ combines both the dynamics of GDP within the same release and the revision in the estimates of GDP between successive releases, and represents a problem of special concern for forecasters.

reduced to approximately 70 days after a quarter has ended.

For example, under the measurement model (8) in Appendix B.1, if $Y_{t,i}$ is unbiased for Y_t (that is, $Y_{t,i} = Y_t + U_{t,i}$, where $U_{t,i}$ is a zero-mean measurement error, independent of Y_t but possibly serially correlated), we have

$$Y_{t,1} - Y_{t-1,2} = X_t + U_{t,1} - U_{t-1,2},$$

$$Y_{t-1,2} - Y_{t-1,1} = U_{t-1,2} - U_{t-1,1},$$

and

$$Y_{t,1} - Y_{t-1,1} = X_t + U_{t,1} - U_{t-1,1} = X_t + (U_{t,1} - U_{t-1,2}) + (U_{t-1,2} - U_{t-1,1}),$$

where $X_t = Y_t - Y_{t-1}$ is the quarter-on-quarter GDP growth rate.

2.3 News and noise in GDP revisions

Mankiw, Runkle and Shapiro (1984) and Mankiw and Shapiro (1986) introduced the distinction between news and noise in GDP announcements (see also Maravall & Pierce 1986, and Sargent 1989). According to the noise view, preliminary estimates of GDP are error-corrupted measurements of GDP, where the "noise" $Y_{t,i} - Y_t$ $(1 \leq i < d)$ need not have zero mean. This view goes back to Howrey (1978) (see also Conrad & Corrado 1979, Harvey *et al.* 1981, De Jong 1987, Mork 1989, Bordignon & Trivellato 1989, Mariano & Tanizaki 1995, and Patterson 1995) and is the basis of the state-space approach outlined in Appendix B.1. Under the measurement equation (8), the "noise" is equal to

$$Y_{t,i} - Y_t = \alpha_i + (\beta_i - 1)Y_t + U_{t,i}. \tag{1}$$

Hence, the "noise" has zero mean only if $\alpha_i = 0$ and $\beta_i = 1$, that is, the preliminary estimate is unbiased. This pair of restrictions may be tested using the classical Mincer-Zarnowitz test of (weak) forecast efficiency (Mincer & Zarnowitz 1969).

According to the news view, the statistical agency produces the preliminary data $Y_{t,i}$ optimally using all the information available up to time $t + i$, including (possibly incorrect) models of the economy and the measurement error process. Thus

$$Y_t = Y_{t,i} + e_{t,i},$$

where $e_{t,i}$ is a "revision error" which reflects news that arrive between time $t+i$ and time $t+d$, that is, after the ith preliminary estimate has been compiled. Notice that, under the news view, the revision error $e_{t,i}$ is correlated with Y_t. If the models of the economy and the measurement error process are correct, then $e_{t,i}$ has the strong property of being uncorrelated with $Y_{t,i}$ and all the information available up to time $t+i$. Thus, a stronger test of forecast efficiency looks at evidence of predictability of future revision errors using current information, including the preliminary data currently available.

Faust, Rogers and Wright (2000) point out an important limitation of both types of test, namely the fact that data construction methods are constantly being revised, which means that past predictability need not be evidence of future predictability. In fact, revisions of the data construction methods tend to produce better estimates of GDP and GDP components, a finding already documented by Stekler (1967).

3 Extent and importance of GDP revisions in Italy

This section presents some evidence on the extent of GDP revisions in Italy, with particular emphasis on revisions of the quarterly national accounts series. This required the not trivial task of creating a "real-time data set" containing the national accounts data available to a policy-maker, forecaster, or academic researcher at various points in time.

3.1 Constructing a "real time data set" for Italy

The sample length of an internally consistent "real-time data set" for Italy is conditioned by the adoption, in 1999, of the new European System of National Accounts (ESA 1995).[13] The early releases of the quarterly national accounts consistent with the new system were characterized by short time series. Later, time series starting in 1982 were issued, but still following the old "cif-fob" scheme. The new accounts expressed according to the "fob-fob" framework were released by Istat with the estimate of the 4th quarter 1999. This is why our real-time data set starts from release 1999q4. Of course, more than the few consecutive releases that we have would be required for a detailed and accurate quantitative evaluation

[13] The first release of the national accounts annual estimates according to the updated ESA 1995 definitions was disseminated by Istat on April 30, 1999 (see Istat 1999).

of the revision process. However, by cutting the revisions sample to 1999q4, we ensure that the data set is homogeneous and that the estimates and their revisions are comparable.

¿From the practical point of view, construction of our "real-time data set" required collecting all the ESA 1995 releases as disseminated by Istat, checking their comparability, and keeping track of all the components of the resources and uses account (the GDP formation from the expenditure side) according to the fob-fob scheme. The resulting data set consists of 13 consecutive releases covering the period from 1999q4 to 2002q4. This allows us to compare the properties of the recent ESA 1995 data vintages with those of the ESA 1979 vintages analyzed by Busetti (2001).[14]

In order to give some flavor of the practical meaning of data revisions, in Figures 2 and 3 we plot the quarter on quarter percentage changes of GDP and its main components for the four quarters of 1999, as estimated in the 13 consecutive releases starting with 1999q4. The choice of 1999 is simply for illustration purposes and is motivated by the fact that 1999 is the most recent completed year in the first release we consider in this paper. The figures clearly show the nontrivial effect of revisions on the estimated quarterly growth rates of GDP and its components. To be more specific, the releases relative to the four quarters move in a range of about 0.2% for the first and third quarters, 0.3% for the second, and 0.6% for the last quarter. Furthermore, the ranges for the single GDP components can be much larger.

3.2 The revision scheme of Italian national accounts

Italian quarterly national accounts are revised every time a new vintage is released. The official revision scheme[15] is such that the quarters of the current year and those of the preceding two years are revised quarterly. When annual estimates are revised, the quarters of the two years preceding the oldest changed annual figure are also updated.[16]

Figure 4 plots the quarterly revisions of the quarter-on-quarter GDP growth rates for the 13 releases starting from 1999q4. Prior to the 2000q4 release, data series started in 1982q1.

[14] Busetti (2001) examines the last 12 vintages of the ESA 1979 version of the Italian quarterly national accounts. The comparison with York and Atkinson (1997) is not fully appropriate, since they use much longer revisions series. Di Fonzo, Pisani and Savio (2002) also contains an analysis of ESA 1979 Italian quarterly revisions.

[15] See *e.g.* the IMF Dissemination Standards at `http://dsbb.imf.org/country/ita/nagmeth.htm`.

[16] For example, if 1998 is the oldest changed annual figure, then 1996 and 1997 are also updated.

Release 2000q4 is the first vintage that includes data since 1970q1. In that occasion, annual data changed only from 1997 onward, but quarterly series were significantly revised for the whole period 1970–2000. This is because Italian quarterly national accounts are estimated using an indirect method which is a derivation of Chow and Lin (1971, 1976) (see Barbone, Bodo & Visco 1981, and Cainelli & Lupi 1999a, 1999b).

The method implies the specification and estimation of econometric equations on annual national accounts items and a set of related indicators (*e.g.* industrial production by sector, household consumption expenditures by category, etc.) at a very disaggregate level.[17] Prolonging the annual series backward implies the re-specification of these "measurement" equations, possibly using new and more accurate indicators. Furthermore, seasonal adjustment procedures give different results when applied to series starting in 1970 rather than 1982. In this sense, the revision implied in the 2000q4 release has an exceptional character, but shows the role of indirect methods in shaping the pattern of revisions to national accounts estimates.

Although this procedure is consistent with the prescriptions of the ESA 1995 (see Eurostat 1996, §12.04), a recent manual on the compilation of quarterly national accounts published by the IMF (Bloem, Dippelsman & Maehle 2001) suggests that quarterly national accounts should be estimated using

> timely and accurate source data that *directly* cover a high proportion of the totals. *Econometric methods* [...] *are not a substitute for data collection* (our emphasis).

3.3 Statistical properties of revision errors

We henceforth assume that the most recent vintage is the most accurate, being based on the largest information set, and compute the "revision errors" as the difference

$$e_{t|i,j} = Y_{t,j} - Y_{t,i}, \qquad 1 \leq i < j,$$

between the estimates of log GDP at time t as reported in vintage i and in the most recent estimate j. Three aspects of these revision errors are typically of interest (see, *e.g.*, Zellner

[17] See Appendix A for details.

1958). The first is the size of $e_{t|i,j}$ (the "error in the level"), for example the mean, standard error (sd) and mean absolute deviation (mad) over either i or t. The second is the size of the revision error on the quarterly growth rates, that is,

$$(Y_{t,j} - Y_{t-1,j+1}) - (Y_{t,i} - Y_{t-1,i+1}) = e_{t|i,j} - e_{t-1|i+1,j+1}$$

(the "error in the amount of change"). The third is the fraction of cases when $Y_{t,j} - Y_{t-1,j+1}$ and $Y_{t,i} - Y_{t-1,i+1}$ have opposite sign (the "error in the direction of change"). The last aspect may be especially important near turning points of the business cycle (Diebold & Rudebusch 1991b, and Dynan & Elmendorf 2001). All three aspects may be inversely related to the speed with which the data are reported, and may be affected by changes in the seasonal adjustment procedures.

Table 1 presents, for the quarterly growth rates of the main GDP components, the mean, standard deviation (sd), and mean absolute deviation (mad) of the revision errors relative to the most recent vintage, computed over the period from 1994q1 to the release date (the date at the top of each column). For example, the rows labelled mad report the mean absolute deviation $Q^{-1} \sum_{t=1994q1}^{1999q3+i} |Z_{t,j} - Z_{t,i}|$ computed for the Q quarters between 1994q1 and 1999q3+i, $1 \leq i \leq j$, where $j = 13$ denotes the "final" (most recent) estimate of the growth rate. Thus, for $i = 3$, the statistics are computed over the sample ranging from 1994q1 to 2000q2. In this case, the figures in the table represent the average "corrections" made to the growth rates, relative to the most recent vintage, over the whole period from 1994q1 to 2000q2. The table does not show any strong bias in the preliminary estimates (the mean revision error is nearly zero), confirming a similar finding reported by Busetti (2001) for the late ESA 1979 vintages.

A closer look to the table reveals that, while revisions on average tend to compensate, sometimes they can be sizeable, as highlighted by their standard errors and mean absolute deviations. In particular, the largest revisions involve fixed investments (above all transport equipments and metal products and machinery), imports and exports. Although consumption growth rates are less exposed to data revisions, small changes in this series can be important because national final consumption represents about 77% of GDP. However, the values reported in the table depend on the fact that averages are taken using a fixed origin (1994q1) up to the relevant release. In this way, there is a longer and longer fraction of the time series that becomes stably equal to the final release in the most recent estimates. When fixed 12-quarter wide windows are taken into account, the results change slightly, as reported in Table 2, and some systematic biases emerge. For example, the figures show that the investment series tends to be underestimated especially because of a negative bias in the estimate of the Transport equipment and Construction components. For the former, the quarter-on-quarter percentage growth rate is on average underestimated of about 0.7% (per quarter) over all the releases considered in the table.

A graphical summary of all the quarter-on-quarter revisions to the quarter-on-quarter growth rate of the main national accounts variables is presented in Figure 5. The figure plots the densities of the non-null quarter-on-quarter revisions on the sample 1982q1–2002q4 over all the 13 releases. Even considering this larger data set, the quarter-on-quarter revisions appear to be generally symmetric around zero. Again, it is confirmed that the most uncertain estimates are those related to the external trade and, to a lesser extent, to fixed investments.

3.4 Implications for short-term economic analysis

This section investigates how revisions may affect the analysis and interpretation of short-term economic movements. Indeed, from an analyst's viewpoint, it is important to understand to what extent revisions can lead one to modify his/her overall interpretation of the business cycle. In particular, is it possible that one goes completely wrong in the detection of the main items that contribute most to recent economic growth? We shall argue that this can happen, but is seldom the case in practice.[18]

Figures 6–9 show the revisions across releases of the contribution of the main GDP components to aggregate GDP growth. If the revision of the contribution of a specific component represents a significant fraction of GDP growth, then our interpretation of the economy based on the information set available in the previous quarter is also likely to require a revision. Figures 6 and 7 show that, except in rare cases, the estimates of the contribution of consumption and investments to GDP growth can be considered as relatively stable and reliable, at least from a quarter-on-quarter perspective. This means that revisions

[18] However, as shown below, there are other, more subtle ways in which revisions can adversely and significantly affect economic analyses and forecasts.

are unlikely to lead us to substantially revise our assessment of the role of internal final demand in shaping the dynamics of GDP. An important exception is the first quarter of 2001, when the contribution of national final consumption was initially estimated as being null, but was lately corrected into +0.36%, nearly half of GDP growth. Another exception is the second quarter of 2000, when the estimated contribution of final consumption more than doubled (to 0.64%) with respect to the previous vintages in correspondence with the dissemination of release 2000q4 (GDP growth was then estimated to be about 0.20%), while the contribution of inventories was significantly revised downward.

The matter is slightly different when we consider the contribution to growth of inventories and net exports (see Figures 8 and 9). Indeed, the estimates of the contribution of net exports are relatively volatile, reflecting the higher uncertainty associated with imports and exports estimates. Further, they often represent a significant fraction of GDP growth. The revisions to the contribution of inventories mirror those of net exports. Since inventories are estimated as the quantity that balances the resources and uses account, changes in this component might erroneously include significant bits of the external trade that have not been allocated in the different demand items.

In short-term economic analysis, when dealing with the data of the early quarters of the year, it is common practice to compute the average annual growth rate under the hypothesis that growth for the variable of interest will be zero in the remaining quarters. This forecast corresponds to the hypothesis that the variable of interest follows a pure random walk.[19] This quantity gives a rough idea of what the economic performance would be under a *status quo* condition. It is useful to compare it with formal annual forecasts in order to get an idea of the gap to be filled for the forecasts to be fulfilled.

How is this practice influenced by data revisions? Can these naïve forecasts be contradicted by later data changes? In Figures 10 and 11 we compute the implied annual growth rates for each quarter of 1999 for GDP and some of its main components when the naïve forecasts are applied to the different releases. Again, year 1999 is taken as an illustration, being the first completed year after the ESA 1995 revision.

On average, revisions influence estimated percentage growth of GDP by about two

[19] In Italy, such practice is called “calcolare l’acquisito”.

decimal points. This may be considered as an acceptable margin, given the uncertainty with which GDP itself is estimated. However, if one looks at the data in more detail, then large revisions can arise from one release to another, in particular for some investments series, imports, and households' consumption. This means that if we wanted to assess the dynamics of the main GDP components using this widespread technique, then we could easily go wrong, substantially missing the information carried by the latest releases. For example, according to a forecaster's viewpoint, the naïve forecast-based annual growth rate computed over the third quarter should be a rather solid benchmark for other predictions. Look at what happened to fixed investments in transport equipment between release 2001q3 and 2001q4: the estimated average annual growth rate estimated using the information embodied in the first three quarters of 2001 in the two different releases jumped from 7.3 to 15.8%. This revision is largely due to the revision of the first three quarters of 2001 in the two consecutive releases.[20] The actual annual growth rate, as of release 2002q2, is 16.8%, not too far from the naïve forecast based on release 2001q4, but more than double with respect to the corresponding forecast computed on release 2001q3.

Similar considerations hold also as far as investments in construction and imports series are considered. Somewhat disturbingly, sizeable and relatively stable components of GDP can also be subject to this kind of effects. Consider for example the naïve forecast-based annual growth rate of households' consumption computed from releases 2000q3 and 2000q4. In this case, the estimate increases from 1.6% to 2.1%. Given the weight and dynamics of this variable, a revision of the annual growth rate of +0.5 percentage points must be considered as large.

Of course, we know that these deviations are due to the updating of the annual figures, where better and more complete sources are used. However, with reference to the forecaster's viewpoint, the evidence suggests that anticipating the dynamics of GDP component may be a very difficult task. If naïve instruments like the one just described may fail badly, there is no special reason to believe that more sophisticated ones may do much better, unless a special treatment for possible data revisions is taken into account. Furthermore, the same evidence also suggests that the estimated short-term dynamics of some quarterly national

[20] Of course, there is also an effect due to the revision of the quarters of year 2000.

accounts variables may become really informative only *ex-post*, when a first estimate of the annual dynamics is available. Special caution should therefore be exerted when interpreting these data, possibly comparing them with alternative indicators, when these are available.

4 Comparisons with other countries

In this section we compare the Italian evidence with the available evidence from other countries.

4.1 Evidence from the USA

In this section we examine the revisions to US GDP as they stem from the real-time macro data set maintained by the Federal Reserve Bank of Philadelphia (see Croushore & Stark 1999, 2001) and freely downloadable from `http://www.phil.frb.org/econ/forecast/`.[21]

The choice of comparing Italy with the USA is motivated by the high reputation of the Federal agency that compiles the US national accounts (the Bureau of Economic Analysis at the Department of Commerce), and the fact that estimation procedures for US and Italian quarterly data represent, in a sense, two polar cases.[22]

First, unlike Italy, a significant fraction of the US quarterly national accounts items is estimated on the basis of a direct estimation method. Benchmarking interpolation techniques are also used to ensure consistency between the annual aggregates and the quarterly ones, primarily in correspondence with annual revisions.

Second, the US and Italian revision schemes are rather different. Exception made for the comprehensive "benchmark" revisions of the national income and product accounts (such as in the first quarter of 2000), in the USA only the estimates of the last quarter are revised quarterly, while in correspondence with the release of the second quarter of the year, 13 quarters back are changed (the first quarter of the current year plus the quarters of the preceding three years), in order to take into account the revisions of the previous years' annual data.

[21] To our knowledge, the only other source of real-time macro data is for the UK. The data are available from `http://www.econ.cam.ac.uk/dae/keepitreal` and data descriptions are provided in Egginton, Pick and Vahey (2002).

[22] See Seskin and Parker (1998) for a detailed description of the US national income and product accounts.

Third, quarterly national accounts are released in the USA no later than 31 days after the end of the reference quarter, while in Italy they are currently released about 70 days after the end of the reference quarter.[23]

Figure 12 plots the quarter-on-quarter GDP revisions to US GDP from 2000q1 to 2002q2. Revisions can occasionally be large, even larger than what we have seen for Italy. In spite of this, revisions to the contributions to GDP growth are negligible, and the interpretation of the main factors determining GDP growth from the demand side results extremely stable. For brevity, in Figure 13 we report only the revisions to the contributions of consumption to GDP growth, the other demand components showing a similar pattern.

Because of the trade-off between timeliness and accuracy, a fair comparison between data revisions in Italy and in the US should take into account the much earlier release of the US data. Further, average GDP growth was significantly more sustained in the USA than in Italy during the period considered.

The history of revisions to the estimates of 1999 GDP is reported in Figures 14 and 15. In particular, the estimates of the growth in 1999q4 have been revised in a range of about half a percentage point. Again, it should be pointed out that economic growth in the USA in that period has been much higher than in Italy.

Finally, the estimated densities of the quarter-on-quarter revisions (on the series 1982q1-2002q2) are reported in Figure 16 for comparison with the Italian case. Given the different revisions scheme, the US estimates are based on smaller samples. The estimated distributions are rather similar to the Italian ones, the only important difference being the smaller variance of the revisions of imports and exports in the USA.

4.2 Evidence from other countries

Cross-country comparisons are not easy to carry out after the adoption of the ESA 1995 by the European countries. However, some evidence based on the ESA 1979 national accounts can be gathered from previous studies.

York and Atkinson (1997) analyze the revisions process in quarterly national accounts of the G-7 countries in the period 1980–1993.[24] Their paper highlights that, in the period

[23] In both cases, we abstract from "flash" or "advance" estimates.
[24] Data start in 1987 for Italy and in 1982 for the UK.

considered there, there were no evidence of systematic bias in the preliminary national accounts figures, but the mean absolute revision was, for some country, a non trivial value. In terms of mean absolute GDP revision, Italy ranks second, just after France. However, if the average GDP growth is taken into account (by dividing the mean absolute GDP revision by the average growth), all the G-7 countries maintain approximately the same rankings, with the exception of Italy and Japan that switch their positions, Italy ranking fifth and Japan second. In two cases (Germany and the UK), the mean absolute revision is even larger than the average GDP growth.

Surprisingly enough, the highly reputed British statistical agency (the UK Central Statistical Office, at that time) ranks seventh in both the absolute and the relative ratings.[25] Where differences become relevant is in the timing of preliminary estimates dissemination. The authors report that, at the time their paper was written, the US statistical agency was the quickest in publishing the preliminary estimates just approximately four weeks after the quarter's end, the average delay being about 9 weeks. At that time, the Italian quarterly national accounts were released about fourteen weeks after the end of the reference quarter.[26]

These results are by and large confirmed by Faust, Rogers and Wright (2000), who also use data from OECD's *Main Economic Indicators*. Contrary to York and Atkinson (1997), this paper suggests that in some countries (notably Italy, Japan, and the UK) revisions are rather predictable. In particular, in these countries large (low) preliminary estimates are likely to be revised downward (upward), consistently with the hypothesis that initial estimates are significantly corrupted by noise, which is lately corrected when new information accrues.

This observation points again to the timeliness versus accuracy tradeoff. In this respect, it should be highlighted that the timing of the first estimate of the quarterly national accounts still varies substantially among countries. The US statistical agency is the quickest (within 31 days after the end of the reference quarter), followed by the Japanese (45 days), the Dutch (49 days), the British (56 days), and the German (58 days).[27] The "slowest"

[25] In this respect, the authors correctly remind that less revised data are not necessarily more accurate.

[26] We already pointed out that Istat has reached some significant results in reducing this delay, but timing differences should always be taken into account when performing reliability comparisons.

[27] These delays do not consider advance or flash estimates and represent the timeliness of the estimates as

agencies are the Swedish (90 days after the end of the reference quarter), the Belgian (98 days), the Irish and the Portuguese (about 120 days). Istat (Italy) and INSEE (France), with their 70-day delay, are in an intermediate position, the average delay computed for the EU15 countries, Canada, Japan and the USA being about 72 days.

5 Effects of GDP revisions

Three areas where data revisions can have important consequences are the definition of key empirical regularities, forecasting and policy analysis. The three areas are intimately connected with each other. For example, failure to properly identify the turning points of the business cycle may lead to policy decisions that appear incorrect when revised data become available (Dynan & Elmendorf 2001). On the other hand, if policy makers ignore erratic, short-term volatility in macroeconomic indicators, then revision errors of relatively large size may have small consequences (see, *e.g.*, Maravall & Pierce 1986, and Sargent 1989).

Existing evidence indicates that many empirical findings may depend on the particular data set employed. For example, Christoffersen, Ghysels, and Swanson (2000) show that the influence of economic news in financial markets can be evaluated correctly only using real-time data sets. Croushore and Stark (1999) highlight that some key macroeconomic empirical results (the so-called "stylized facts") depend crucially on the data vintage used. Testing Hall's (1978) life-cycle/permanent-income hypothesis is an example. Also Blanchard and Quah's (1989) permanent-transitory decomposition appears to be sensitive to the particular data release. This outcome seems to be even more general than the previous one, being apparently related to the sensitiveness of structural VARs to small changes in the data.

An important point that has been addressed only recently is that policy analysis using real-time data and current-vintage data can differ substantially. In order to study past economic policies and understand how and why policymakers reacted the way they did to particular situations, it is essential to consider the same information that the policymakers

reported by the IMF Data Dissemination Standards. The British statistical agency produces also a version of the quarterly national accounts whose delay is only 4 weeks after the reference quarter. However, this first vintage is more similar to a "flash" estimate than to a regular one.

had at the time they made their decisions. Using updated information would be unfair and unrealistic (see for example Runkle 1998).

In the remainder of this section, we give a flavor of the effects of GDP revisions on various aspects of the distribution of some GDP-related measures that are often of considerable public policy interest.

5.1 Basic descriptive statistics

Consider first the effects of GDP revisions on the mean and the variance of some GDP-related measures. Recall that the mean is the best unconditional predictor of a random variable under quadratic loss, and the variance is the risk associated with this predictor. The general message is that, even when they contain no systematic bias, preliminary data are more noisy than true data.

The first measure is simply the (log) level Y_t of GDP. The measurement equation (8) in Appendix B.1 implies that

$$\mathrm{E}\, Y_{t,i} = \alpha_i + \beta_i \,\mathrm{E}\, Y_t, \qquad \mathrm{Var}\, Y_{t,i} = \beta_i^2 \;\mathrm{Var}\, Y_t + \mathrm{Var}\, U_{t,i}.$$

Thus, unless $\alpha_i = 0$ and $\beta_i = 1$, preliminary data are subject to systematic bias. If $\mathrm{Var}\, U_{t,i} / \mathrm{Var}\, Y_t > 1 - \beta_i^2$ (for example, if $\beta_i \geq 1$), then they are also more noisy than the true data.

The second are ratios to GDP, such as the money stock to GDP ratio or the ratio of government budget deficit to GDP. After taking logs, these ratios are of the form $Z_t = W_t - Y_t$. Suppose for simplicity that the ith release of the data contains unbiased estimates $Y_{t,i}$ and $W_{t,i}$ of Y_t and W_t respectively, that is, $Y_{t,i} = Y_t + U_{t,i}$ and $W_{t,i} = W_t + V_{t,i}$, where $U_{t,i}$ and $V_{t,i}$ are zero-mean measurement errors, independent of Y_t and W_t, but possibly correlated.[28] In this case, the ith preliminary estimate $Z_{t,i} = W_{t,j} - Y_{t,i}$ is unbiased for Z_t, that is, $\mathrm{E}(Z_{t,i} - Z_t) = 0$. However,

$$\mathrm{Var}\, Z_{t,i} - \mathrm{Var}\, Z_t = \mathrm{Var}(V_{t,i} - U_{t,i}) \geq 0,$$

that is, the preliminary estimates of the ratio are more noisy than the actual ratio. Notice

[28] Positive correlation is especially likely if W_t is a sizeable component of GDP, such as private consumption.

that the excess variability of $Z_{t,i}$ is higher if $U_{t,i}$ and $V_{t,i}$ are negatively correlated than if they are positively correlated.

The third is the quarter-on-quarter GDP growth rate $X_t = Y_t - Y_{t-1}$. To analyze the effects of GDP revisions in this case, again suppose that preliminary estimates are unbiased, and consider the simple AR(1) plus noise model for GDP growth discussed in Appendix B.1. This model implies that

$$X_{t,i} = X_t + \epsilon_{t,i}, \tag{2}$$

where $X_{t,i} = Y_{t,i} - Y_{t-1,i+1}$ is the ith preliminary estimate at time $t+i$ of the GDP growth rate and $\epsilon_{t,i}$ is a white noise process, uncorrelated with X_t, with mean zero and finite variance.

It follows immediately from (2) that, although preliminary estimates of GDP growth are unbiased for X_t, that is, $\mathrm{E}(X_{t,i} - X_t) = 0$, we have

$$\mathrm{Var}\, X_{t,i} - \mathrm{Var}\, X_t = \mathrm{Var}\, \epsilon_{t,i} \geq 0,$$

that is, preliminary estimates of GDP growth are more noisy than actual GDP growth. Notice that the autocovariances of GDP growth are not affected if $\epsilon_{t,i}$ is a white-noise process, but not under more general processes.

5.2 Forecasting

We now consider the problem of forecasting future values of GDP growth $X_t = Y_t - Y_{t-1}$ on the basis of the information available up to time t.[29] This section summarizes the results presented in Appendix B.2.

At each point in time, several forecasts are available. A sub-optimal forecast ignores the preliminary data completely and uses the forecasting model and the finally revised observations. A naïve forecast ignores the distinction between preliminary and final data, and is typically constructed on the basis of the assumed model using the most recent preliminary observation of the unobserved true values.

Suppose for example that GDP growth (or, more precisely, the deviations of GDP growth from its mean) follows a zero-mean AR(1) process with known autoregressive parameter

[29] Given a forecast of X_{t+1}, a forecast of (log) GDP at time $t+1$ is easily obtained from the fact that $Y_{t+1} = Y_t + X_{t+1}$.

$|\phi| < 1$. If GDP were fully observable with no delay, then the best (i.e. minimum MSE) forecast of X_{t+j} available at time t would be

$$\hat{X}_{t+j|t} = \phi^j X_t.$$

If the finally revised data are released with a delay of $d \geq 1$ periods, the sub-optimal forecast that completely ignores the preliminary data is

$$\hat{X}^*_{t+j|t} = \phi^{j+d} X_{t-d}.$$

When ϕ is not too close to unity and d (the delay in the availability of the final data) is large, this forecast is essentially useless because indistinguishable from the unconditional mean.

Under the AR(1) plus noise model, the naïve j-step ahead forecast is

$$\tilde{X}_{t+j|t} = \phi^j X_{t,1}.$$

If $d = 1$, the best forecast is

$$\tilde{X}^*_{t+j|t} = \phi^j \left[\beta X_{t,1} + (1-\beta)\phi X_{t-1}\right],$$

with $0 \leq \beta \leq 1$ ($\beta = 1$ if GDP is fully observable with no delay).

The best forecast may be written

$$\tilde{X}^*_{t+j|t} = \phi^j \left[\beta X_{t,1} + (1-\beta)\hat{X}_{t|t-1}\right],$$

where $\hat{X}_{t|t-1} = \phi X_{t-1}$ is the one-step-ahead forecast from an AR(1) model. This expression suggests two alternative methods for improving upon the naïve forecast in more general models. The first method consists of modifying the forecast initial conditions. In the AR(1) plus noise model, this involves replacing the naïve forecast by $\phi^j X^*_t$, where

$$X^*_t = \beta X_{t,1} + (1-\beta)\hat{X}_{t|t-1}$$

is a convex combination of the preliminary estimate $X_{t,1}$ and the one-step-ahead forecast from the model. The second method consists of an intercept correction. In the AR(1) plus noise model, this involves replacing $\phi^j X_{t,1}$ by $\phi^j(X_{t,1} + c_t)$, where

$$c_t = X^*_t - X_{t,1} = -(1-\beta)(X_{t,1} - \hat{X}_{t|t-1}),$$

with $c_t = 0$ if $X_{t,1} = \hat{X}_{t|t-1}$.

Both methods are fully optimal for the AR(1) with noise model but need not be so for more general models. Busetti (2001) explores their properties using the Bank of Italy Quarterly Econometric Model.

5.3 Impulse response analysis

Sofar, we assumed complete knowledge of the data generating process. In fact, data revisions may have strong implications for the properties of estimated econometric models. Indeed, revisions can induce parameters changes or even different specifications of the model for the outcome variable of interest. However, when revisions are primarily concentrated in the most recent data, one should not expect the identified and estimated models to be too different.

We briefly investigate this issue using a simple univariate framework. We represent each GDP component (in log first differences) as

$$\varphi(L)\, X_t = \mu + wd_t + \vartheta(L) v_t, \tag{3}$$

where X_t denotes the first differences of the national accounts item under scrutiny, $\varphi(L)$ and $\vartheta(L)$ are polynomials in the lag operator L, wd_t represents a variable for (first differenced and seasonally adjusted) trading days and v_t is a white-noise process. The reason for including wd_t in the model is that Italian national accounts data are not adjusted for calendar effects, so that wd_t may have explanatory power that can be fruitfully exploited.

For each GDP component, we identify and estimate two separate models using a common sample (1982q1–1999q4) extracted from the 1999q4 and 2002q4 releases, respectively. Model identification is attained by selecting the model with the "best" BIC criterion and no residuals autocorrelation over all the possible combinations of AR and MA components up to order 5. Estimation is always performed by exact maximum likelihood.

In order to summarize the empirical results, we use the impulse response functions of the estimated models (Figures 17–19). We think that they are adequate to represent the behavior of dynamic forecasts derived from the estimated models. Though model parameters and parameterizations are sometimes different, the impulse response functions of the models estimated on the 1999q4 release are generally not dramatically different from those derived from the models estimated on the most recent release (2002q4). Significant differences arise only with respect to fixed investments (particularly transport equipment and construction) and stock changes. Also total resources show rather different impulse response functions in the two cases, but it is seldom the case that this variable is forecasted directly rather than being derived from the forecasts of its main components.

With the exceptions outlined above, the impulse response functions estimated on the preliminary and the “final” releases do not differ much. Hence, it appears that the most relevant differences in terms of forecasting are confined to those related to the change of the forecast origin. A proper use of intercept corrections might therefore be advisable.

6 Concluding remarks

In this paper we study why national accounts data revisions arise, offer some assessments of their extent in Italy as compared to the United States and other industrialized countries, and indicate their implications for data users. We do not address the issue of the *accuracy* of the national accounts estimates. In fact, revisions *per se* are not a measure of accuracy.

The focus of the paper is on quarterly data, although revisions stem from modifications in both the annual and the quarterly information set. In those countries, like Italy, where quarterly national accounts items are estimated using indirect methods, a change in the annual data affect the quarterly profiles in three ways. First, via the direct effect of the change in the annual average. Second, through the re-estimation of the measurement equations that are at the core of the econometric time disaggregation methods. Third, via the “propagation” autoregressive mechanism that characterizes many time series disaggregation procedures.

Even at the quarterly frequency, one of the main sources of revision is the arrival of new and more complete information. This is true also when indirect methods are used, especially when forecasted values for the main indicators are substituted by the actual ones as soon as they become available.

Quarterly series are also subject to quarter to quarter changes for the effect of seasonal adjustment procedures that typically give slightly different results when applied to time series of different length. In this respect, it should be recognized that some research has been carried out at Istat to minimize the extent of this component of the overall revision (see Di Palma & Savio 2000).

We focus on the 13 most recent quarterly Italian national accounts data vintages, evaluated according to the ESA 1995 accounting scheme. The extent of data revisions is comparable to that experienced by the US national accounts. However, it should be stressed that the publication delay of US data is much shorter and that economic growth in Italy has been considerably lower than in the US in the period covered by our sample. Furthermore, while in Italy there are a few instances in which the interpretation of the main causes of GDP growth have been changed by later data, in the US, the contribution of the main demand components to GDP growth is more stable, and no issue of this kind arise in the period considered in our study. Evidence gathered from previous papers (referring to the old ESA 1979 accounting scheme) suggests that Italy is in a mid-ranking position among the most industrialized countries, as far as national accounts data revisions are concerned.

We investigate the theoretical and empirical implications of data revisions for forecasting. On the empirical side, we show that some commonly used naïve forecasting techniques may be misleading in the presence of data revisions of the size experienced in Italy. On the other hand, the impulse response functions of simple univariate models estimated on different data releases appear to be rather similar, except for some specific series. This finding suggests that the most serious problem for forecasting is related to the revisions of the forecast origin, rather than to the dynamics implicit in the forecasting model.

On the theoretical side, it is even difficult to think of an uncontroversial way of comparing the quality of alternative forecasts, when these are corrupted by data revisions. This is because the dynamics implicit in the first estimate of a quarter with respect to the forecast origin combine both the dynamics of the variable within the same release and the revision between successive releases. Which series should be taken as the "true" one? The most recent? The one that includes the first estimate of forecast data point? Both choice could be criticized, but they imply different conclusions in terms of forecasting performance.

Data revisions do adversely affect the quality of forecasts. However, how large is this influence is an empirical matter that depends on the particular series being forecast. Not using at all preliminary estimates is not a viable solution. We show that forecasts combination and intercept corrections may offer better answers. We have analyzed this issue only in a simplified univariate framework. The negative influence of data revisions is likely to be more pronounced in large econometric models. This would suggest using simple forecasting tools to obtain more reliable forecasts. At the moment, however, these are simply conjectures and open research issues.

We conclude with some policy recommendations, distinguishing between those for data producers and those for data users. Our first set of issues regards how to monitor and regulate the quality of the production process of national accounts data.

In science, a necessary condition for quality is replicability. For example, scientific journals typically require data and programs to be made publicly available. Replicability would require availability of the raw data and a complete documentation of the procedures being used.

Although the quality of the available documentation has certainly improved, complete replicability is likely to be hard or even impossible, given the complexity of the estimation procedures of national accounts data. Microdata come from different sources, often there are privacy issues; procedures are highly standardized but sometimes unexpected problems may arise with respect to which ad hoc decisions have to be made; the exact timing with respect to which national accountants receive the necessary information is often unpredictable and not replicable, with serious consequences on the replicability of the estimates.

If we accept the view that there is a limit to complete replicability of national accounts data, how big should this limit be? How far are we from attaining the limit? Large deviations are likely to signal inefficiencies in data production. How can these inefficiencies be reduced? Given the complexity of the data production process, it is extremely difficult to address these issues. Moreover, they concern not only Istat, but also Ministries, public and private offices and enterprizes, and even international organizations. In this respect, it should be acknowledged that the European Statistical Agencies have recently become well aware of the problem, suggesting that

> The need for standardisation of methodologies used in the estimate of aggregates required by the accounting model of the NA and the need for an Integrated Information System are determined not only by the will to use the highest

> informative power of existing sources, but also by the need for "repeatable" estimates. Above conditions are deemed necessary so that the estimates could be considered part of a methodological model which is clear to users and able to measure statistic effectiveness. (Calzaroni & Puggioni 1998.)

There is no doubt that some structural problems, namely the weight of small firms in the economy, the size of the underground economy, and the extent of regional inequality, make GDP measurement particularly difficult in Italy. At the same time, the informational problems that affect our understanding of the economic issues involved, make it harder to find appropriate policies and reasonable solutions.

Turning to the specific question about how the data production process could be improved, three actors may play an important role, besides Istat itself: the Commission for the Protection of Statistical Information (Commissione per la Garanzia dell'Informazione Statistica or CGIS), international organizations (Eurostat, OECD, IMF, ECB, United Nations, etc.), and the scientific community. The role of the scientific community is, in our view, especially important. It is often the case that not all the national accounts items can be scrutinized with the same accuracy during the production process. Serious data users can provide a thorough check on data quality. This is simply because of the extensive and careful use they make of these data. If only for this reason, feedback from the scientific community should be appreciated and valued by statistical agencies.

The last remark, takes us to our second set of issues, those aimed at the data users. As far as policy rules are concerned, the available evidence suggests "prudence" and the use of rules rather than "activist" policies. For example, Swanson, Ghysels and Callan (1998) show that if the Fed used a Taylor-type rule and based policy decisions on changes in the index of industrial production, then policy would have improved significantly if policy makers had waited for data to be revised, rather than react to newly released data. This is particularly the case when data are contaminated by noise.

Last but not least, when analyzing economic performance it is advisable not to rely on a single indicator. Even if the focus is on real economic activity, it is advisable to have alternative measures, possibly related to synthetic leading and coincident indicators, rather than just quarterly GDP. However, bear in mind that also the forecasting performance of leading indicators worsens dramatically when real-time data are considered, independently of the levels or the turning points being the quantities to be forecast (see *e.g.* Diebold & Rudebusch 1991a, 1991b).

References

Barbone L., Bodo G., and Visco I. (1981), "Costi e profitti nell'industria in senso stretto: un'analisi di serie trimestrali", *Bollettino della Banca d'Italia*, 36, 465–510.

Blanchard O.J., and Quah D. (1989), "The dynamic effects of aggregate demand and supply disturbances", *American Economic Review*, 79, 655–673.

Bloem A.M., Dippelsman R.J., and Maehle N.O. (2001), *Quarterly National Accounts Manual: Concepts, Data Sources, and Compilation*. Washington DC: International Monetary Fund.

Bordignon S., and Trivellato U. (1989), "The optimal use of provisional data in forecasting with dynamic models", *Journal of Business and Economic Statistics*, 7, 275–286.

Borgioli S. (1996), "Il bilanciamento dei conti nazionali: la metodologia Stone e le sue estensioni", *Quaderni di Ricerca Istat*, No. 3/1996, 139–174.

Bruno G., Cubadda G., Giovannini E., and Lupi C. (2002), "The flash estimate of the Italian Gross Domestic Product", in Eurostat, *Workshop on Quarterly National Accounts*, 225–235.

Busetti F. (2001), "The use of preliminary data in econometric forecasting: an application with the Bank of Italy Quarterly Model", Banca d'Italia, Temi di Discussione No. 437.

Cainelli G., and Lupi C. (1999a), "Aggregazione contemporanea e specificazione econometrica nella stima trimestrale dei conti economici nazionali", *Statistica*, 59, 239–267.

Cainelli G., and Lupi C. (1999b), "The choice of the aggregation level in the estimation of quarterly national accounts", *Review of Income and Wealth*, 45, 483–492.

Calzaroni M., and Puggioni A. (1998), "Evaluation and analysis of the quality of the national accounts aggregates", Document CPBN 250 for the 31st meeting of the GNP Committee.

Carson C.S. (2001), "Towards a framework for assessing data quality", International Monetary Fund, Working Paper WP/01/25.

Chow G., and Lin A.L. (1971), "Best unbiased interpolation, distribution and extrapolation of time series by related series", *Review of Economics and Statistics*, 53, 372–375.

Chow G., and Lin A.L. (1976), "Best linear unbiased estimation of missing observations in an economic time series", *Journal of the American Statistical Association*, 71, 719–721.

Christoffersen P., Ghysels E., and Swanson N.R. (2000), "Let's get 'real' about using economic data", mimeo

Conrad W., and Corrado C. (1979), "Application of the Kalman filter to revisions in monthly retail sales estimates", *Journal of Economic Dynamics and Control*, 1, 177–198.

Croushore D., and Stark T. (1999), "A real-time data set for macroeconomists", Federal Reserve Bank of Philadelphia, Working Paper No. 99-4.

Croushore D., and Stark T. (2001), "A real-time data set for macroeconomists", *Journal of Econometrics*, 105, 111–130.

De Jong P. (1987), "Rational economic data revisions", *Journal of Business and Economic Statistics*, 5, 539–548.

Di Fonzo T., Pisani S., and Savio G. (2002), "Revisions to Italian quarterly national accounts aggregates: some empirical results", in Eurostat, *Workshop on Quarterly National Accounts*, 7–31.

Diebold F.X., and Rudebusch G.D. (1991a), "Forecasting output with the composite leading index: A real-time analysis", *Journal of the American Statistical Association*, 86, 603–610.

Diebold F.X., and Rudebusch G.D. (1991b), "Turning point prediction with the composite leading index: An Ex Ante Analysis", in K.Lahiri and G.H.Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records.* Cambridge: Cambridge University Press.

Di Palma F., and Savio G. (2000), "Temporal aggregation, seasonal adjustment and data revisions", Istat, mimeo.

Drudi I. (2002), "Indagine sulla percezione della qualità degli utenti di dati statistici", in C. Filippucci (ed.), *Strategie e modelli per il controllo della qualità dei dati.* Milano: Franco Angeli.

Dynan K.E., and Elmendorf D.W. (2001), "Do provisional estimates of output miss turning points?", Board of Governors of the Federal Reserve System, Working Paper.

Egginton D.M., Pick A., and Vahey S.P. (2002), " 'Keep it real!': A real-time UK macro data set", *Economics Letters*, 77, 15–20.

Eurostat (1996), *European System of Accounts ESA 1995.* Luxembourg: Eurostat.

Eurostat (2001), "Task force on accuracy assessment of national accounts statistics: Final report", Luxembourg: Eurostat.

Faust J., Rogers J.H., and Wright J. (2000), "News and noise in G-7 GDP announcements", Board of Governors of the Federal Reserve System, International Finance Discussion Paper No 690.

Filippucci C. (Ed.) (2002), *Strategie e Modelli per il Controllo della Qualità dei Dati.* Milano: Franco Angeli.

Gallo G.M., and Marcellino M. (1999), "Ex post and ex ante analysis of provisional data", *Journal of Forecasting*, 18, 421–433.

Hall R.E. (1978), "Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence", *Journal of Political Economy*, 86, 971–987.

Harvey A.C., McKenzie C.R., Blake D.P.C., and Desai M.J. (1981), "Irregular data revisions", in A. Zellner (ed.), *Applied Time Series Analysis of Economic Data*, Washington, DC: US Department of Commerce, Economic Research Report ER-5.

Howrey E.P. (1978), "The use of preliminary data in econometric forecasting", *Review of Economics and Statistics*, 60, 193–200.

Istat (1999), "Revisione dei conti nazionali e adozione del SEC95", *Note Rapide*, 4, No. 4, 1-8.

Mankiw N.G., Runkle D.E., and M.D. Shapiro (1984), "Are preliminary announcements of the money stock rational forecasts?", *Journal of Monetary Economics*, 14, 15–27.

Mankiw N.G., and M.D. Shapiro (1986), "News and noise: An analysis of GNP revisions", *Survey of Current Business*, 66, 20–25.

Maravall A., and Pierce D.A. (1986), "The transmission of data noise into policy noise in U.S. monetary control", *Econometrica*, 54, 961–979.

Mariano R.S., and Tanizaki H. (1995), “Prediction of final data with use of preliminary and/or revised data”, *Journal of Forecasting*, 14, 351–380.

Mincer J., and Zarnowitz V. (1969), “The evaluation of economic forecasts”, in J. Mincer (ed.), *Economic Forecasts and Expectations*, New York: NBER.

Mork, K.A. (1987), “Ain't behavin': Forecast errors and measurement errors in early GNP estimates”, *Journal of Business and Economic Statistics*, 5, 165–175.

Patterson K.D. (1995), “An integrated model of the data measurement and data generation process with an application to consumers' expenditure”, *Economic Journal*, 105, 54–76.

Runkle D.E. (1998), “Revisionist history: How data revisions distort economic policy research”, *Federal Reserve Bank of Minneapolis Quarterly Review*, 22, 3–12.

Sargent T.J. (1989), “Two models of measurements and the investment accelerator”, *Journal of Political Economy*, 97, 251–287.

Seskin E.P., and Parker R.P. (1998), “A guide to NIPA”, *Survey of Current Business*, 78, 26–68.

Seskin E.P., and Parker R.P. (1999), “Annual revision of the national income and product account”, *Survey of Current Business*, 79, 6–33.

Stekler H.O. (1967), “Data revisions and economic forecasting”, *Journal of the American Statistical Association*, 62, 470–483.

Stekler H.O., and Burch S.W. (1968), “Selected economic data – Accuracy vs. reporting speed”, *Journal of the American Statistical Association*, 63, 436–444.

Stone R. (1990), ”Adjusting the national accounts”, *Annali di Statistica*, 9, 311–325.

Stone R., Champernowne D.G., and Meade J.E. (1942), ”The precision of national income estimates”, *Review of Economic Studies*, 9, 111–125.

Swanson N.R., Ghysels E., and Callan M. (1999), “A multivariate time series analysis of the data revision process for industrial production and the composite leading indicator”, in R.F. Engle and H. White (eds.), *Cointegration, Causality and Forecasting.* Oxford: Oxford University Press.

York R., and Atkinson P. (1997), “The reliability of quarterly national accounts in seven major countries: a user's perspective”, OECD, Working Paper No. 171.

Zellner A. (1958), “A statistical analysis of provisional estimates of gross national product and its components, of selected national income components, and of personal savings”, *Journal of the American Statistical Association*, 53, 54–65.

Table 1: Summary statistics of revisions errors of GDP components percentage growth rates computed with respect to the last (13th) available vintage. Statistics are calculated over the sample 1994q1 $\leq t \leq$ date of release. The figures in the table represent the average "corrections" to the growth rates over the whole time series starting from the first quarter 1994.

		1999q4	2000q1	2000q2	2000q3	2000q4	2001q1	2001q2	2001q3	2001q4
	mean	0.04	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01
GDP	sd	0.23	0.21	0.22	0.21	0.11	0.11	0.12	0.12	0.08
	mad	0.20	0.19	0.19	0.19	0.06	0.06	0.06	0.06	0.06
	mean	0.14	0.26	0.09	0.03	0.00	0.01	0.02	0.01	0.01
Imports	sd	1.29	1.14	1.90	1.69	1.40	1.19	1.40	1.34	0.68
	mad	1.37	1.34	1.53	1.50	0.62	0.61	0.58	0.61	0.59
	mean	0.06	0.08	0.04	0.03	0.01	0.01	0.01	0.01	0.01
Total resources	sd	0.35	0.30	0.46	0.40	0.29	0.27	0.33	0.31	0.17
	mad	0.28	0.36	0.41	0.37	0.15	0.11	0.09	0.13	0.12
	mean	0.09	0.06	0.07	0.08	0.02	0.03	0.02	0.02	0.01
National consumption	sd	0.32	0.30	0.31	0.29	0.16	0.18	0.17	0.17	0.11
	mad	0.27	0.28	0.30	0.29	0.08	0.15	0.11	0.12	0.06
	mean	0.11	0.09	0.09	0.10	0.02	0.02	0.00	0.00	0.00
Households' consumption	sd	0.39	0.37	0.38	0.35	0.17	0.17	0.18	0.18	0.15
	mad	0.34	0.44	0.42	0.40	0.11	0.12	0.12	0.17	0.11
	mean	-0.01	-0.02	-0.01	0.00	0.01	0.05	0.08	0.08	0.05
Public consumption	sd	0.23	0.23	0.24	0.25	0.21	0.31	0.34	0.32	0.18
	mad	0.22	0.21	0.21	0.21	0.12	0.11	0.15	0.14	0.06
	mean	0.05	0.09	0.04	0.07	0.03	0.07	0.10	0.10	0.00
Fixed investments	sd	0.91	0.89	0.83	0.83	0.45	0.47	0.47	0.46	0.40
	mad	0.54	0.69	0.37	0.54	0.36	0.43	0.46	0.36	0.22
	mean	0.03	-0.02	-0.04	-0.02	-0.03	-0.04	-0.02	-0.01	0.01
Machinery	sd	1.64	1.55	1.51	1.50	0.64	0.54	0.55	0.54	0.45
	mad	1.16	1.01	1.05	1.01	0.45	0.23	0.39	0.39	0.24
	mean	0.16	0.18	0.19	0.20	0.14	0.45	0.52	0.56	0.05
Transport equipments	sd	3.07	3.00	2.93	2.75	1.38	2.05	2.11	2.15	1.45
	mad	2.29	2.15	2.03	1.77	0.76	0.95	1.00	0.58	0.46
	mean	0.03	0.15	0.05	0.12	0.08	0.09	0.11	0.08	-0.03
Construction	sd	0.70	0.83	0.72	0.77	0.74	0.74	0.74	0.71	0.62
	mad	0.56	0.58	0.72	0.65	0.37	0.40	0.40	0.35	0.21
	mean	0.11	0.19	0.10	0.06	-0.06	-0.01	-0.06	-0.02	-0.03
Exports	sd	1.15	1.25	1.50	1.33	1.50	1.47	1.61	1.53	0.79
	mad	0.85	1.08	1.17	0.97	0.37	0.35	0.35	0.35	0.22

Table 2: Summary statistics of revisions errors of GDP components percentage growth rates computed with respect to the last (13th) available vintage. Statistics are calculated. Statistics are computed over 12-quarter moving samples.

		1999q4	2000q1	2000q2	2000q3	2000q4	2001q1	2001q2	2001q3	2001q4
	mean	0.06	0.08	0.05	0.07	0.02	0.03	0.02	0.02	0.03
GDP	sd	0.21	0.25	0.26	0.26	0.12	0.10	0.15	0.14	0.11
	mad	0.22	0.20	0.18	0.23	0.10	0.10	0.19	0.11	0.10
	mean	0.06	0.38	0.58	0.11	-0.12	0.05	0.11	0.18	-0.05
Imports	sd	1.62	1.05	2.10	2.29	2.08	1.81	2.18	2.07	0.98
	mad	1.38	0.91	2.06	2.61	1.24	1.27	1.21	1.23	1.22
	mean	0.06	0.14	0.17	0.08	0.00	0.04	0.04	0.05	0.01
Total resources	sd	0.40	0.35	0.57	0.55	0.40	0.37	0.49	0.46	0.25
	mad	0.43	0.46	0.71	0.62	0.22	0.22	0.15	0.25	0.33
	mean	0.14	0.20	0.13	0.17	0.00	0.02	0.06	0.06	0.03
National consumption	sd	0.35	0.30	0.31	0.21	0.13	0.17	0.18	0.19	0.15
	mad	0.38	0.31	0.32	0.27	0.08	0.16	0.19	0.16	0.10
	mean	0.18	0.26	0.17	0.22	0.01	0.03	0.04	0.01	0.01
Households' consumption	sd	0.42	0.38	0.41	0.29	0.15	0.19	0.18	0.22	0.22
	mad	0.57	0.42	0.50	0.30	0.11	0.22	0.15	0.22	0.21
	mean	0.01	0.02	0.02	0.02	-0.02	0.02	0.15	0.20	0.09
Public consumption	sd	0.28	0.25	0.25	0.26	0.20	0.24	0.44	0.43	0.24
	mad	0.29	0.24	0.22	0.19	0.24	0.24	0.30	0.28	0.21
	mean	0.10	0.16	0.19	0.15	0.11	0.07	0.19	0.31	0.04
Fixed investments	sd	0.79	0.76	0.66	0.66	0.39	0.40	0.44	0.43	0.52
	mad	0.54	0.72	0.48	0.34	0.51	0.44	0.28	0.41	0.62
	mean	0.11	0.12	0.00	-0.02	-0.01	-0.05	-0.09	0.07	0.08
Machinery	sd	1.11	0.99	0.98	1.02	0.66	0.46	0.41	0.47	0.70
	mad	1.25	0.94	0.99	1.08	1.12	0.51	0.58	0.39	0.62
	mean	0.04	0.56	0.78	0.94	0.49	0.43	1.36	1.57	0.04
Transport equipments	sd	3.24	2.50	2.38	1.94	1.56	1.68	2.95	3.02	2.34
	mad	2.09	1.68	1.57	1.07	1.26	1.20	1.33	1.22	1.51
	mean	0.08	0.09	0.24	0.14	0.15	0.11	0.17	0.23	-0.02
Construction	sd	0.92	0.90	0.75	0.70	0.61	0.60	0.61	0.61	0.28
	mad	0.48	0.59	0.53	0.90	0.40	0.44	0.51	0.63	0.25
	mean	-0.05	0.22	0.56	0.03	-0.09	-0.14	0.02	-0.06	-0.07
Exports	sd	1.18	1.49	1.67	1.59	2.32	2.29	2.57	2.51	1.30
	mad	0.86	1.32	1.95	1.24	0.79	0.93	0.77	0.81	0.64

Figure 1: Annual and quarterly time-series revisions: Release 1999q4 (broken line) and release 200q4 (solid line). The annual figures (horizontal lines) are represented by quarterly average.

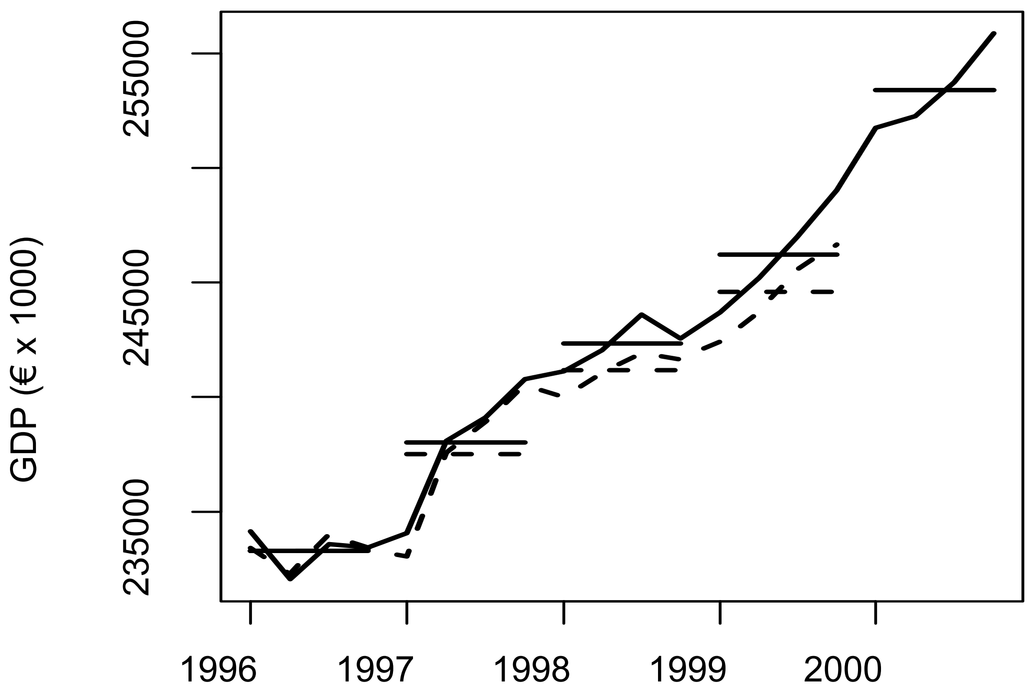


Figure 2: Estimates of the four 1999 quarters of Italian GDP in the last 13 revisions (quarter-on-quarter percentage changes): solid line, first quarter; long dashed, second quarter; dashed, third quarter; dot-dashed, fourth quarter.

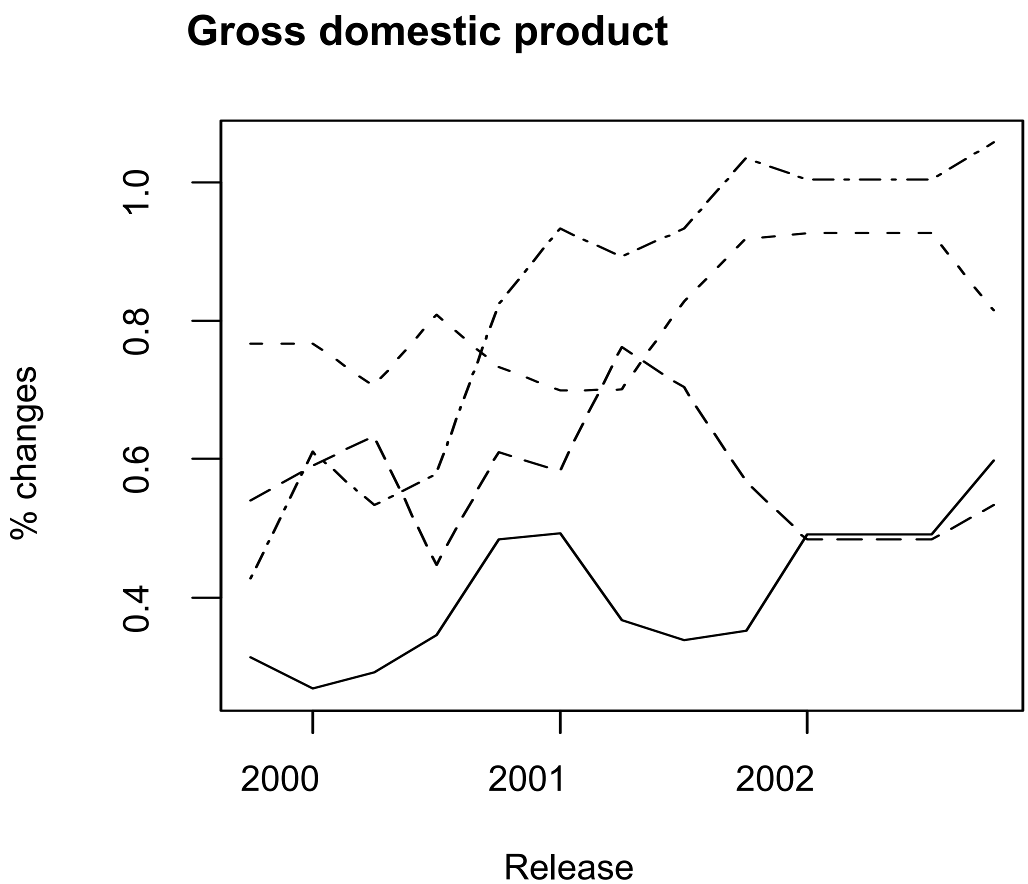


Figure 3: Estimates of the four 1999 quarters of Italian GDP main components in the last 13 revisions (quarter-on-quarter % changes): solid line, first quarter; long dashed, second quarter; dashed, third quarter; dot-dashed, fourth quarter.

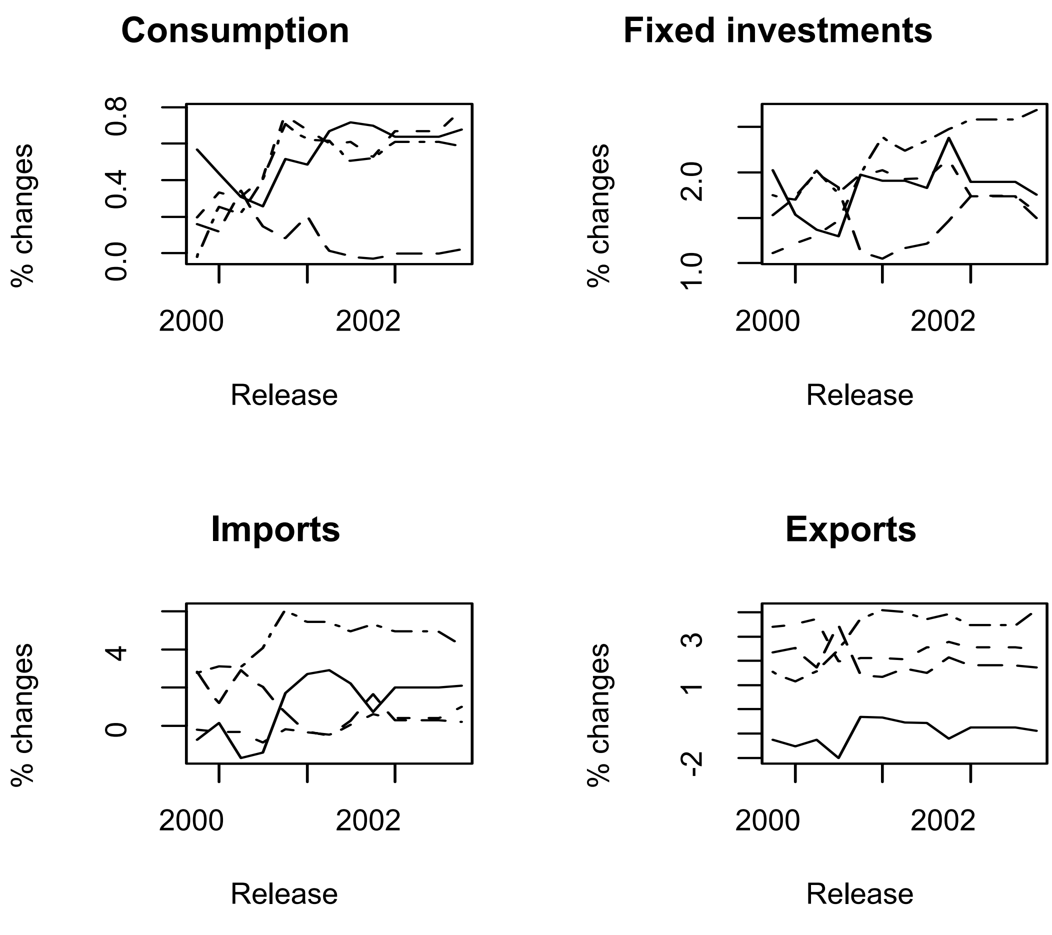


Figure 4: Quarter-on-quarter revisions of Italian GDP (% changes).

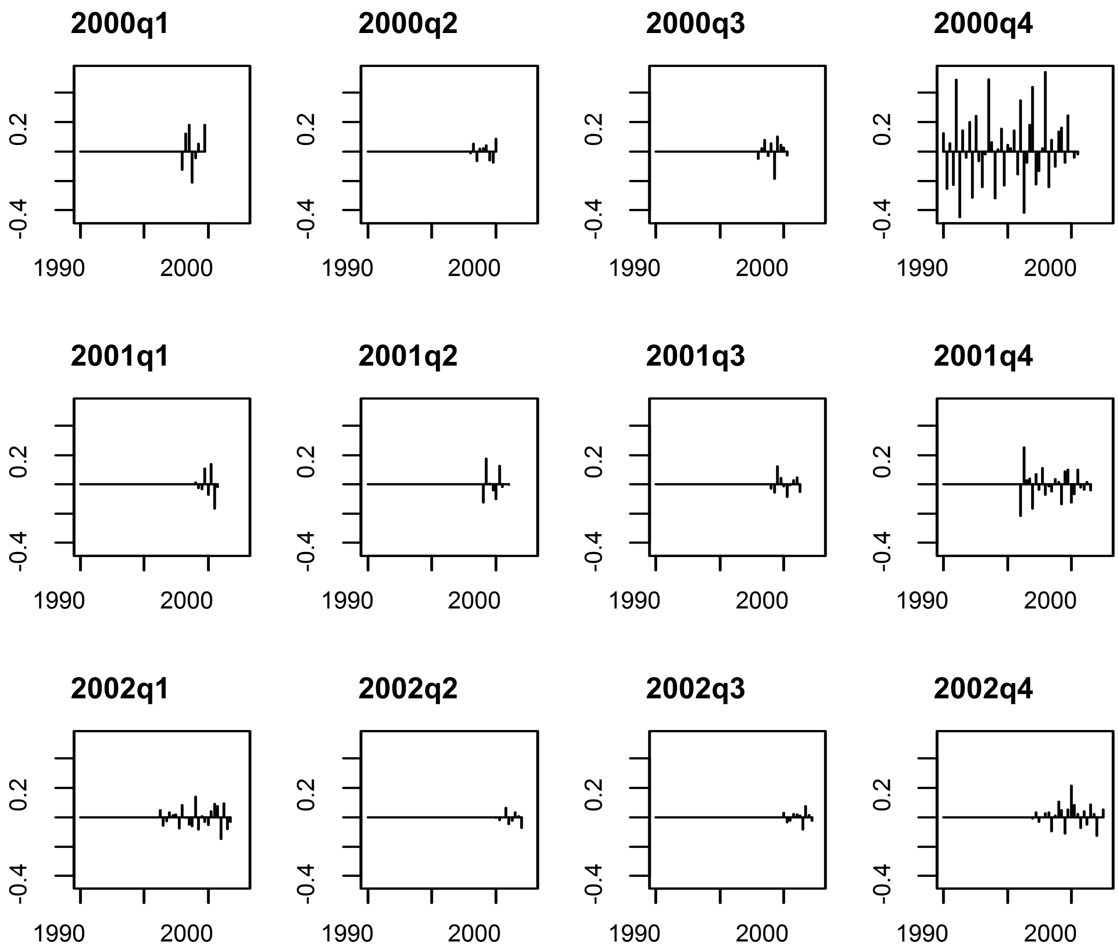


Figure 5: Estimated kernel densities of the quarter-on-quarter revisions in the quarter-on-quarter % growth rates: Gaussian kernel with biased cross-validation bandwidth selection. Normal densities with same mean and variance of the observed data are reported for comparison (dashed lines).

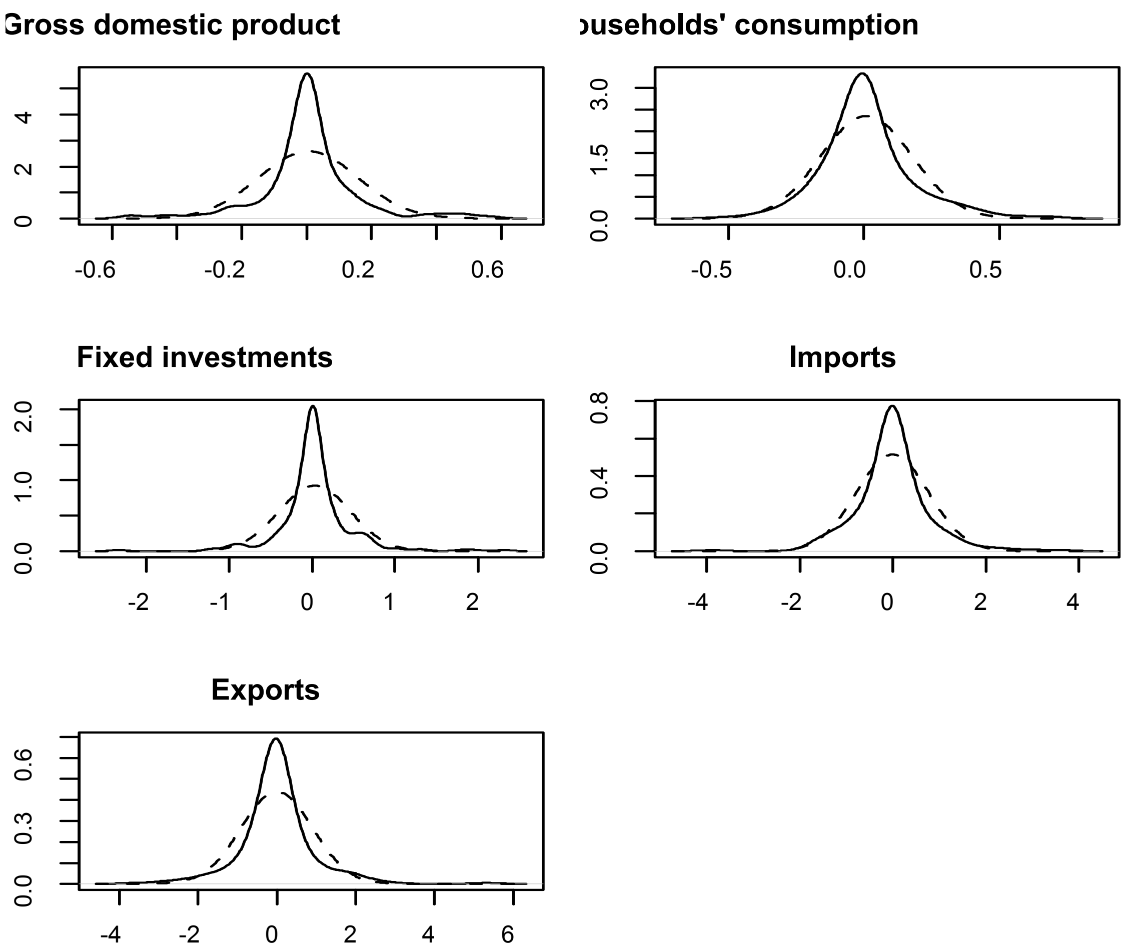


Figure 6: Quarter-on-quarter Italian GDP growth (solid line), and quarter-on-quarter revisions of contribution to GDP growth: Consumption.

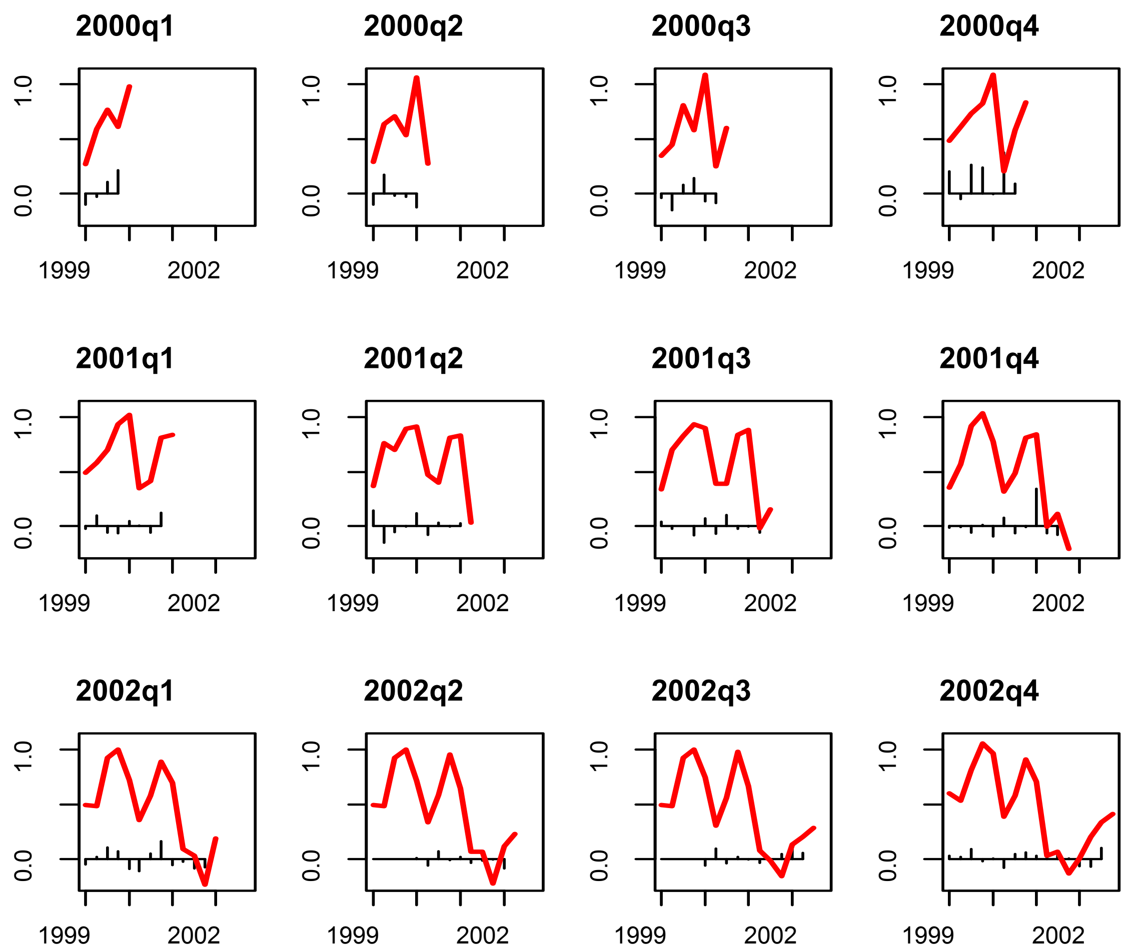


Figure 7: Quarter-on-quarter Italian GDP growth (solid line), and quarter-on-quarter revisions of contribution to GDP growth: Fixed investments.

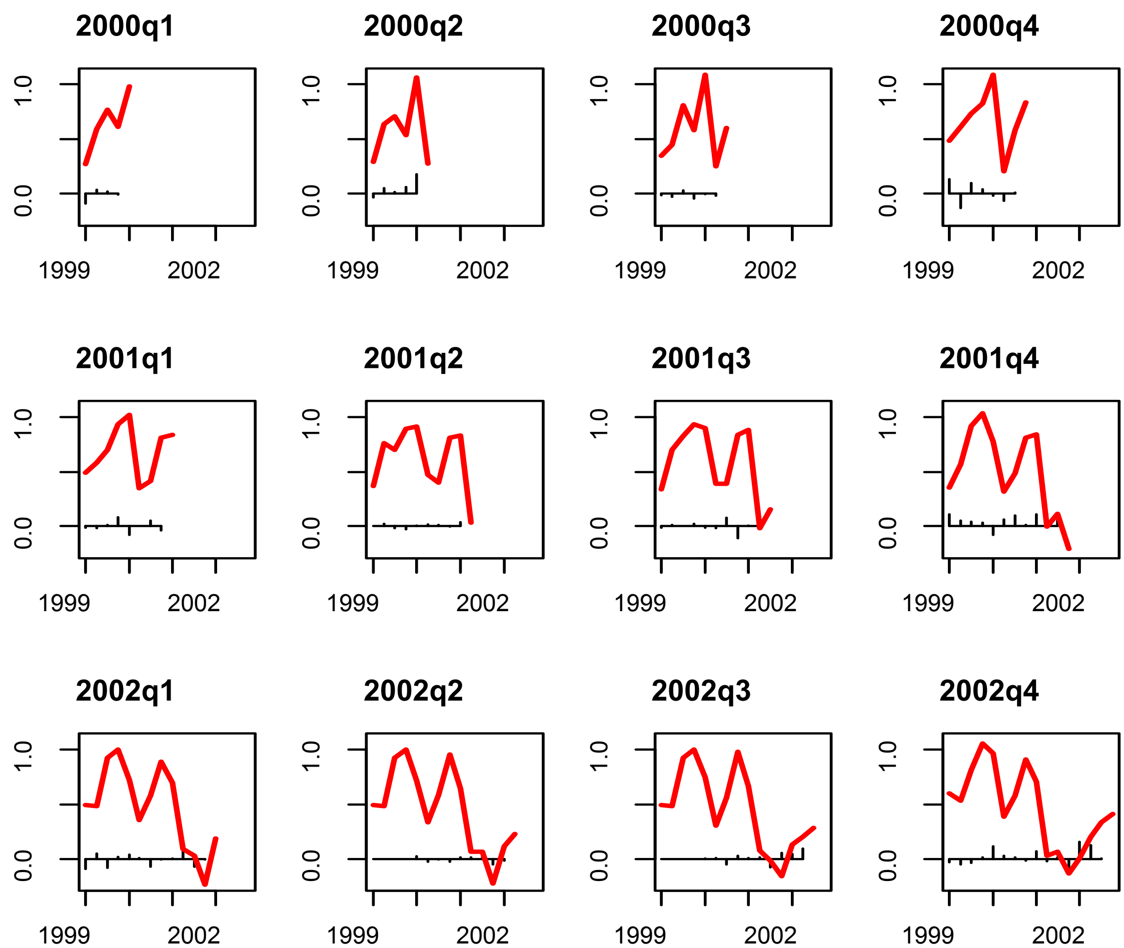


Figure 8: Quarter-on-quarter Italian GDP growth (solid line), and quarter-on-quarter revisions of contribution to GDP growth: Net exports.

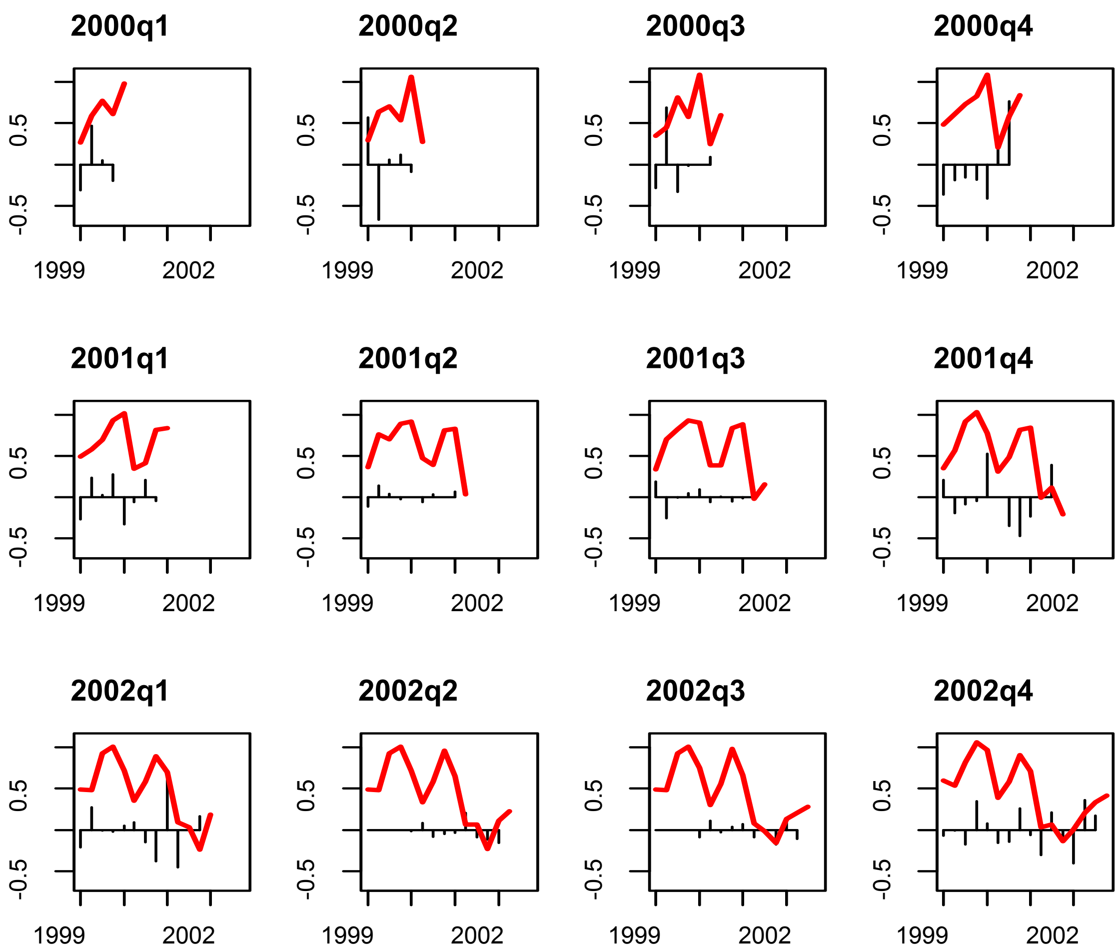


Figure 9: Quarter-on-quarter Italian GDP growth (solid line), and quarter-on-quarter revisions of contribution to GDP growth: Inventories.

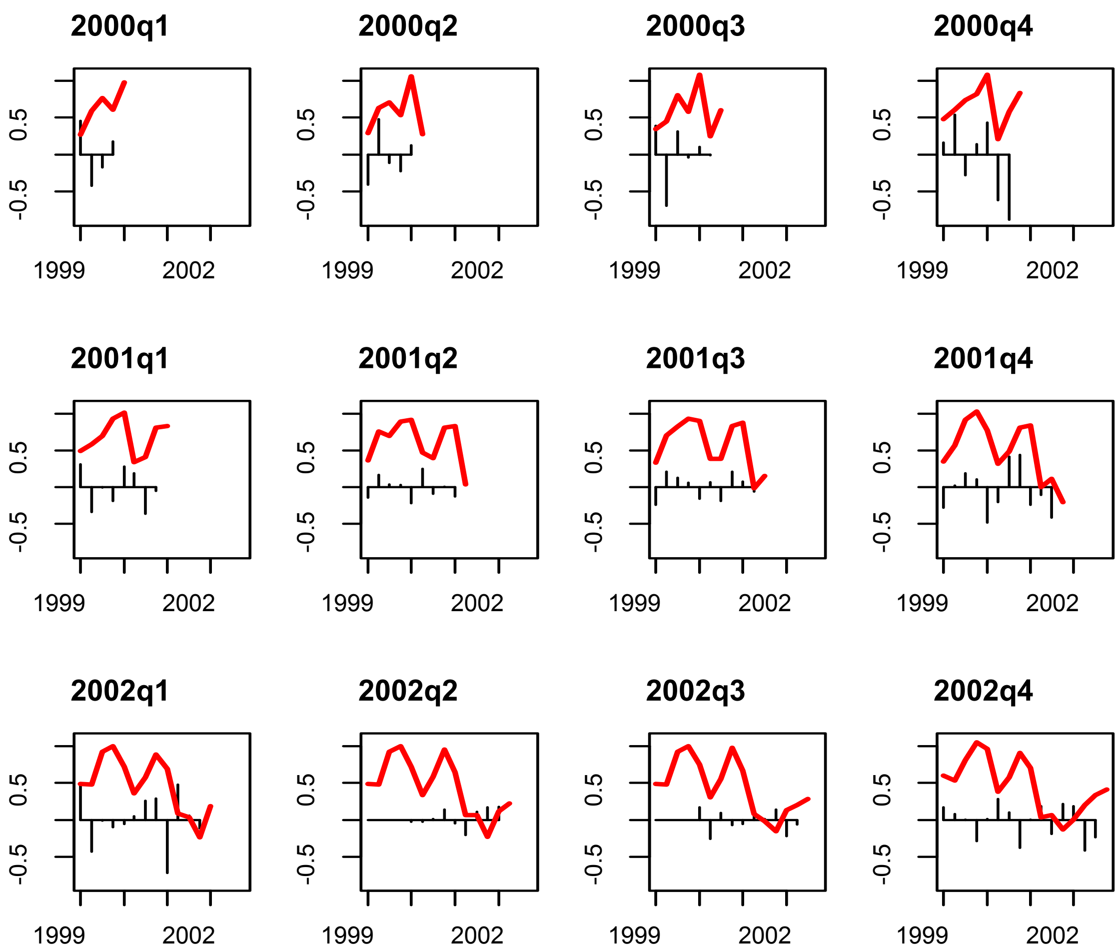


Figure 10: Naïve forecasts of annual growth rates of Italian GDP and its main components in 1999 over the last 13 revisions: solid line, first quarter; long dashed, second quarter; dashed, third quarter; dot-dashed, fourth quarter. The data for the fourth quarter coincide with the estimates of actual annual growth rates.

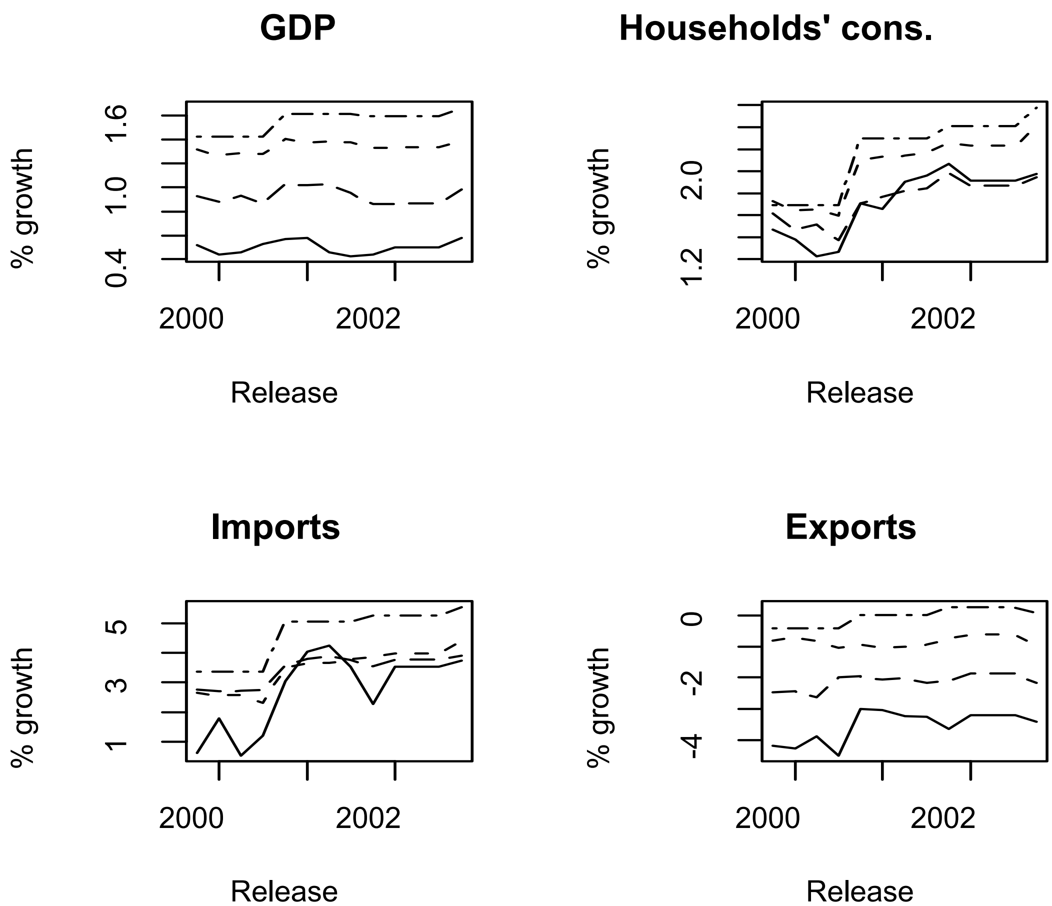


Figure 11: Naïve forecasts of annual growth rates of Italian fixed investments in 1999 over the last 13 revisions: solid line, first quarter; long dashed, second quarter; dashed, third quarter; dot-dashed, fourth quarter. The data for the fourth quarter coincide with the estimates of actual annual growth rates.

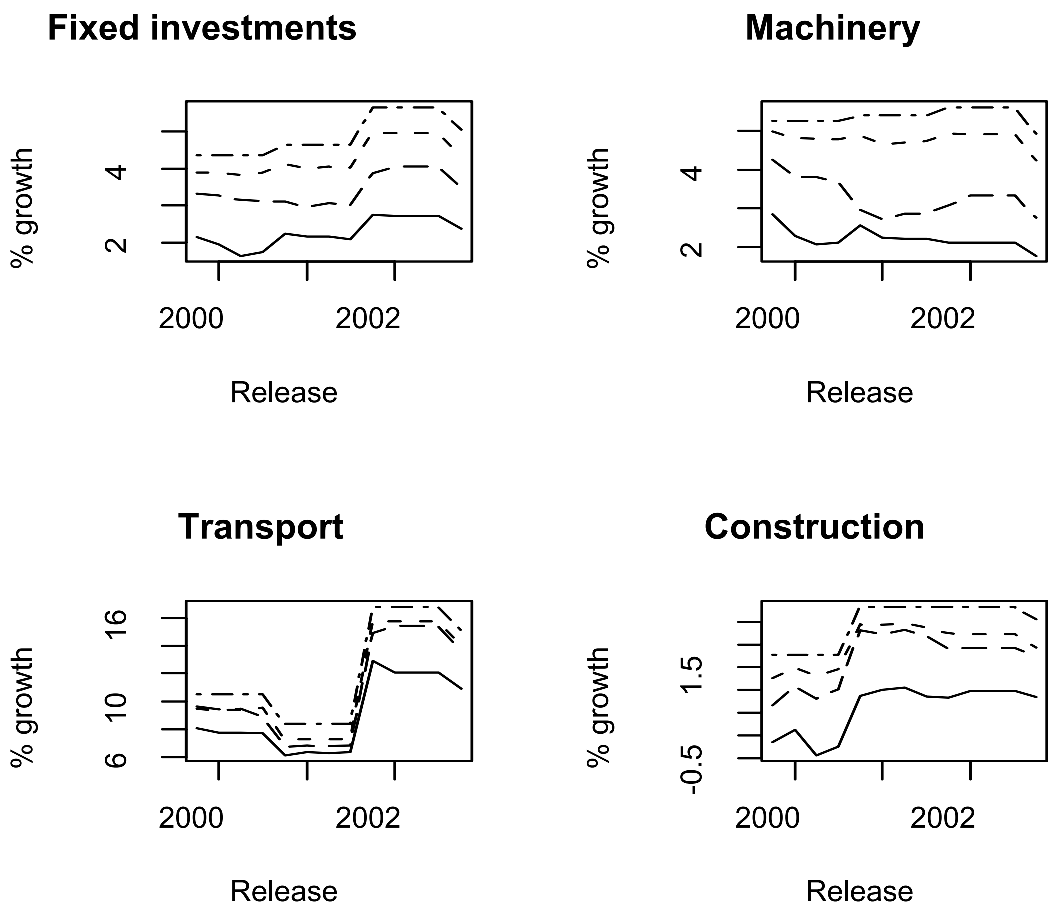


Figure 12: Quarter-on-quarter revisions of US GDP (% changes).

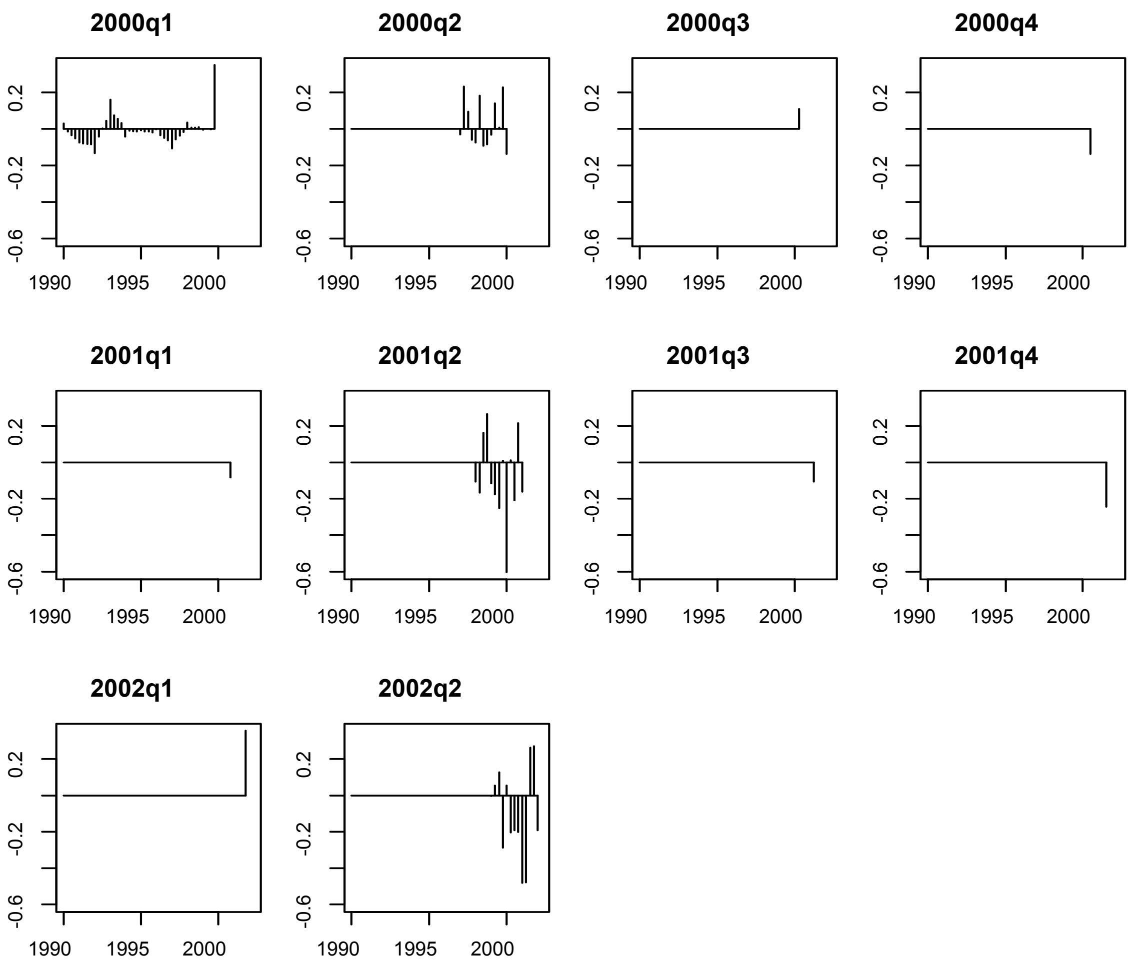


Figure 13: Quarter-on-quarter US GDP growth (solid line), and quarter-on-quarter revisions of contribution to GDP growth: Personal consumption.

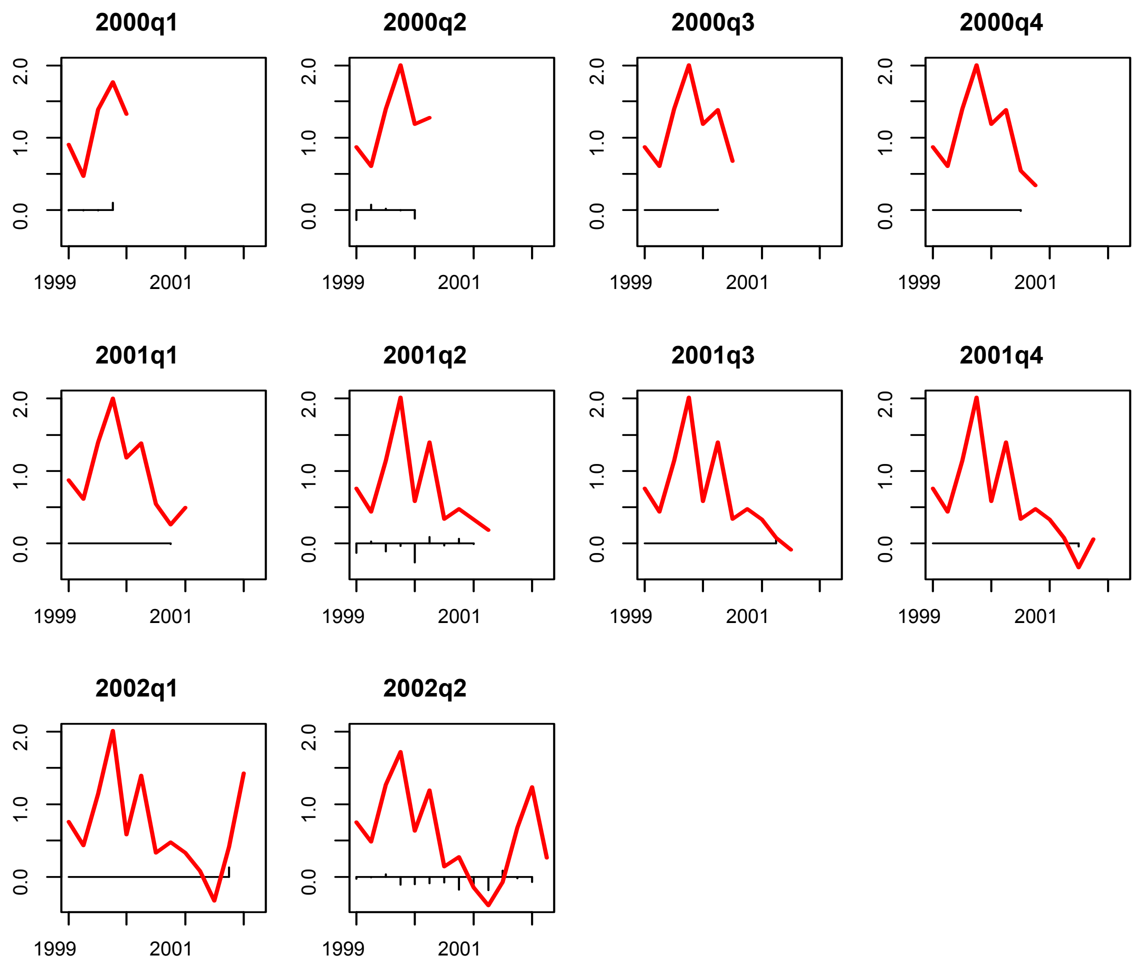


Figure 14: Estimates of the four 1999 quarters of US GDP in the last 11 revisions (quarter-on-quarter % changes): solid line, first quarter; long dashed, second quarter; dashed, third quarter; dot-dashed, fourth quarter.

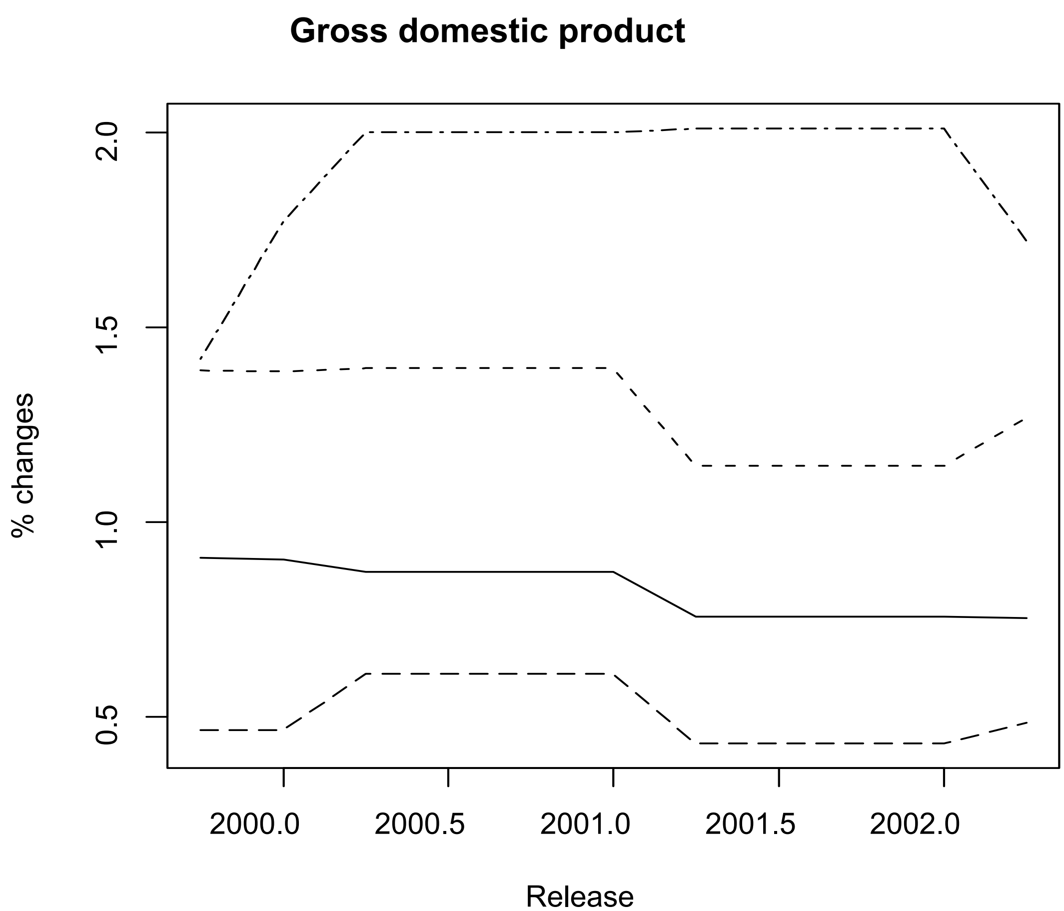


Figure 15: Estimates of the four 1999 quarters of US GDP main components in the last 11 revisions (quarter-on-quarter % changes): solid line, first quarter; long dashed, second quarter; dashed, third quarter; dot-dashed, fourth quarter.

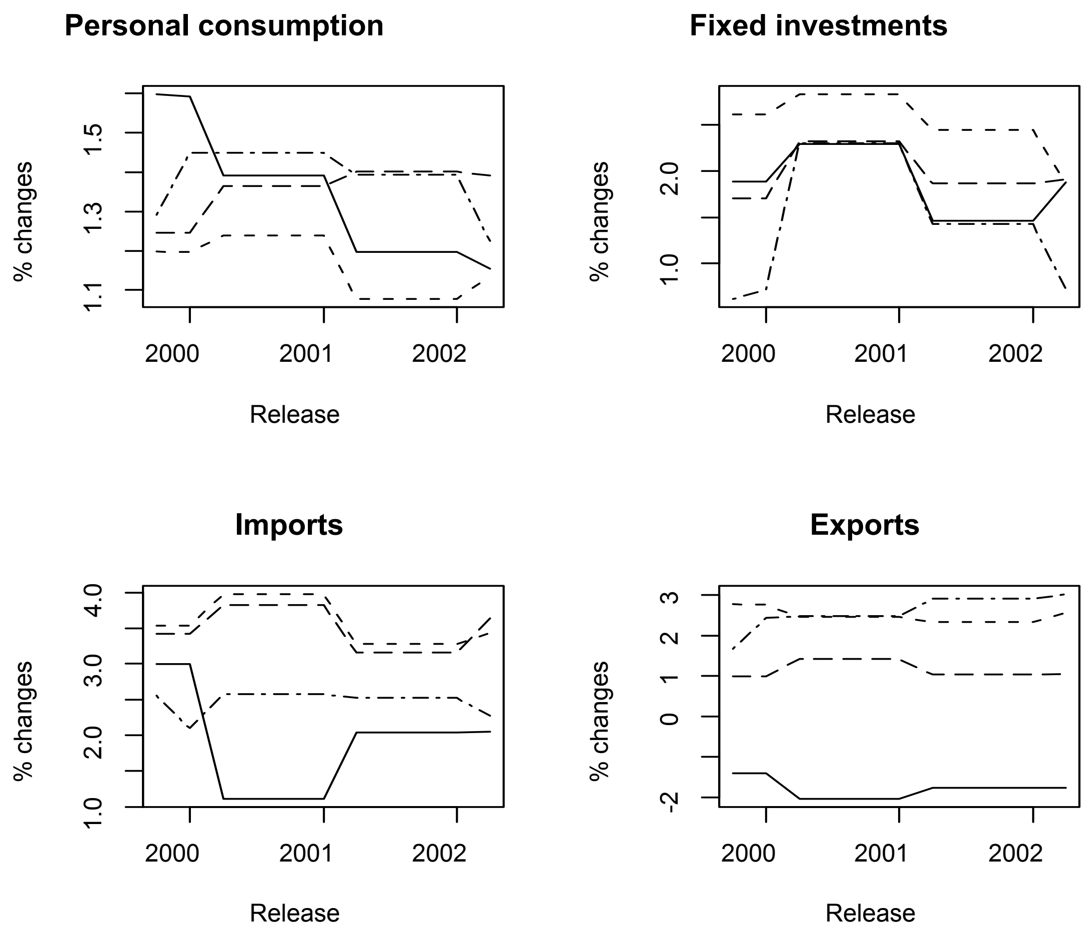


Figure 16: Estimated kernel densities of the quarter-on-quarter revisions of the quarter-on-quarter growth rates (% changes) of US GDP and its main components: Gaussian kernel with biased cross-validation bandwidth selection. Normal densities with same mean and variance of the observed data are reported for comparison (dashed lines).

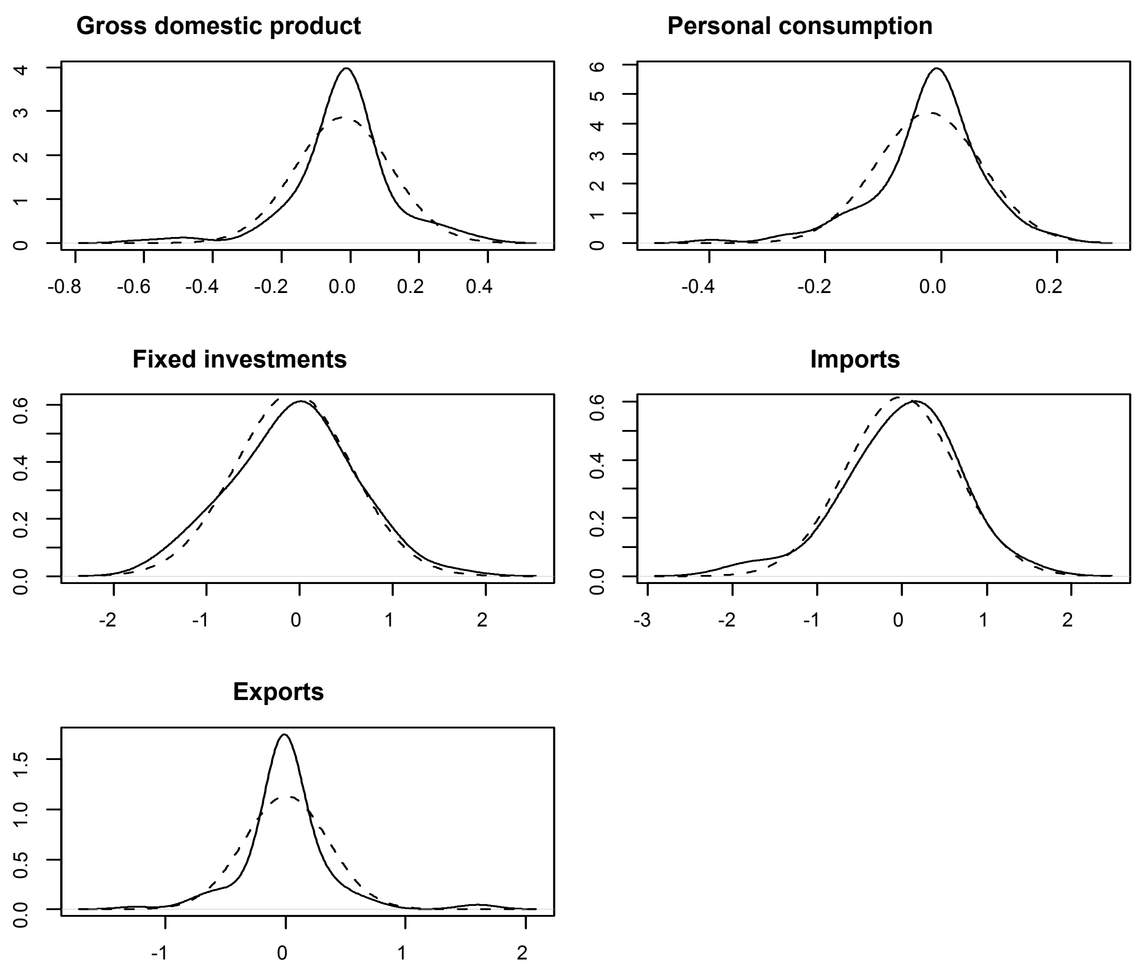


Figure 17: Estimated impulse response functions.

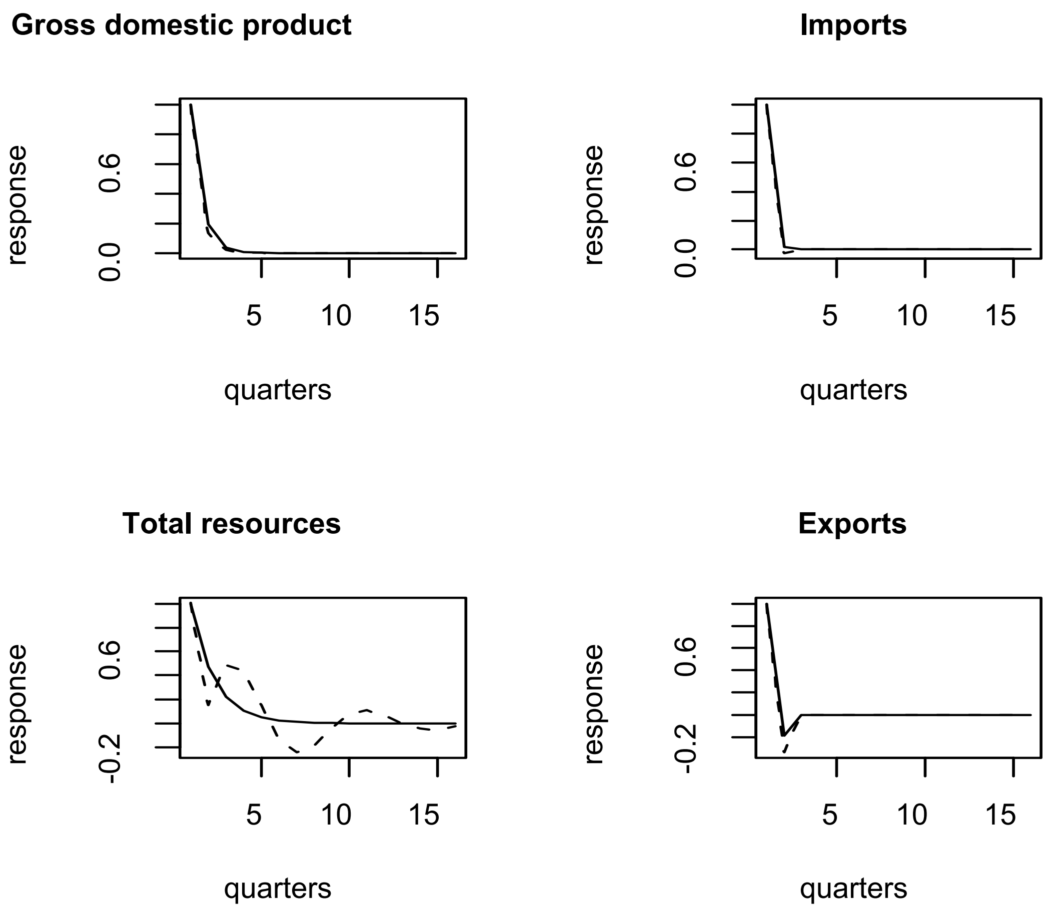


Figure 18: Estimated impulse response functions.

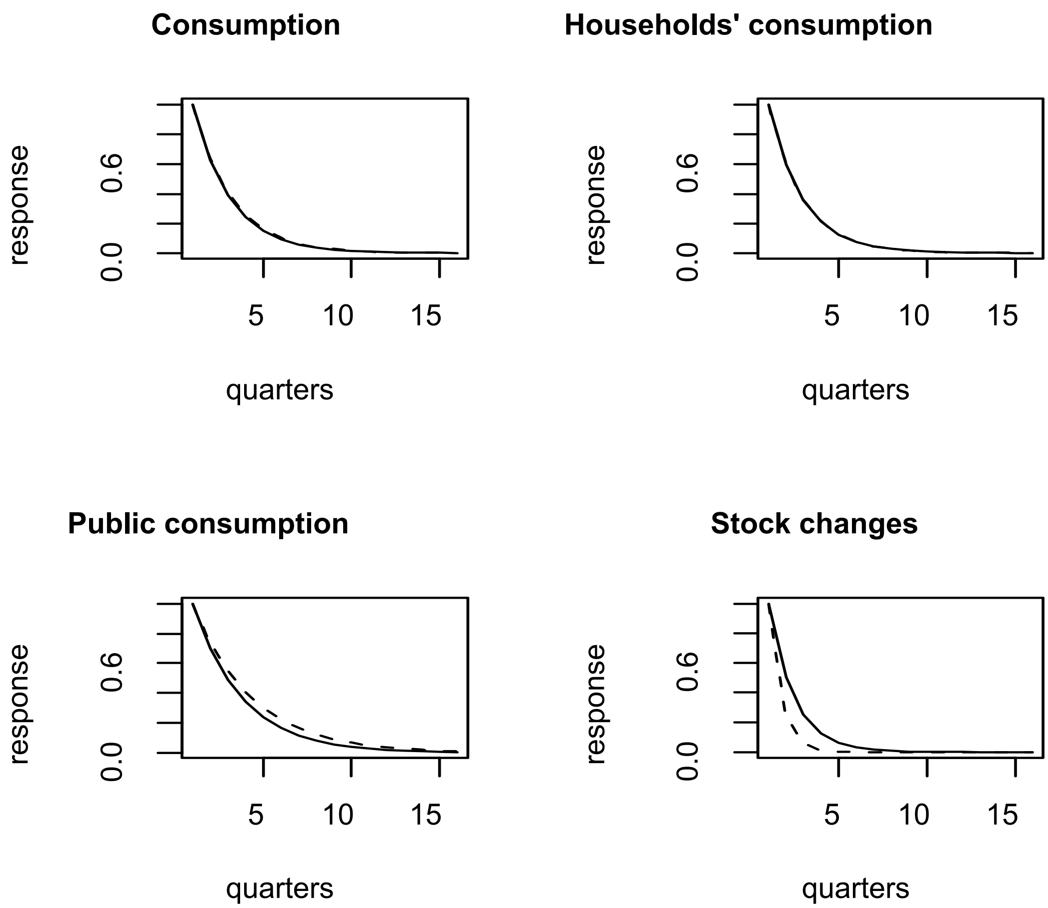
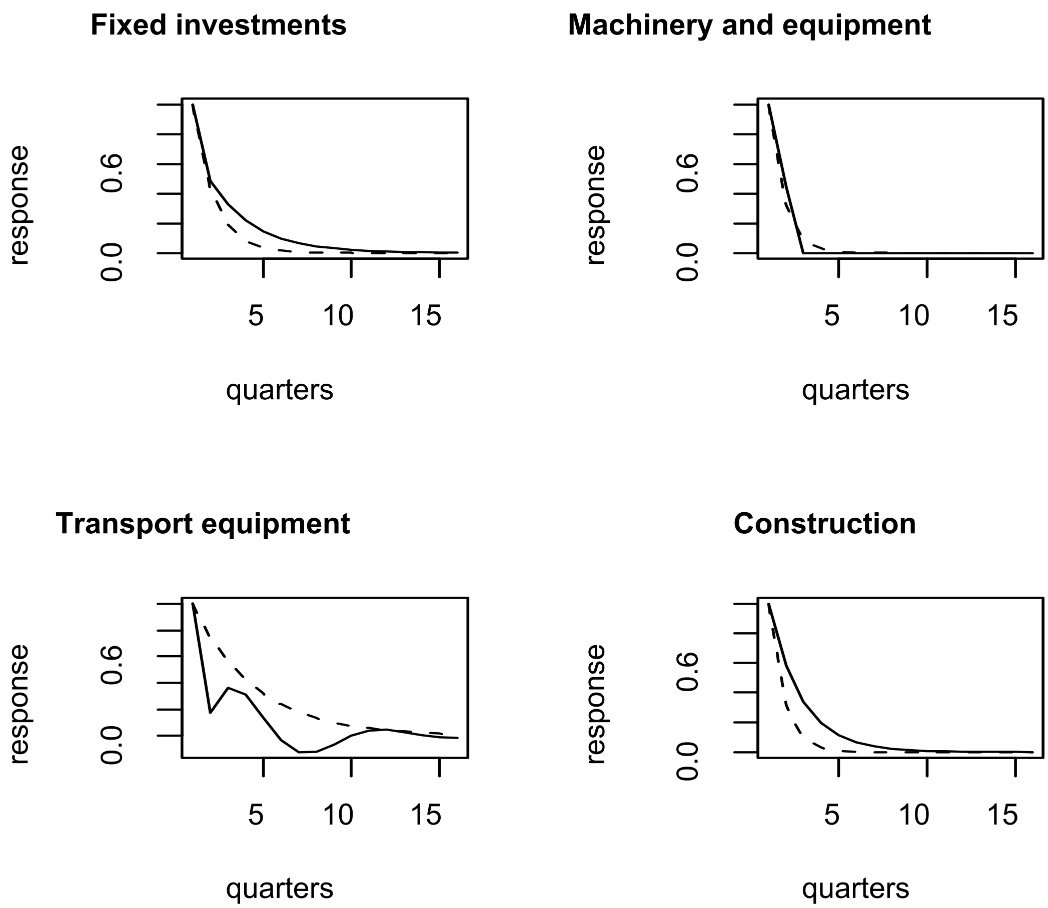


Figure 19: Estimated impulse response functions.



A The indirect estimation approach

In Italy, the main sources for the estimation of the annual national accounts are not available on a quarterly basis. Therefore, consistently with the prescriptions of ESA 1995 (Eurostat 1996, §$\tilde{1}$2.04), Italian quarterly national accounts are estimated using an indirect approach. Indeed, this is a common situation for a number of statistical agencies in Europe (see *e.g.* Bruno *et al.*, 2002). The indirect approach used by Istat traces back to the seminal contribution of Chow and Lin (1971), as modified by Barbone, Bodo and Visco (1981). A detailed discussion of the theoretical limitations of this methodology can be found in Cainelli and Lupi (1999a).

The Chow-Lin approach to time series disaggregation can be summarized as follows. Let $\{Z_t\}_{t=1}^{T}$ be a scalar time series (the "variable of interest") observed with frequency 1 (annual) from time 1 to time T. Let $\{W_t\}_{t=1}^{T}$ be a vector of $k \geq 1$ time series (the so-called "indicators"), which are related to $\{Z_t\}$.[30] By related, we mean that the indicators measure essentially the same phenomenon captured by the variable of interest, but with different features. Notably, the variable of interest is observable only at annual frequency, while the indicators are observed also at monthly or quarterly frequency. A typical example is when the variable of interest is the national accounts estimate of the production of industry i, and the indicator is the industrial production index for the same industry. The national account estimate is available only at annual frequency, while the industrial production index is published on a monthly basis. The nature of the two quantities differs for a number of reasons (coverage, numeraire, etc.), but both measure industrial output.

Let us now use lower cases to denote the high-frequency (quarterly) quantities, so that $\{w_t\}_{t=1}^{4T+\tau}$ is the indicator(s) series at the disaggregate frequency, $\{z_t\}_{t=1}^{4T+\tau}$ is the unobserved quarterly variable, and τ is an integer $(0 < \tau < 4)$ which is different from zero if we do not observe the annual value of the current year but we have τ observations for the disaggregate indicator(s). In what follows, we consider explicitly only the case $\tau = 0$. The extension to $\tau \neq 0$ is straightforward, but requires a slight complication in the notation.

We indicate with boldface fonts the whole time series of the variable and its indicator(s)

[30] Of course, when there is a single indicator, so that $k = 1$, then $\{W_t\}$ is also a scalar time series. However, this occurs rarely in real applications since the common practice is to consider a constant term among the indicators. Therefore, in general, $k \geq 2$.

so that, for example, $\mathbf{z}$ and $\mathbf{w}$ are, respectively, the vector of the unknown values of $\{z_t\}_{t=1}^{4T}$ and the matrix of observations on $\{w_t\}_{t=1}^{mT}$ at the quarterly frequency. The time aggregation from $\mathbf{z}$ to $\mathbf{Z}$ and from $\mathbf{w}$ to $\mathbf{W}$ is obtained by applying the linear transforms $\mathbf{Z} = A\mathbf{z}$ and $\mathbf{W} = A\mathbf{z}$, where the aggregation matrix A is such that $A = I_T \otimes (1, 1, 1, 1)$, with I_T the $T \times T$ identity matrix and $\otimes$ the Kronecker product.[31]

The Chow-Lin approach is based on the hypothesis that the relation between the disaggregate variable and its indicator(s) is

$$\mathbf{z} = \mathbf{w}\beta + \mathbf{u}, \tag{4}$$

where $\mathbf{u}$ is a vector of residuals such that $\mathrm{E}\,\mathbf{u} = 0$ and $\mathrm{E}\,\mathbf{u}\mathbf{u}^\top = \Omega_\mathbf{u}$. The practical implementation of the method used by Istat also assumes that

$$u_t = \rho u_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \text{iid } \mathcal{N}(0, \sigma_\varepsilon^2) \tag{5}$$

(see *e.g.* Barbone, Bodo & Visco 1981). This assumption is both an identifying restriction and a practical way to obtain rather smooth disaggregated series (see Cainelli & Lupi 1999a). Given (4), the aggregated relation is

$$\mathbf{Z} = \mathbf{W}\beta + \mathbf{U}, \tag{6}$$

with $\mathbf{U} = A\mathbf{u}$, $\mathrm{E}\,\mathbf{U} = 0$, $\mathrm{E}\,\mathbf{U}\mathbf{U}^\top = \Omega_\mathbf{U} = A\Omega_\mathbf{u}A^\top$. When $\Omega_\mathbf{U}$ is known, from (6) it is possible to derive the GLS estimator of β

$$\hat{\beta} = (\mathbf{W}^\top\Omega_\mathbf{U}^{-1}\mathbf{W})^{-1}\mathbf{W}^\top\Omega_\mathbf{U}^{-1}\mathbf{Z}.$$

Finally, the Chow-Lin estimator of $\mathbf{z}$ is given by

$$\hat{\mathbf{z}} = \mathbf{w}\hat{\beta} + L\hat{\mathbf{U}}, \tag{7}$$

where $\hat{\mathbf{U}} = \mathbf{Z} - \mathbf{W}\hat{\beta}$ are the GLS residuals from (6) and $L = \Omega_\mathbf{u}A^\top\Omega_\mathbf{U}^{-1}$. Chow and Lin (1971) prove that not only $A\hat{\mathbf{z}} = \mathbf{Z}$ (so that the estimated values sum up to the known temporally aggregated ones), but also that, if $\Omega_\mathbf{u}$ is known,[32] then (7) gives the best linear unbiased

[31] The form of the matrix A reported in the text corresponds to the aggregation by sum of quarterly values. The extension to the aggregation by averaging and to different frequencies is straightforward.
[32] The issues related to this hypothesis are discussed in detail in Cainelli and Lupi (1999a).

estimator for $\mathbf{z}$. In practice, $\Omega_{\mathbf{u}}$ is unknown and the solution advocated by Barbone, Bodo and Visco (1981) is to estimate $\Omega_{\mathbf{U}}$ by assuming $u_t \sim$ AR(1), as in (5), and by minimizing $\hat{\mathbf{U}}^\top \Omega_{\mathbf{U}}^{-1} \hat{\mathbf{U}}$ iteratively with respect to ρ, by iterating on a grid of values for ρ.

The estimation of the quarterly national accounts uses the Chow-Lin approach to temporally disaggregate national accounts variables at a very detailed level, so that the whole process requires the application of the method over thousands of time series.[33]

B State-space modelling and forecasting

The state-space representation and the Kalman filter have become a standard approach to modelling and forecasting GDP in the presence of data revisions. The approach, initially proposed by Howrey (1978) and Conrad and Corrado (1979), can easily be extended to multivariate settings (see *e.g.* Patterson 1995, and Busetti 2001).

B.1 State-space modelling

Let Y_t denote the true value of log GDP at time t (as defined on the basis of a given system of conventions, such as ESA 1995), and let $Y_{t,i}$ denote its preliminary estimate as of release or "vintage" i ($i = 1, 2, 3, \ldots$). We assume that $Y_{t,i}$ is available at time $t+i$ and that $Y_{t,d} = Y_t$, that is, Y_t is fully observed with a delay of $d \geq 1$ periods. The relationship between the ith preliminary estimate and true GDP is described by the "measurement equation"

$$Y_{t,i} = \alpha_i + \beta_i Y_t + U_{t,i}, \qquad i = 1, \ldots, d-1, \tag{8}$$

where $U_{t,i}$ is a zero-mean random variable, independent of Y_t, representing the error made in release i. To capture the serial correlation of revisions, the $U_{t,i}$ may be serially correlated. Patterson (1995) and Gallo and Marcellino (1999) notice that, if $\{Y_{t,i}\}$ and $\{Y_t\}$ are integrated processes and $\{U_{t,i}\}$ is a stationary process, then (8) is a cointegrating relationship whose cointegrating vector $(-\alpha_i, -\beta_i)$ can be estimated superconsistently by OLS.

The assumption that Y_t is fully observed with a delay of d periods implies that (8) also hold for $i = d$, with $\alpha_d = 0$, $\beta_d = 1$ and $\operatorname{Var} U_{t,d} = 0$. A preliminary estimate $Y_{t,i}$ is said to be unbiased if $\alpha_i = 0$ and $\beta_i = 1$, in which case $\mathrm{E}(Y_{t,i} - Y_t) = 0$. A test of unbiasedness of

[33] The issues related to the "right" (cross section) aggregation level to be considered in the estimation of the quarterly national accounts are addressed in Cainelli and Lupi (1999a, 1999b).

the ith release based on the OLS estimates of α_i and β_i in a regression of $Y_{t,i}$ on Y_t is also known as a Mincer-Zarnowitz test.

Log GDP is assumed to obey the dynamic model

$$\Phi(L)\, Y_t = \mu + V_t, \tag{9}$$

where $\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ is a polynomial in the lag operator L and $\{V_t\}$ is a stationary process. Notice that some roots of $\Phi_p(L)$ may lie on the unit circle, that is, $\{Y_t\}$ may be an integrated process. Model (9) may easily be put in "companion form", thus completing the state-space representation of the system. Optimal forecasts of Y_{t+j} based on the information available up to time t may then be computed recursively using the Kalman filter (see *e.g.* Harvey 1989).

An important special case of (8)–(9) is when $\Phi(L) = (1-\phi L)(1-L)$ and $\{V_t\}$ is a white noise process with zero mean and finite variance σ^2_V. In this case, the growth rate of GDP, $X_t = Y_t - Y_{t-1}$, follows a stationary stable AR(1) process with mean equal to $\mu/(1-\phi)$, that is,

$$X_t = \mu + \phi X_{t-1} + V_t.$$

If the preliminary estimates of GDP are unbiased, then

$$X_{t,i} = X_t + \epsilon_{t,i},$$

where $X_{t,i} = Y_{t,i} - Y_{t-1,i+1}$ is the ith preliminary estimate of the GDP growth rate available at time $t+i$ and $\epsilon_{t,i} = U_{t,i} - U_{t-1,i+1}$. The AR(1) plus noise model for GDP growth corresponds to the case when $\epsilon_{i,t}$ is a white noise process with zero mean and finite variance σ^2_ϵ.

B.2 Forecasting

If GDP growth follows a stable invertible process with MA(∞) representation $X_t = \sum_{h=0}^{\infty} \psi_h V_{t-h}$, where $\{V_t\}$ is a white-noise process, then its best (i.e. minimum MSE) j-step-ahead forecast is

$$\hat{X}_{t+j|t} = \sum_{h=0}^{\infty} \psi_{j+h} V_{t-h} = \sum_{h=0}^{\infty} \psi_{j+h} (\sum_{s=0}^{\infty} \pi_s X_{t-h-s}) = \sum_{h=0}^{\infty} \pi_h^{(j)} X_{t-h},$$

where $\{\pi_h\}$ are the weights in the AR(∞) representations of the process, $\pi_0^{(j)} = \psi_j$, $\pi_1^{(j)} = \psi_j\pi_1 + \psi_{j+1}$, $\pi_2^{(j)} = \psi_j\pi_2 + \psi_{j+1}\pi_1 + \psi_{j+2}$, etc. The associated forecast MSE is

$$\text{MSE}(\hat{X}_{t+j|t}) = \sigma_V^2 \sum_{h=1}^{j} \psi_{h-1}^2.$$

A naïve j-step-ahead forecast may be constructed on the basis of model (9) using the most recent preliminary observation of the unobserved true values. This forecast ignores the distinction between provisional and final data. If the preliminary estimate is unbiased, then the naïve j-step-ahead forecast of GDP growth is

$$\tilde{X}_{t+j|t} = \sum_{h=0}^{d-1} \pi_h^{(j)} X_{t-h,h+1} + \sum_{h=d}^{\infty} \pi_h^{(j)} X_{t-h} = \hat{X}_{t+j|t} + \sum_{h=1}^{d-1} \pi_h^{(j)} \epsilon_{t-h,h+1},$$

where $\epsilon_{t-h,h+1} = X_{t-h,h+1} - X_{t-h}$. The general expression for the associated forecast MSE is given in Busetti (2001).

In the special case of an AR(1) plus noise model with $d = 1$, the naïve forecast of X_{t+j} is

$$\tilde{X}_{t+j|t} = \phi^j X_{t,1},$$

and the associated forecast MSE is

$$\text{MSE}(\tilde{X}_{t+j|t}) = \sigma_V^2 \sum_{h=1}^{j} \phi^{2(h-1)} + \phi^{2j}\sigma_\epsilon^2.$$

Notice that the 2-step-ahead naïve forecast $\tilde{X}_{t+2|t}$ has smaller MSE than the 1-step-ahead naïve forecast $\tilde{X}_{t+1|t}$ whenever

$$\text{Var}\, X_t = \frac{\sigma_V^2}{1-\phi^2} < \sigma_\epsilon^2 = \text{Var}\, \epsilon_{ti}.$$

If $d = 1$, the best forecast is instead

$$\tilde{X}^*_{t+j|t} = \phi^j \left[\beta X_{t,1} + (1-\beta)\phi X_{t-1}\right],$$

where $\beta = \sigma_V^2/(\sigma_V^2 + \sigma_\epsilon^2)$, and the associated forecast MSE is

$$\text{MSE}(\tilde{X}^*_{t+j|t}) = \sigma_V^2 \sum_{h=1}^{j} \phi^{2(h-1)} + \beta\sigma_\epsilon^2\phi^{2j}.$$

The best forecast of X_{t+j} may be written

$$\tilde{X}^{*}_{t+j|t} = \phi^{j}\,[\beta X_{t,1} + (1-\beta)\hat{X}_{t|t-1}],$$

where $\hat{X}_{t|t-1} = \phi X_{t-1}$ is the one-step-ahead forecast from an AR(1) model. This expression suggests two alternative methods for improving upon the naïve forecast in more general models. The first method consists of modifying the forecast initial conditions. In the AR(1) plus noise model, this involves replacing the naïve forecast $\phi^{j} X_{t,1}$ by $\phi^{j} X^{*}_{t}$, where

$$X^{*}_{t} = \beta X_{t,1} + (1-\beta)\hat{X}_{t|t-1}$$

is a convex combination of the preliminary estimate $X_{t,1}$ and the one-step-ahead forecast from the model. The second method consists of an intercept correction. In the AR(1) plus noise model, this involves replacing $\phi^{j} X_{t,1}$ by $\phi^{j}(X_{t,1} + c_t)$, where

$$c_t = X^{*}_{t} - X_{t,1} = -(1-\beta)(X_{t,1} - \hat{X}_{t|t-1}),$$

with $c_t = 0$ if $X_{t,1} = \hat{X}_{t|t-1}$.

Both methods are fully optimal for the AR(1) with noise model but need not be so for more general models.