Predicting Recoveries

and the Importance of Using Enough Information

Xiaoming Cai and Wouter J. Den Haan*

October 22, 2009

Abstract

Several papers that make forecasts about the long-term impact of the current financial crisis rely on models in which there is only one type of financial crisis. These models tend to predict that the current crisis will have long lasting negative effects on economic growth. This paper points out the deficiency in this approach by analyzing the ability of "one-type-shock" models to correctly forecast the recovery from past economic downturns. It is shown that these models often overestimate the long-run impact of recessions and that slightly richer models that allow the effects of recessions to be both persistent and transitory predict recoveries much better.

Key Words: forecasting, unit root, financial crisis, great recession,

JEL Classification: C51,C53,E37

*Cai: Tinbergen Institute. E-mail: `xcai@mail.tinbergen.nl`. Den Haan: University of Amsterdam and CEPR. E-mail: `wdenhaan@uva.nl`. We would like to thank Bas Jacobs, Coen Teulings and Sweder van Wijnbergen for useful comments.

1 Introduction

The sharp economic downturn and aggressive stimulus packages have lead to substantial increases in budget deficits and government debt levels. Given this deterioration of government finances, the amount of economic growth countries will experience in the upcoming decade is of vital importance for several reasons. One of those reasons is that economic growth determines tax revenues and claims on government services. Consequently, economic growth is key in understanding what will happen with the budget deficit and government debt. If governments are serious about preventing deficits and debt levels from remaining too high for too long, then the amount of economic growth will determine how much action governments will have to undertake in terms of cutting government expenditures and/or raising taxes.

Not surprisingly, government agencies are keen to make predictions about economic growth. Given that the current situation is unusual, this is clearly not an easy task. One of the central questions is whether the current economic crisis will have permanent level and/or growth effects. There are both empirical and theoretical arguments that such permanent effects could occur (and also that they should not occur).

Forecasts about economic growth not only provide information for the latitude policy makers have in setting fiscal policy. They are also an important input for the expectations of the private sector. For these reasons, it is important that such forecasts are done with utmost care; forecasts that are too pessimistic or too buoyant could induce the wrong decisions and be quite harmful. This would especially be the case if the forecasts made by government (and other types of) institutes are publicized to the world without making explicit the underlying uncertainty in them.

Predictions for GDP are typically obtained with time series models estimated using past data. In the last section of this paper, we give a comprehensive discussion on the elements serious forecasting exercises should include and we will argue that one cannot rely on time series models alone. Until we reach that point of the paper, however, we limit the focus of our paper to time series models.

Medium and long-term forecasts of the effects of the current financial crisis differ widely.

According to Figure 12 in CBO (2009), the U.S. economy would be back to its pre-crisis trend around 2015. Chapter 4 in IMF (2009) points out that countries have experienced widely different growth patterns after financial crises, ranging from medium-term declines in GDP of 26% to *increases* of +6%.[1] The discussion in IMF (2009) also makes clear that there is not one type of crisis and that the damage of the crisis depends on, for example, the strength of the monetary and fiscal stimulus injected at the onset of the crisis.

On his blog,[2] Greg Mankiw questioned positive forecasts by the Council of Economic Advisors (CEA) by reminding the reader of the analysis in Campbell and Mankiw (1987) that reaches the following conclusion:

> The data suggest that an unexpected change in real GNP of 1 percent should change one's forecast by over 1 percent over a long horizon.

The analysis in Campbell and Mankiw (1987) is based on a linear univariate forecasting model for output of the following form:

$$y_t = a_1 y_{t-1} + ... + a_k y_{t-k} + e_t, \tag{1}$$

where e_t is the unexpected shock that hit the economy. This time series model has the following properties: (*i*) there is only one-type of shock, that is, the response of output to e_t is always the same independent of why there is a shock to output, (*ii*) the responses of output are linear in the magnitude of the shock, that is, if the shock is twice as large then the responses are simply doubled, and (*iii*) the only variables that are used to predict output are lagged values of output. The specification in Equation (1) looks ridiculously simple, but for *short*-term forecasts (say on a monthly or perhaps quarterly basis), it often does a decent job. In fact, Campbell and Mankiw (1987) find that

$$y_t = y_{t-1} + e_t \tag{2}$$

[1] Countries for which the long-term growth path improved after the crisis are Chili after the 1981 crisis, Argentina after the 1989 crisis, Mexico after the 1994 crisis, and Uganda after the 1994 crisis. All four countries did implement some serious structural reforms ranging from implementing the NAFTA trade agreement by Mexico to major financial reforms by Argentina.

[2] See http://gregmankiw.blogspot.com/2009/03/team-obama-on-unit-root-hypothesis.html.

is not a bad representation of the data, that is, U.S. output follows a random walk. We will refer to the time series model of Equation (1) as the one-type-shock model. As will be made clear below, this name makes clear the model's main deficiency.

Many forecasts currently being made are based on a time series model like the one given in Equation (1) or modified versions. For example, van Ewijk and Teulings (2009) allow the time-series model to have a regular shock and a financial crisis shock that are allowed to have different effects on output. But their specification still only allows for one type of financial crisis, which—as is made clear in IMF (2009), is not a sensible assumption.

This paper consists of a methodological part and an application. In the methodological part, we argue that models like the one given in Equation (1) are not suited to make long-term predictions. In particular, we show that approaches based on models like the one given in Equation (1) have the following problems.

- They are likely to generate biased long-term predictions, because for variables like GDP that are sums of other variables they are likely to be misspecified. In such a situation, there are clear benefits of modelling the components directly.

- Since there is only one-type-of-shock, or in the modified specification only one-type-of-financial-crisis, an unexpected drop in output of a certain size will always lead to the same long-term predictions, completely independent of what other variables do. This is obviously a strong limitation, because—as pointed out in IMF (2009)—there are financial crises that only have transitory effects on output and even financial crises that have positive long-term effects. We will show that by incorporating other variables into the time series model, one can much better determine whether economic downturns are due to shocks with permanent or only transitory consequences.

- Moreover, even if the time series model specified in Equation (1) is the correct *linear* process (i.e., there is no misspecification as discussed in the first bullet point) and there are no other variables available (i.e., the type of efficiency gains discussed in the second bullet point are not possible), then it is possible that one makes systematic mistakes when forming predictions on one-type-shock models. In particular, we give an example of an environment for which Equation (2) is the correct linear univariate time series model, but conditional on being in a recession, the forecasts based on Equation (2) systematically overpredict the long-term negative consequences.

These two objections are not novel, but given the arguments used in the current forecasting debate, it seems useful to bring these arguments to the forefront. We rely in particular on the work of Nobel prize winner Clive Granger, a former colleague of one of the authors.

What is novel in this paper is the application with which we document the enormous quantitative importance of our arguments for forecasting U.S. GDP during economic downturns. We compare the one-type-shock model with a model that predicts the components of GDP and obtains forecasts for GDP by explicitly aggregating the forecasts of the components. We will refer to this model as the "components" model. This model not only allows for different types of shocks, it also is more likely to be correctly specified.

The forecasts of these two models at the troughs of NBER recessions are compared with each other and with the actual realizations. We find that during economic downturns the one-type-shock model is in general too pessimistic, that is, it underestimates the ability of the economy to bounce back. It is important to realize that both models capture the fact that GDP is, or at least is close to, an I(1) process. That is, in the components model there are shocks that have a permanent effect. In contrast to the one-type-shock model, the components model realizes that not all shocks have large permanent effects and that some shocks even have no permanent effects at all. This aspect of the components model makes it possible to forecast better how the economy will get out of a recession.

Section 2 contains an allegory on predicting the consequences of epidemics to intuitively explain the main points of this paper. In Section 3, we will provide some theoretical background useful in understanding the drawbacks of one-type-shock models. In Section 4, we discuss the differences in the forecasts made during economic downturns including the current one. In the last section, we give our views on the elements a responsible forecasting exercise should contain.

2 Allegory; predicting the consequences of epidemics

The main point of our paper is that models that directly forecast aggregate output allowing for only one type of financial crisis are likely to predict that the impact of the shock is permanent even when there are shocks that do not have permanent effects. In this section, we illustrate the reason with a simple allegory. We first describe what the world actually looks like and then we describe what the econometrician knows and how he can best use the available information to make predictions.

Our hypothetical world is occasionally faced with an epidemic, the analogue of the financial crisis. There are only two types of illnesses in this world and consequently only two types of epidemics. Initially, these illnesses have indistinguishable effects on humans health, say high fever, weight loss, and loss of energy. The long-term consequences, however, are very different. One illness hardly ever leads to death, but the other one unfortunately does.[3] For readability we will refer to these as the "flu" and the "AIDS" epidemic, but the reader should understand that in our hypothetical world these two illnesses only have the characteristics we explicitly attribute to them.

In this world, the health of person i in period t can be measured with an index, $h_{i,t}$. A person's health index is equal to 1 if he is perfectly healthy and less than 1 if he is not. Thus, a reduction in $h_{i,t}$ means that person i's illness is getting worse. If the health index is equal to 0, then the person is dead, a state from which he cannot recover. Aggregate health in this country is given by the sum of all health indices, that is,

$$H_t = \sum_{i=1}^{I} h_{i,t}. \tag{3}$$

This is, of course, the analogue to GDP.

We consider two types of econometricians. Both types know whether an epidemic has started. But we focus on the initial phase of the epidemic at which point, it is still impossible to figure out what type of illness a sick person has. The first type of econometrician only uses data on H_t and he also knows whether an epidemic has started.

[3] That is, one type of epidemic is like the financial crisis of Sweden and the other one is like the financial crisis of Japan.

The information used by the other econometrician will be described below. What will this first type of econometrician predict when the epidemic hits a country. This econometrician will use data on past epidemics and since he does not know what type of epidemic has hit his country, he will use data on *all* past epidemics, that is, of both types. His predictions will then be some kind of average of the outcomes of the two types of epidemics. This means that at the outbreak of *every* epidemic, this econometrician will predict that people will die, even though there are epidemics without fatalities in our world. Similarly, we will show that econometricians that (i) use past data and (ii) directly try to forecast GDP allowing for only one type of financial crisis will tend to find permanent effects. The reason is that there are indeed shocks to GDP that have a permanent effect (just like there are fatalities in some epidemics) and this disastrous outcome will always get some weight.

The other econometricians does not focus directly on H_t, but uses data on the health index of adults, $H_{a,t}$, and the health index of children, $H_{c,t}$. Note that

$$H_t = H_{a,t} + H_{c,t}. \tag{4}$$

To make the point of the paper most clear, assume that children never get AIDS, but can get the flu. This means that the econometrician that observes both $H_{a,t}$ and $H_{c,t}$ would immediately know whether it is a flu or an AIDS epidemic. Even if observations are noisy, e.g., because children sometimes do get AIDS, it is clear that by directly focusing on the two components, $H_{a,t}$ and $H_{c,t}$, this econometrician can quickly figure out whether his country was hit by an AIDS or a flu epidemic, even when a physician would not yet be able to tell whether a sick person is suffering from AIDS or the flu.

In this example, we assumed that children could not get AIDS. But this is not necessary for the point we make. First note that this assumption does not affect the first econometrician's prediction. This econometrician's prediction is always that there will be fatalities. All that is needed for the second econometrician to do better is that two components, $H_{a,t}$ and $H_{c,t}$, respond differently to the two types of epidemics. For example, if children get the flu as quickly as adults, but take more time to get sick when infected with the AIDS virus, then the time series properties of $H_{a,t}$ and $H_{c,t}$ will reveal much quicker than the time series properties of H_t whether a new epidemic will result in fatalities or not.

The lesson of this parable is that econometricians should not just focus on GDP. If one does this, then one is bound to predict that the current crisis will have permanent effects.

3 Some econometric background

In this section, we review some results from the econometric literature that demonstrate how easily one-type-shock models can miss important information and that one-type-shock models for aggregate variables are likely to be misspecified.

3.1 Lack of essential information in one-type-shock time series models

3.1.1 Showing how forecastable components can disappear

In the introduction, we already mentioned that there are prominent economists that believe that GDP behaves like a random walk. That is, output, y_t, behaves according to

$$\begin{aligned} y_{t+1} &= y_t + e_{t+1}, \\ \mathrm{E}\left[e_t\right] &= 0 \\ \mathrm{COV}\left[e_{t+1}, e_{t-j}\right] &= 0 \text{ for } j = 0, 1, \cdots . \end{aligned} \tag{5}$$

where e_t is the shock. If the only information available about y_t is this law of motion, then a one unit drop in output caused by a negative shock to e_t should lead to a reduction in the forecast of all future output levels with one unit. That is, the shock is predicted to have permanent consequences.

Our argument is not that this is representation for GDP is that bad. Although there may be some initial transitional dynamics, one can find empirical support for laws of motions close to this specification for several countries. The point we want to make is that even if the behavior of y_t is close to that of a random walk, there still may be important forecastable changes in y_t and that some shocks may have no permanent effects. To see why this is possible, consider the following data generating process (*dgp*) for y_t:[4]

$$
\begin{aligned}
y_t &\equiv x_t + z_t \\
x_t &= \rho_x x_{t-1} + \sigma_{e,x} e_{x,t} \\
\Delta z_t &= \rho_z \Delta z_{t-1} + \sigma_{e,z} e_{z,t} \\
\mathrm{E}_t\left[e_{x,t}\right] &= \mathrm{E}_t\left[e_{z,t}\right] = 0 \\
\mathrm{E}_t\left[e_{x,t}^2\right] &= 1 \\
\mathrm{E}_t\left[e_{z,t}^2\right] &= 1 \\
\mathrm{E}_t\left[e_{x,t} e_{z,t}\right] &= 0 \\
|\rho_x| \leq 1 &\text{ and } |\rho_z| < 1
\end{aligned}
\tag{6}
$$

If

$$\rho_x = \rho_z = \rho,\ \sigma_{e,x}^2 = \rho\sigma_e^2,\ \sigma_{e,z}^2 = (1-\rho)^2\sigma_e^2,\ , \text{ and } \rho \in [0,1) \tag{7}$$

then y_t follows a random walk as in Equation (5).[5] That is,

$$\mathrm{COV}\left[\Delta y_t \Delta y_{t-j}\right] = 0 \text{ for } j \geq 1, \tag{8}$$

which implies that if one restricts oneself to linear univariate models that following a negative shock to e_t, the best prediction for the long-term impact of this shock for y_t is indeed the most recent value. But as long as $\rho_z \neq 0$ and $\rho_x \neq 1$, then shocks to both $e_{x,t}$ and $e_{z,t}$ induce predictable changes in x_t and z_t, respectively. Consequently, shocks to $e_{x,t}$ and $e_{z,t}$ lead to predictable changes in y_t. Without loss of generality, we will set $\sigma_e = 1$ in the rest of the paper.

[4] This example is from Blanchard, L'Huillier, and Lorenzoni (2009).

[5] If this condition holds, then

$$\mathrm{VAR}(\Delta y_t) = \frac{2\sigma_{e,x}^2}{1+\rho} + \frac{\sigma_{e,z}^2}{1-\rho^2} = \sigma_e^2$$

and

$$\mathrm{COV}(\Delta y_t, \Delta y_{t-j}) = -\frac{\rho^{j-1}(1-\rho)\sigma_{e,x}^2}{1+\rho} + \frac{\rho^j \sigma_{e,z}^2}{1-\rho^2} = 0 \text{ for } j \geq 1.$$

If all the autocovariances are zero, then the specturm is flat and Δy_t is white noise and y_t is a random walk.

3.1.2 Showing how forecastable components can reappear

Suppose that one would also observe another variable $\tilde{y}_t$ which is equal to

$$\tilde{y}_t = \alpha_x x_t + \alpha_z z_t. \tag{9}$$

Then one could solve for $e_{x,t}$ and $e_{z,t}$ from

$$\begin{bmatrix} e_{x,t} \\ e_{z,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \alpha_x & \alpha_z \end{bmatrix}^{-1} \begin{bmatrix} y_t \\ \tilde{y}_t \end{bmatrix}. \tag{10}$$

That is, one does not have to observe x_t and z_t themselves. All that is needed is that one observes one more variable. Then one can actually correctly identify the transitory and the permanent component.

3.1.3 Is the predicted long-run impact correct on average?

To make the discussion as simple as possible, we focus now on a particular version of the *dgp* given in Equation (6). First, we assume that Equation (7) is satisfied, so that the specification in Equation (5) is the correct linear univariate model for y_t. Second, we assume that $e_{x,t}$ and $e_{z,t}$ can take on only two values, namely -1 and $+1$, both with equal probability. Finally, we assume that $\rho = 0.381966$. For this value of ρ, Equation (7) implies that $\sigma_{e,x} = \sigma_{e,z}$. Consequently, if $e_{x,t}$ and $e_{z,t}$ have the opposite sign, the value of y_t remains unchanged. Let $y^f_{t,t+\tau}$ be the τ-period ahead forecast for y_t according to Equation (5), and let $\mathrm{E}_t\left[y_{t,t+\tau}\right]$ be the conditional expectation according the true *dgp*. Also, note that ρ_z and σ_z are such that a unit shock in $e_{z,t}$ has a long-run impact on y_t equal to 1 and, of course, of the same sign.

The one-type-shock *dgp* specified in Equation (5) clearly misses some useful information; if one would observe one more variable, then one typically can determine what the values of x_t and z_t are and make much better forecasts. Although not the most efficient, it may still be the case that the model given in Equation (5) generates (long-term) predictions that are *on average* correct. That is, sometimes one overestimates the long-run impact of a shock and sometimes one underestimates it, but these errors would average out.

As long as Equation (7) holds, then the specification in Equation (5) is the correct linear univariate model for y_t.[6] Consequently, long-term predictions are not biased. Nevertheless, it is still true that in recessions, they systematically overpredict the long-term consequences of the shock.

The idea is the following. Consider the four possible combinations of positive and negative shocks for $e_{x,t}$ and $e_{z,t}$. If the signs of the two shocks differ, then a positive and a negative shock exactly offset each other since $\sigma_{e,x} = \sigma_{e,z}$, that is, $y_t - y_{t-1} = 0$. In this case, the long-run prediction of the effect of the shock according to the model of Equation (5) would not be affected. That is,

$$\begin{aligned} y^f_{t,t+\tau} &= y_t \quad \forall \tau \geq 0 \text{ or} \\ y^f_{t,t+\tau} - y_t &= 0 \quad \forall \tau \geq 0. \\ \lim\nolimits_{\tau \longrightarrow \infty} y^f_{t,t+\tau} - y_t &= 0 \end{aligned} \tag{11}$$

Consequently, the long-run impact of the shock would be estimated to be equal to

$$\lim_{\tau \longrightarrow \infty} y^f_{t,t+\tau} - y_{t-1} = 0. \tag{12}$$

What would be the estimate of the long-run impact of the shock if one could observe $e_{z,t}$? In this case, one would get

$$\begin{aligned} \lim\nolimits_{\tau \longrightarrow \infty} \mathrm{E}_t \left[y_{t+\tau} \right] - y_{t-1} &= +1 \text{ if } e_{z,t} = +1 \text{ and} \\ \lim\nolimits_{\tau \longrightarrow \infty} \mathrm{E}_t \left[y_{t+\tau} \right] - y_{t-1} &= -1 \text{ if } e_{z,t} = -1 \end{aligned} \tag{13}$$

If one does not observe z_t, then one does not know whether $e_{z,t}$ is equal to -1 or $+1$ and one cannot do better then to predict that there is no long-term change in output. Not a great situation, but at least one is not systematically wrong. That is, half of the time one misses that there is a long-term decline in y_t and half of the time one misses that there is a long-term increase, but on average the prediction errors cancel out.

Now consider the case in which $e_{x,t} = e_{z,t} = -1$. Thus, $y_t - y_{t-1} = -\sigma_{e,x} - \sigma_{e,z} = -1.23607$. The long-run prediction according to the one-type-shock model is again that

[6] In Section 3.2, we discuss that the possibility forecasts are systematically biased, because the correct time series model for an aggregate could be a lot more complex than those allowed for with the class of models considered in Equation (1), but that is not an issue here.

output will not change, thus, the permanent impact of the shock would be estimated to be

$$\lim_{\tau \longrightarrow \infty} y^f_{t,t+\tau} - y_{t-1} = -1.23607. \tag{14}$$

But note that the true long-run impact of the shock is much smaller, namely

$$\lim_{\tau \longrightarrow \infty} \mathrm{E}_t\left[y_{t+\tau}\right] - y_{t-1} = -1. \tag{15}$$

That is, in a recession, the one-type-shock model systematically overpredicts the long-run impact of the shock. The simplicity of this example makes clear that one obviously could do better than the one-type-shock model if one would observe that $\Delta y_t = -1.23607$. If output drops, then it must be the case that $e_{x,t} = e_{z,t} = -1$. Consequently, it would be silly not to use this information. Note that there is still a sense in which the long-term prediction errors of the one-type-shock model average out. While one overestimates the negative long-run impact when $e_{x,t} = e_{z,t} = -1$, one also underestimates the positive long-run impact when $e_{x,t} = e_{z,t} = 1$. The problem with the one-type-shock *dgp* given in Equation (5) is that it does not take advantage of the fact that the magnitude of Δy_t actually has information about the values that $e_{x,t}$ and $e_{z,t}$ have taken on.[7]

To make the discussion as simple as possible, we made several simplifying assumptions. There is a much more general point to be made, however. Suppose that we now only assume that y_t is the sum of a transitory and a permanent component, a quite weak assumption, and that both components are linear processes. Even if the underlying processes are linear, then this does not mean that the univariate time-series process for the sum of them, i.e., y_t, is linear. The point being made in this subsection is that there could very well be information in the *magnitude* of the unexpected change in y_t. That is, if one observes a very large drop in y_t, then the chance is higher that $e_{x,t}$ and $e_{z,t}$ were both negative than that $e_{x,t}$ was positive and $e_{z,t}$ was extremely negative. That is, the larger the economic downturn the larger the probability that a certain fraction of this downturn is driven by the transitory shock, that is, the larger the probability that a fraction of the

[7] That is, the linearity of the *dgp* restricts the law of motion to be the same for small and large changes in y_t.

downturn will be recovered.[8]

3.2 Difficulty of correctly specifying aggregated processes

$ARMA$ processes are a standard way to model time series processes. The troublesome feature for aggregated processes is that the sum of two $ARMA$ processes is typically a more complex ARMA process. Granger and Morris (1976) show that if x_t is an $ARMA(p_x, q_x)$ and z_t is an $ARMA(p_z, q_z)$, then $y_t \equiv x_t + z_t$ is an $ARMA(p, q)$ with

$$p \leq p_x + p_z \text{ and } q \leq \max\{q_x + p_z, q_z + p_x\}. \tag{16}$$

The inequality could be strict, but unfortunately, the orders of the $ARMA$ representation of a sum of variables is typically higher than the highest orders of the $ARMA$ representations of the individual components, except if particular restrictions on the parameter values are satisfied. Granger (1980) argues that if the variable of interest is the sum of many components, as is the case for macroeconomic variables, then the process may even exhibit long memory.[9]

One might think that the solution to this dilemma is simply to use more complex $ARMA$ processes. The problem is that the processes have to be estimated with a finite amount of data, consequently the values of p and q cannot be too high. But if the values of p and/or q are too low, then the dgp is misspecified.[10] That such misspecification can easily lead to very bad long-term forecasts is documented by the following example.

---

[8] Key is that the components, x_t and z_t, are linear processes. If one would start with non-linear processes for x_t and z_t, then anything could happen, but the magnitude of the change in y_t is still likely to be informative.

[9] One aspect that is ignored in this literature is that the *dgp*s of the individual components may be "aligned". For example, if markets are complete then market prices will align agents' marginal rates of sustitution even if agents face very different income processes.

[10] The misspecification is likely to be worse than indicated in this section. Typically, log-linear processes are more suitable than linear processes. But if $y_t \equiv x_t + z_t$ and x_t and z_t are log-linear processes, then neither y_t nor $\ln(y_t)$ is a linear process and the convention of modelling $\ln(y_t)$ as a linear process is, thus, not correct. In fact, the effects of shocks on y_t would be time-varying. These issues are further discussed in den Haan, Sumner, and Yamshiro (2009).

Consider the following *dgp*:

$$
\begin{aligned}
y_t &\equiv x_t + z_t \\
x_t &= \lambda_x x_{t-1} + e_{x,t} \\
z_t &= \rho_{z,1} z_{t-1} + \rho_{z,2} z_{t-2} + e_{z,t} \\
\mathrm{E}_t\left[e_{x,t}\right] &= \mathrm{E}_t\left[e_{z,t}\right] = 0 \\
\mathrm{E}_t\left[e_{x,t}^2\right] &= \sigma_{e,x}^2 \\
\mathrm{E}_t\left[e_{z,t}^2\right] &= \sigma_{e,z}^2 \\
\mathrm{E}_t\left[e_{x,t} e_{z,t}\right] &= 0
\end{aligned}
\tag{17}
$$

It will be useful to rewrite the law of motion for z_t as follows:

$$(1-\lambda_{z,1}L)(1-\lambda_{z,2}L)z_t = e_{z,t} \tag{18}$$

with $\lambda_{z,1}+\lambda_{z,2}=\rho_{z,1}$, $\lambda_{z,1}\lambda_{z,2}=-\rho_{z,2}$, and without loss of generality we assume that $|\lambda_{z,1}| \geq |\lambda_{z,2}|$. The correct univariate *dgp* for y_t is in general an $ARMA(3,2)$. The dominant autoregressive root of this $ARMA(3,2)$ specification is equal to

$$\lambda_y^{\max} = \max\{|\lambda_x|, |\lambda_{z,1}|\}. \tag{19}$$

Thus, this univariate $ARMA(3,2)$ process would predict that a shock to y_t vanishes according to $(\lambda_y^{\max})^t$, while y_t actually vanishes according to $(\lambda_x)^t$ if y_t changes because of a shock in $e_{x,t}$ and vanishes according to $(\lambda_{z,1})^t$ if y_t changes because of a shock in $e_{z,t}$. That is, the long-run impact according to the *correct* univariate specification of y_t is only correct for one type of shock.

We will now show that the long-run impact even can be misleading following both types of shocks if one misses the complexity of the higher-order process. In particular, suppose that one misses the complexity and estimates an $AR(1)$.[11] It is straightforward to show that asymptotically the value of the $AR(1)$ coefficient is given by

$$\hat{\lambda}_y = \lambda_x \frac{\sigma_x^2}{\sigma_x^2+\sigma_z^2} + \frac{\lambda_{z,1}+\lambda_{z,2}}{1+\lambda_{z,1}\lambda_{z,2}} \frac{\sigma_z^2}{\sigma_x^2+\sigma_z^2}, \tag{20}$$

[11] It is unlikely that one is so constrained in the number of the degrees of freedom that one can estimate an $AR(1)$ and cannot estimate an $AR(2)$. Our point about misspecification carries over to higher-order processes where degrees of freedom are a constraint, but is easier to illustrate for lower-order processes.

where σ_x and σ_z are the standard deviations of x_t and z_t, respectively. Now suppose for simplicity that $\lambda_x = \lambda_{z,1} > \lambda_{z,2}$ and $\sigma_x = \sigma_z$. That is, the dominating root of z_t is assumed to be equal to the root for x_t. This implies that in the long run the effects of both $e_{x,t}$ and $e_{z,t}$ vanish according to $(\lambda_x)^t$. In this particular case, the correctly specified univariate *dgp* of y_t would predict the same. But according to the misspecified $AR(1)$ process for y_t the effects of a shock vanishes according to $\left(\hat{\lambda}_y\right)^t$ with

$$\hat{\lambda}_y > \lambda_x$$

as long as $\lambda_{z_2} > 0$.

Thus, if a variable y_t is the sum of a set of other variables, then it would make more sense to directly model the laws of motion for the components of y_t and obtain forecasts for y_t by *explicitly* aggregating the forecasts for the components.

Figure 1 illustrates graphically the long-term consequences of misspecification. First, consider the true responses of y_t to shocks in $e_{x,t}$ and $e_{z,t}$.[12] The short-term responses are very different, but in the long run the effects of the shock disappear at the same rate (because the dominating root was chosen to be the same for x_t and z_t). The figure also plots the response of y_t to a shock according to the one-type-shock process. The one-type-shock $AR(1)$ process grossly overpredicts the persistence of the shocks. The reason is that the one-type-shock process tries to fit both the hump-shaped short-run response to a shock in $e_{z,t}$ and the mean-reverting short-run response to a shock in $e_{x,t}$. Without sufficient complexity in the estimated *dgp* for y_t, this leads to an overestimation of the long-run effects of the shock.

In this example, the dynamics of the components are still very simple and the solution would be to estimate an $ARMA(3,2)$. But if the two components are, for example, both an $AR(4)$, which is not implausible, one would have to estimate an $ARMA(8,4)$, and if y_t is the sum of three $AR(4)$ processes, then one would have to estimate an $AR(12,8)$. In the next section, we document that a better strategy is to estimate separate *dgp*s for the components and than explicitly aggregate the forecasts of the components to obtain forecasts for the aggregated variables.

[12] The parameter values are as follows: $\lambda_x = \lambda_{z,1} = 0.85$, $\lambda_{z,2} = 0.6$, and $\sigma_{e,x} = \sigma_{e,z} = 1$.

4 Documenting the advantages of additional information

4.1 Empirical specifications

The purpose of this section is *not* to construct the best possible forecasting model; the point is to document that time series models that only have one type of shock (or one type of financial crisis) are inadequate and tend to predict that shocks have very persistent effects even when the impact of the shock is short lived and that models that allow for more than one type of shock also allow different types of recovery patterns. The following *dgp* is estimated:

$$\ln(y_t) = a_0 + \sum_{j=1}^{4} a_j \ln(y_{t-j}) + e_t, \tag{21}$$

where $\mathrm{E}_t\left[e_{t+1}\right] = 0$. The key feature of this law of motion is that there is only one type of shock. Thus, if output turns out to be unexpectedly lower than expected, i.e., $e_t < 0$, then the predicted effect on y_t is *always* the same.

The specification of the components model is given by

$$\ln(s_t) = b_0 + \sum_{j=1}^{4} B_j \ln(s_{t-j}) + e_{s,t}, \tag{22}$$

where s_t is a 5×1 vector containing the expenditure components, that is,

$$s_t' = [c_t, i_t, g_t, x_t, m_t]$$

and $\mathrm{E}_t[e_{s,t+1}] = 0$, for $s \in \{c, i, g, x, m\}$. The forecast for $y_{t+\tau}$ follows directly from

$$y_{t+\tau} \equiv e^{\ln(c_{t+\tau})} + e^{\ln(i_{t+\tau})} + e^{\ln(g_{t+\tau})} + e^{\ln(x_{t+\tau})} - e^{\ln(m_{t+\tau})}, \tag{23}$$

where we make explicit that the forecasts for the components are for the logs.

Both time series processes are estimated with ordinary least squares (OLS). Given that the variables could very well be integrated, it is important to add enough lags to ensure that the shocks are stationary and spurious regression results are avoided. If the series are known to be (co)integrated, then efficiency gains are possible by imposing the implied restrictions. *If* these restrictions are correct, then the estimated parameter values in an unrestricted system that does not impose them will converge towards these restrictions at rate T, that is, there is superconsistency. If the restrictions are not correct and are nevertheless imposed, then the system is misspecified and the estimated system will typically not converge towards the correct answer. Because of the superconsistency, we prefer not to impose these types of restrictions on the system and let the data decide.

4.2 Impulse response functions

The impact of a negative one-standard-deviation shock to e_t on (the log of) output, i.e., the impulse response function (IRF), is displayed in Figure 2.[13] Even though the specification in Equation (21) does not impose a unit root, the estimated specification documents that the impact to the shock e_t is very persistent.[14] It is exactly this type of result that underlies the argument of Greg Mankiw that one should expect the current crisis to have permanent effects.

If output is generated by the components model, i.e., according to Equations (22) and (23), then there are five shocks and consequently, there are five IRFs, that is, five different ways in which output could respond. There are fierce debates in the economic literature on how to interpret shocks, but for our purposes the interpretation of the shocks is not important. For our purposes it is sufficient to document that this richer model allows for very different types of responses to output. For convenience, we will label the shocks according to the variable being estimated. For example, we will refer to $e_{c,t}$ as the consumption shock, but this is just a label and not even meant to hint at a structural interpretation.

The five IRFs are plotted in Figure 3. The figure makes clear that according to the components model there are also shocks that have a very persistent impact on output. The figure also makes clear, however, that there are shocks that have a transitory impact on output. That is, the figure clearly illustrates two key points of this paper: (i) models that allow for only one type of shock (or one type of financial crisis) will not discover that the long-term impact is not the same for each type of shock and (ii) that a slightly richer

[13] The results in this subsection are based on quarterly U.S. data from 1947Q1 to 2009Q2. See Appendix A for further details on data sources.

[14] The impact of the shock remains persistent if a deterministic time trend is included in the specification.

model does discover this difference.

4.3 Forecasting recoveries from past recessions

The analysis above showed that—in contrast to what is suggested by the IRF of the one-type-shock *dgp* of Equation (21)—the dynamics of U.S. GDP according to the components model are rich and diverse. The question arises whether this actually matters for how the U.S. economy recovers from a recession. After all, it may very well be the case that the shocks that got the U.S. economy into a recession were those that indeed do have a very persistent impact on the economy.

This is the question we address in this subsection. The findings of this section can be summarized as follows.

1. Most economic downturns did indeed lead to a downward adjustment of the expected long-term growth path (relative to the most optimistic growth path recently expected).

2. This downward adjustment is smaller for the components model than for the one-type-shock model *and* the components models predicts a much faster convergence towards this new long-term growth path.

3. There is only one recession where the one-type-shock model outperforms the components model, which is the 2001 recession.

Using "real-time" data. We use the two time series models to make predictions about future output at the troughs of NBER recessions, t_{trough}. These predictions are compared with the predictions made at $t_{\text{optimistic}}$ where $t_{\text{optimistic}}$ is an earlier date, i.e., $t_{\text{optimistic}} < t_{\text{trough}}$, at which the long-term predictions were most optimistic.

At each forecasting point, t_{trough}, we only use data up to date t_{trough}. It is a bit of abuse of terminology to refer to this as real-time data, because we use data up to date t_{trough} as it is available now, that is, we ignore that at date t_{trough}, econometricians would not have access to the revised data in our sample.[15]

By using real-time data, we have to limit the number of recessions we can look at, because we do not have enough data to estimate a sensible model for the earlier recessions.[16] The first recession considered is the 1973-75 recession for which we can make forecasts based on a model estimated with 109 quarterly observations.

Explaining the figures. For each recession, we have one figure with two panels. The top panel plots the results for the components model and the bottom panel the results for the model with only one type of shock. There are two vertical lines indicating the two forecasting points, with the trough of the recession corresponding to the one most to the right. The thick solid line plots the actual data. Each panel also plots the predicted growth path for the two forecasting dates. Note that these coincide with the actual data at the forecasting point (the vertical line). To facilitate the comparison between the forecasts of the two models, the bottom panel also repeats the forecast according to the components model.

1973-75 recession. The results for the 1973-75 recession are shown in Figure 4. According to both time series models, the forecasts made in 1975Q1, the trough of the recession, were substantially below those in 1973Q2, when forecasts were most optimistic. At this trough date, the forecasts of the one-type-shock model are much more pessimistic than those of the components model. The one-type-shock model basically predicts that none of the loss in GDP will be made up. In contrast, the components model predicts several periods of high growth after which the growth rate is predicted to level off, a pattern basically followed by actual GDP.

[15] Being able to use revised data may very well improve the quality of the forecasts, but is unlikely to affect the comparison of the two different types of time series models.

[16] For the earlier recessions, we compared the performance of the two models when both were estimated using the full sample. The results for these earlier recessions are if anything even stronger than those for the recessions discussed below: We find that the predictions of the one-type-shock model at the trough dates are too pessimistic and that the recoveries predicted by the components model are much better.

1980 recession. The results for the 1980 recession are shown in Figure 5. They resemble the results of the 1973-75 recession. Again the predictions of the one-type-shock are more pessimistic than those of the components model, but now even the predictions of the components model are too pessimistic, that is, actual GDP performed better than what was expected, even by the less pessimistic predictions of the components model (ignoring, of course, the temporary downturn of the 1981-82 recession).

1981-82 recession. As documented in Figure 6, this recession did not lead to a substantial downward revision of long-term output predictions according to the components model, a prediction that turned out to be confirmed by actual GDP realizations. In contrast, the one-type-shock model interprets this transitory shock as a permanent one, leading to a substantial reduction in long-term output forecasts, a prediction which turned out to be way too pessimistic.

1990-91 recession. As documented in Figure 7, the forecasts of the two models are very similar for this recession. Both models downgrade their long-term output forecasts, a revision that turned out to be not justified. The long-term predictions that were made before the crisis hit turned out to be a better forecast of *long-term* output developments than those made at the end of the recession.

2001 recession. As documented in Figure 8, the forecasts of the two models made at the end of this short-lived recession differ substantially. The components model interprets the downturn in large part as a temporary shock and the long-run growth path is not revised downward by much. For this recession, this turned out to be a mistake. The high growth rates of the second half of the nineties were not repeated in the first decade of the new millennium. By interpreting the downturn as permanent, the one-type-shock does a better job. However, the components models quickly realizes it made a mistake. Figure 9 plots the forecasts one year after the trough, i.e., 2002Q4. The components model has at this point revised its long-term forecasts for output considerably and its forecasts are quite good up to the beginning of the recent recession.

Current recession. Figure 10 gives the forecasts of the two models made using data up to the second quarter of 2009. The differences in the forecasts are striking. The one-type-shock model predicts a much larger permanent effect of the crisis. Relative to the forecasts made at the third quarter of 2007, the one-type-shock model reduces its forecast for output in 2014 down with 10%, whereas the components model does with only 6%. These forecasts are meant to illustrate the difference between the one-type-shock model and the richer model. We want to stress that we do not present these as the best possible forecasts of the effects of the current crisis. First note that the downward revisions are measured relative to the point at which the long-term forecasts were most optimistic. Second, it is easy to come up with specifications in which the long-term impact of the current recession is much smaller, but we always find that there remains a substantial difference between the two time series models. An easy modification is to add a linear deterministic time trend to the two time series models. The results are reported in Figure 11. The one-type-shock model still predicts a substantial permanent effect of the current crisis. In particular, the recent reductions in output led to a reduction of the forecast of output for 2014Q1 (2019Q2) of 7.2% (6.0%) according to the one-type-shock model and to a reduction of only 2.7% (1.4%) according to the components model.

5 Financial crises and current forecasting practice

The time series properties of U.S. post-war data may not be representative for how the U.S. will recover from the current downturn, because none of the post-war recessions were financial crises. Several researchers have argued that it would, thus, be better to look at the consequences of past financial crises to predict recovery, or the lack of it, from the current one. The problem of this approach is, of course, that to have enough observations one has to include financial crises that occurred in a wide range of different countries and/or a long time ago. Moreover, there are many types of financial crises, so it is not clear whether these events are more representative for the current crisis than post-war recessions, but it is definitely worth considering them.

A good example is the analysis of Cerra and Saxena (2008), who estimate the following time series model:

$$g_{i,t} = a_i + \sum_{j=1}^{4} \beta_j g_{i,t-j} + \sum_{s=0}^{4} \delta_j D_{i,t-s} + \varepsilon_{it}, \tag{24}$$

where g_t is the percentage change in real GDP, $\Delta \ln y_{i,t}$, and $D_{i,t}$ is a dummy indicating a financial (or political) crisis. The authors assume that $\ln y_{i,t}$ has a unit root and, thus, impose that crises *always* have a permanent effect. But it is, of course, silly to impose this. The motivation is typically that GDP is known to have a unit root, but in Section 3 we saw that even if a variable has a unit root, then this does not imply that all shocks have permanent effects. Therefore, we focus on the less restrictive version of Equation (24), which is

$$\ln y_{i,t} = a_i + \sum_{j=1}^{5} \beta_j \ln y_{i,t-j} + \sum_{s=0}^{4} \delta_j D_{i,t-s} + \varepsilon_{it}. \tag{25}$$

The drawbacks of forecasts based on this or similar specifications are the following:

1. Most importantly, even the less restrictive specification given in Equation (25) is subject to the *exact* same problems as those discussed in Section 3.

   (a) Because one tries to predict an aggregate variable, $y_{i,t}$, one has to deal with the complexity of GDP induced by aggregation. This would be especially important for predicting the long-term consequences of shocks.

   (b) More important is that the one-type-of-financial crisis model like the one-type-of-shock model is not capable of distinguishing between events with different effects on long-term growth.

To see this, suppose that the law of motion for output, y_t, is now given by

$$\begin{aligned} y_t &\equiv y_t^f + y_t^{nf} \\ y_t^f &\equiv x_t + z_t \\ x_t &= \rho_x x_{t-1} + \sigma_{e,x} e_{x,t} \\ \Delta z_t &= \rho_z \Delta z_{t-1} + \sigma_{e,z} e_{z,t} \\ \mathrm{E}_t\left[e_{x,t}\right] &= \mathrm{E}_t\left[e_{z,t}\right] = 0 \\ \mathrm{E}_t\left[e_{x,t}^2\right] &= 1 \\ \mathrm{E}_t\left[e_{z,t}^2\right] &= 1 \\ \mathrm{E}_t\left[e_{x,t} e_{z,t}\right] &= 0 \\ |\rho_x| \leq 1 &\text{ and } |\rho_z| < 1 \end{aligned} \tag{26}$$

That is, output consists of two components, a financial crisis component, y_t^f, and a non-crisis component, y_t^{nf}. The law of motion of the non-crisis component is irrelevant for the discussion here and is left unspecified. Note that the financial crisis component consists of an $I(0)$ and an $I(1)$ component. Thus, a specification as in Equation (25) will always predict that a financial crisis has permanent effects even though there are financial crises, namely those triggered by shocks in $e_{x,t}$, that do not have permanent effects. As shown in Section 3, it is even possible that y_t^f is a random walk, even though both its components are forecastable.

For past financial crises, it may be difficult to obtain a rich data set, but for every additional variable added to the analysis, one allows for more types of financial crisis. In particular, suppose that in addition to $y_{i,t}$ one has data on $s_{i,t}$, where $s_{i,t}$ is an $n_s \times 1$ vector with $n_s \geq 1$. Then one could estimate

$$\begin{aligned} \begin{bmatrix} y_{i,t} \\ s_{i,t} \end{bmatrix} &= A(L) \begin{bmatrix} y_{i,t-1} \\ s_{i,t-1} \end{bmatrix} + B(L) D_{i,t} \begin{bmatrix} y_{i,t-1} \\ s_{i,t-1} \end{bmatrix} \\ &+ \begin{bmatrix} u_{y,t} \\ u_{s,t} \end{bmatrix} (1 - D_{i,t}) + \begin{bmatrix} \tilde{u}_{y,t} \\ \tilde{u}_{s,t} \end{bmatrix} D_{i,t}, \end{aligned} \tag{27}$$

where $D_{i,t} = 1$ if country i is in a crisis in period t. Note that this system has

$2 \times (n_s + 1)$ IRFs for $y_{i,t}$, namely $(n_s + 1)$ IRFs for shocks occurring during normal times ($D_{i,t} = 0$) and $(n_s + 1)$ financial crises IRFs.

2. This specification assumes that the occurrence of a financial crisis, i.e., $D_{i,t} = 1$, is exogenous. But a crisis is, of course, much more likely to occur if growth prospects are poor. That is, even if without the current financial crisis there would most likely have been a correction in house prices that would have slowed down economic growth. More formally, negative estimates for δ_j could capture that the crisis has a negative impact on future values of $y_{i,t}$, but also can imply that negative *expectations* about future values of $y_{i,t}$ cause the crisis. Dealing with the endogeneity of $D_{i,t}$ is a horrendously difficult task.[17] The good news for the current financial crisis is that the downturn seems to have been caused almost fully by the financial crisis in the sense that beside aspects like excessive risk taking in the financial sector, the economy was healthy in most other aspects. That is, the financial crisis was not triggered by knowledge that the economy itself had some serious problems, which—without doubt—played a role in several financial crises.

Section 6 contains a general discussion on forecasting during times of crisis, but here we give two easy ways in which papers that focus on past crises could be improved upon.

1. It may not be possible to obtain the components of GDP for a sufficiently large set of past financial crises, that is, one may have no choice but to use GDP. But this is no excuse not to add additional variables. In particular, it would be worthwhile to add employment and/or capacity utilization to the analysis and construct a model that predicts these variables *jointly.* As long as the variables included respond differently to different types of shocks, then one is likely to quickly discover the nature of the financial crisis. For example, if output drops because employment or capacity utilization drop, then the long-term consequences are likely to be less severe, than when output drops because of a drop in productivity.

[17] Cerra and Saxena (2008) propose to deal with the endogeneity by estimating an equation describing the probability a crisis occurs, but the information set is so limited that it is unlikely that they accurately capture the link between knowledge about future movements in $y_{i,t}$ and the occurence of the crisis.

2. Even if one only had GDP data, then there is one exercise that should be included in any analysis based on time series models like the one given in Equations (5) and (6). This exercise is a comparison of the long-term impact of non-crisis shocks with the long-term impact of crisis shocks. Note that the specification of Equations (5) and (6) impose that the effects of a regular and a crisis shock are the same, but it is easy to make the specification more flexible to allow for different dynamic effects. In particular, it would be worth comparing the persistence of the effect of a regular shock to that of a financial crisis shock. Whereas the literature argues that the impact of financial crises on output found in other countries is representative for the current situation, it would make more sense to argue that only the *additional* persistence found after a financial crisis shock (compared with the persistence of a regular shock) is similar across countries. That is, if one wants to argue that the impact of the current financial crisis will have a longer lasting impact on economic growth than post-war recessions did on subsequent economic growth, then one should show that this *difference* in persistence was present in countries included in the financial crisis data set. Just showing that financial crises seem to have a persistent impact is not enough. We know that one-type-shock models based on U.S. data also predict that all recessions should have permanent effects, a result that is refuted by richer models that realize that there are also recessions with a strong transitory component.

Predictions of van Ewijk and Teulings (2009) Figure 4.2 in van Ewijk and Teulings (2009) provides estimates of the long-term effects of a financial crisis. They find very persistent effects. Given the importance of such forecasts for policy makers, they have received quite a bit of attention in the Dutch media, so it is worth discussing them in a bit more detail. The analysis of van Ewijk and Teulings (2009) is based on the data set of Caprio and Klingebiel (2003), that only includes relatively recent crises (since the late 70s). A questionable aspect of their analysis is that it seems to be based on numerous developing and newly industrialized countries. The full analysis underlying the depressing predictions of van Ewijk and Teulings (2009) has not yet been made available. For example, it is not clear which exact set of countries of the data set of Caprio and Klingebiel (2003) are included, and in particular whether it includes countries like say Bangladesh, Burkina Faso, Eritrea, Latvia, Niger, Senegal, or Romania.[18]

Although the full analysis is not yet in the public domain, the book's website provides an informal discussion making clear that a somewhat more general specification is used than the one given in Equation (25). This more general specification allows the β coefficients to depend on the occurrence of a crisis. In particular, they use

$$y_{i,t+\tau} = \sum_{k=1}^{8} \beta_k^{(\tau)} y_{i,t-k} + \delta^{(\tau)} b_{i,t} + e_{i,t}, \tag{28}$$

that is, the time path of $y_{i,t}$ during a regular downturn, i.e., one that is initiated by a negative shock to $\varepsilon_{i,t} \neq 0$, is allowed to differ from the time path when the downturn is initiated by a financial crisis, i.e., by a change in $D_{i,t}$. Moreover, this specification follows Jorda (2005) by letting the β coefficients depend on the forecast horizon. This last generalization could possibly alleviate the difficulty of ensuring that the process is complex enough given that we are trying to forecast an aggregated variable.

This generalization, however, does not avoid the key problems discussed in this paper. In terms of Equation (26), this generalization just means that one can disentangle the crisis component, y_t^f, and the non-crisis component, y_t^{nf}, and one can specify two IRFs: one for an $\varepsilon_{i,t}$ shock and one for a financial crisis shock. If y_t^f consists of stationary and permanent components, however, then there are at least two financial crisis IRFs and the more general specification given in Equation (28) would not be able to disentangle these, whereas the analysis in IMF (2009) clearly demonstrates that not all financial crises have permanent effects. That is, being able to condition on there being a financial crisis does

[18] To see whether the set of included countries can reasonably be expected to respond in a way that is representative for what will happen to countries like the Netherlands in the current situation, the authors could do the following. Instead of a dummy for a financial crisis, they could include a dummy for the oil crisis of the seventies. They could compare whether their set of included countries (which excludes the Netherlands) responded to the oil crisis like the Netherlands did. If they find responses to the oil crisis that are more severe than the experience in the Netherlands, then it seems likely that this would happen again, given that structurally the Netherlands was in much better shape before the current than before the seventies oil crisis.

not generate multiple IRFs for different types of financial crises. But if one would observe just one more variable, then one can already distinguish between two types of financial crises, even at the beginning of the crisis.[19] Moreover, since the specification in Equation (28) is still linear in y_t, it does not use information about the magnitudes of the shocks. That is, if output is generated during a financial crisis by a transitory and a permanent component, as in Equation (6), then one would systematically overpredict the impact of a severe and underpredict the impact of a mild crisis as explained in Section 3.1.3.

Finally, the empirical analysis in van Ewijk and Teulings (2009) completely seems to ignore the role of macroeconomic policies being implemented, both in terms of fiscal and monetary stimulus and in terms of governments preventing major banks from collapsing. Is it the view of the CPB that these are not useful? In contrast, the analysis in IMF (2009) does find a role for government policy. Given that there are so many numerous weak aspect to the analysis of van Ewijk and Teulings (2009) one wonders whether it would not have been more sensible to first write down a careful academic paper and invite a serious discussion with the academic community before mobilizing the academic personnel of the CPB and even a journalist to write a popular book to convince the general public of this gloomy outlook.

6 How to make forecasts in the current situation?

The main contribution of this paper consists of documenting the inadequacy of the current practice of making forecasts based on one-type-shock and/or one-type-of-financial-crisis models. Our components model is mainly meant to document the lack of flexibility of the one-type-shock models, not as the best possible forecast.

In the last section of this paper, we stress that a serious forecasting exercise can never simply be based on a time series model, especially not in the current situation. Instead, we argue that a responsible forecasting exercise should consist of the following elements.

- **Empirical component I - time series models:** Time series models estimated

[19] Unless, of course, the additional variable carries the exact same information as output.

with past data are a necessary ingredient to make quantitative forecasts, because economic theories are currently too abstract to produce precise quantitative forecasts. The key lesson we learned from our analysis is that one *cannot* rely on a limited set of variables. It does *not* make sense to simply argue that because output has dropped substantially and we are in a financial crisis, that the best prediction of future output growth is the average post-financial-crisis growth path observed in a wide range of countries; the countries involved are very different from each other and the financial crises differ in key aspects of the current economic crisis. Before using the experience of these other financial crises as the basis for making predictions about the recovery out of the current crisis, one has to make sure that many more variables than just output behave now the way they did in those other crises. Employment, capacity utilization, and productivity are obvious candidates for which data is often available. Ideally one also should include financial variables, like a variable that measures the extent of bankruptcies in the financial sector. The big question after a financial crisis is whether the financial sector will be able to function as before and channel resources to the most productive resources. The numerous bail outs of banks may be bad for reasons of moral hazard, but they ensured that many networks were kept in tact, which should be very helpful. Besides the number of bank failures, bank profits could be an important variable to consider. The fact that several banks are already making profits could very well distinguish this crisis from others.

- **Theoretical component:** On a daily basis, one can find claims by economists on how the economy will develop and what type of (government) action is beneficial or harmful. We are stunned by the fact that even academics are willing to make so many bold claims without ever making clear what the basis of their arguments is. Any academic that takes a stand on how the economy will develop or what the government should do, should have the responsibility to make clear, for example, on his web site, what empirical or theoretical arguments were used in forming her/his judgement. This is also important for forecasting. There are some theories that predict that (large) shocks can have permanent effects. In fact, one of us has worked on such theories. By being explicit about the theories one has in mind, one can start a discussion on whether the assumptions of the theory are currently applicable. For example, in Den Haan, Ramey, and Watson (2003) a model is developed in which a large enough shock could damage the financial sector so much that the economy collapsed, a situation out of which the market cannot by itself recover. The collapse occurs when there is a downward spiral in which investors provide less liquidity which deteriorates the health of the financial sector, which in turn discourages investors to provide liquidity. According to this model, financial crises could have permanent effects. But the model also suggests that there is an important role for the government and that a large supply of liquidity by the government can stop the downward spiral. So although the drastic implications of the model of Den Haan, Ramey, and Watson (2003) may very well be applicable to some past financial crises, the government seems to have done exactly the right thing to prevent the current financial crisis to have permanent effects through the channel emphasized by Den Haan, Ramey, and Watson (2003). den Haan (2007) is another example of a model in which a large negative shock can have permanent effects. This model explains why European unemployment remained high for so long after the recession of the seventies, whereas U.S. unemployment recovered relatively quickly. The idea of this model is that the shock pushed the European economy towards a "low-activity" equilibrium in which high unemployment levels lead to higher government transfers and other related government expenses, which in turn lead to higher tax rates, which in turn reduces the demand for labor, which in turn leads to high unemployment levels. In the U.S., such a low-activity equilibrium does not exist, because unemployment benefits are much lower. For this story to work, there has to be a large enough mass of marginal jobs for which relatively small changes in the tax rate can make all the difference. Possibly, this story was relevant for the seventies, but it seems unlikely that there was such a large mass of marginal jobs before the start of the current crisis in many European countries.[20]

A serious forecasting procedure would consist of this type of reviews of theories and evaluates their relevance. Often this requires additional information, for example, on whether there are many jobs that are right at the point of being unprofitable. This brings us to the next ingredient.

- **Empirical component II - detailed information:** Understanding which economic theory is currently the most relevant, requires detailed information about what is going on *right now*. For example, the main theory that predicts that government spending is beneficial in a crisis relies on rigid prices and/or wages.[21] There is ample evidence that there is some rigidity in prices and wages during normal times, but are prices and wages currently flexible? And if one finds that they are more flexible than normal—which anecdotal evidence suggests—are they flexible enough? Similarly, it would be important to know what limits investments? Is it lack of demand, lack of financing, or the possibility of higher taxes in the future? Money spent on additional surveys could now be worthwhile.

Because of their access to excellent data and experience in forecasting, government agencies with research departments should play a key role in performing these tasks. It goes without saying that during times like the current, there should be coordination and intense interaction between the research departments of government agencies. Moreover, it is essential that there is interaction with university economics departments. Every central bank president and every director of any other type of government agencies with a research group should be held responsible for what their research group has done in terms of helping us to understand the crisis, in terms of developing models to forecast its consequences, and in terms of designing well-motivated policies to alleviate the damage. This is not the right time for research departments to spend their time working with linearizations around the steady state and fine tuning optimal monetary policies, luxury problems for which the

[20] Of course, if another mechanism creates such a large mass of marginal jobs, then this channel could become relevant.

[21] There are others, for example, theories that rely on externalities,

value added was already rapidly approaching zero before the crisis started.

A Data sources

Data are downloaded from the web site of the Federal Reserve Bank of St. Louis. Details of the downloaded series are as follows.

- Consumption: Real Personal Consumption Expenditures
  - Series ID: PCECC96
  - Source: U.S. Department of Commerce: Bureau of Economic Analysis
  - Seasonal Adjustment: Seasonally Adjusted Annual Rate
  - Frequency: Quarterly
  - Units: Billions of Chained 2005 Dollars
  - Last updated: 2009-08-27 11:02 AM CDT
- Investment: Real Gross Private Domestic Investment
  - Series ID: GPDIC96
  - Source: U.S. Department of Commerce: Bureau of Economic Analysis
  - Seasonal Adjustment: Seasonally Adjusted Annual Rate
  - Frequency: Quarterly
  - Units: Billions of Chained 2005 Dollars
  - Last updated: 2009-08-27 10:34 AM CDT
- Government expenditures: Real Government Consumption Expenditures & Gross Investment
  - Series ID: GCEC96
  - Source: U.S. Department of Commerce: Bureau of Economic Analysis

- Seasonal Adjustment: Seasonally Adjusted Annual Rate
- Frequency: Quarterly
- Units: Billions of Chained 2005 Dollars
- Last updated: 2009-08-27 10:34 AM CDT

- EXPORTS: Real Exports of Goods & Services
  - Series ID: EXPGSC96
  - Source: U.S. Department of Commerce: Bureau of Economic Analysis
  - Seasonal Adjustment: Seasonally Adjusted Annual Rate
  - Frequency: Quarterly
  - Units: Billions of Chained 2005 Dollars
  - Last updated: 2009-08-27 10:34 AM CDT

- IMPORTS: Real Imports of Goods & Services
  - Series ID: IMPGSC96
  - Source: U.S. Department of Commerce: Bureau of Economic Analysis
  - Seasonal Adjustment: Seasonally Adjusted Annual Rate
  - Frequency: Quarterly
  - Units: Billions of Chained 2005 Dollars
  - Last updated: 2009-08-27 10:34 AM CDT

- GDP. The GDP series used is the sum of the consumption, investment, government expenditures, and exports series minus the imports series. Adding up these real series generates a series that is extremely close, but not exactly identical to the real GDP series. By using this GDP series we avoid clutter in the paper due to small differences in the series used in the two types of time series models. Now the GDP series used is identical for the two time series models.

References

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2009): "News, Noise, and Fluctuations: An Empirical Exploration," NBER working paper No. 15015.

Campbell, J. Y., and N. G. Mankiw (1987): "Are Output Fluctuations Transitory," *Quarterly Journal of Economics*, 102, 857–880.

Caprio, G. J., and D. Klingebiel (2003): "Episodes of Systemic and Borderline Financial Crises," Unpublished manuscript, Williams College.

CBO (2009): "The Budget and Economic Outlook: Fiscal Years 2009 to 2019," Statement of Robert A. Sunshine, acting director of the Congressional Budget Office, before the Committee on the Budget, United States Senate, January 8, 2009.

Cerra, V., and S. C. Saxena (2008): "Growth Dynamics: The Myth of Economic Recovery," *American Economic Review*, 98, 439–457.

den Haan, W. J. (2007): "Shocks and the Unavoidable Road to Higher Taxes and Higher Unemployment," *Review of Economic Dynamics*, 103, 348–366.

Den Haan, W. J., G. Ramey, and J. Watson (2003): "Liquidity Flows and Fragility of Business Enterprises," *Journal of Monetary Economics*, 50, 1215–1241.

den Haan, W. J., S. W. Sumner, and G. Yamshiro (2009): "Bank Loan Components and the Time-Varying Effects of Monetary Policy Shocks," Unpublished manucript, University of Amsterdam.

Granger, C., and M. Morris (1976): "Time Series Modelling and Interpretation," *Journal of the Royal Statistics Society A*, 139, 246–257.

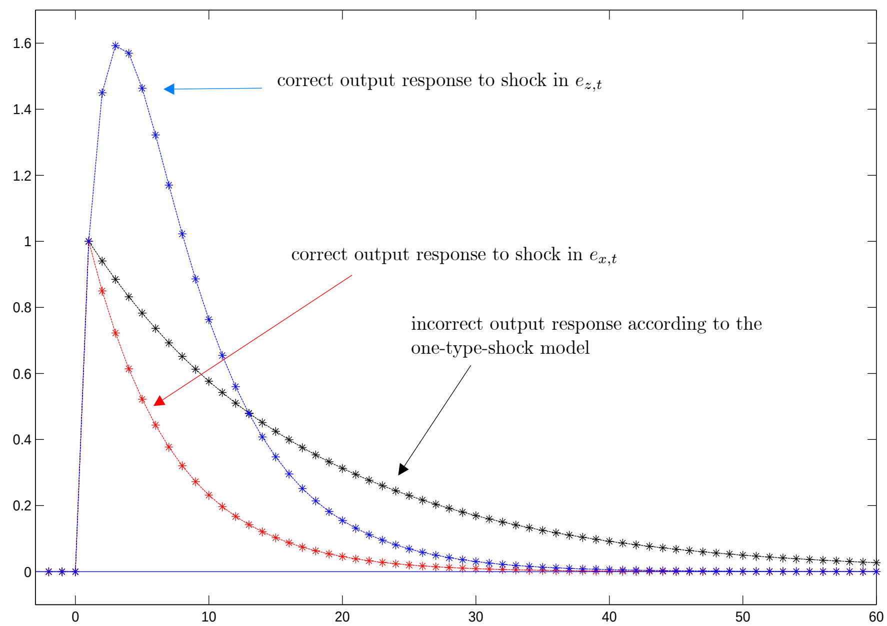
Granger, C. W. (1980): "Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, 14, 227–238.

IMF (2009): "World Economic Outlook," October 2009, International Monetary Fund.

Jorda, O. (2005): "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 95, 161–182.

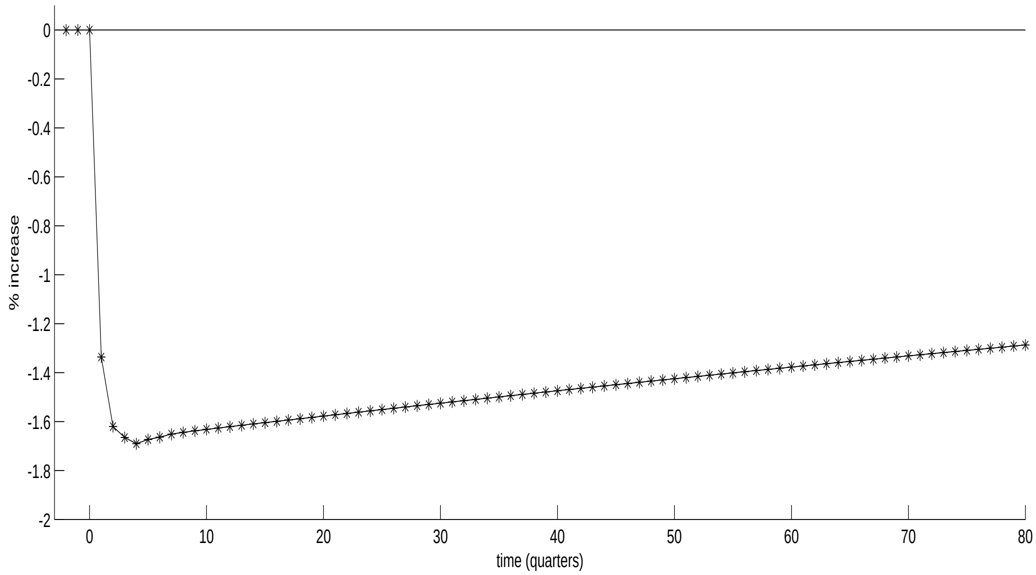
van Ewijk, C., and C. Teulings (2009): *De Grote Recessie*. Balans, Amsterdam.

Figure 1: Correct IRFs and IRF of one-type-shock model



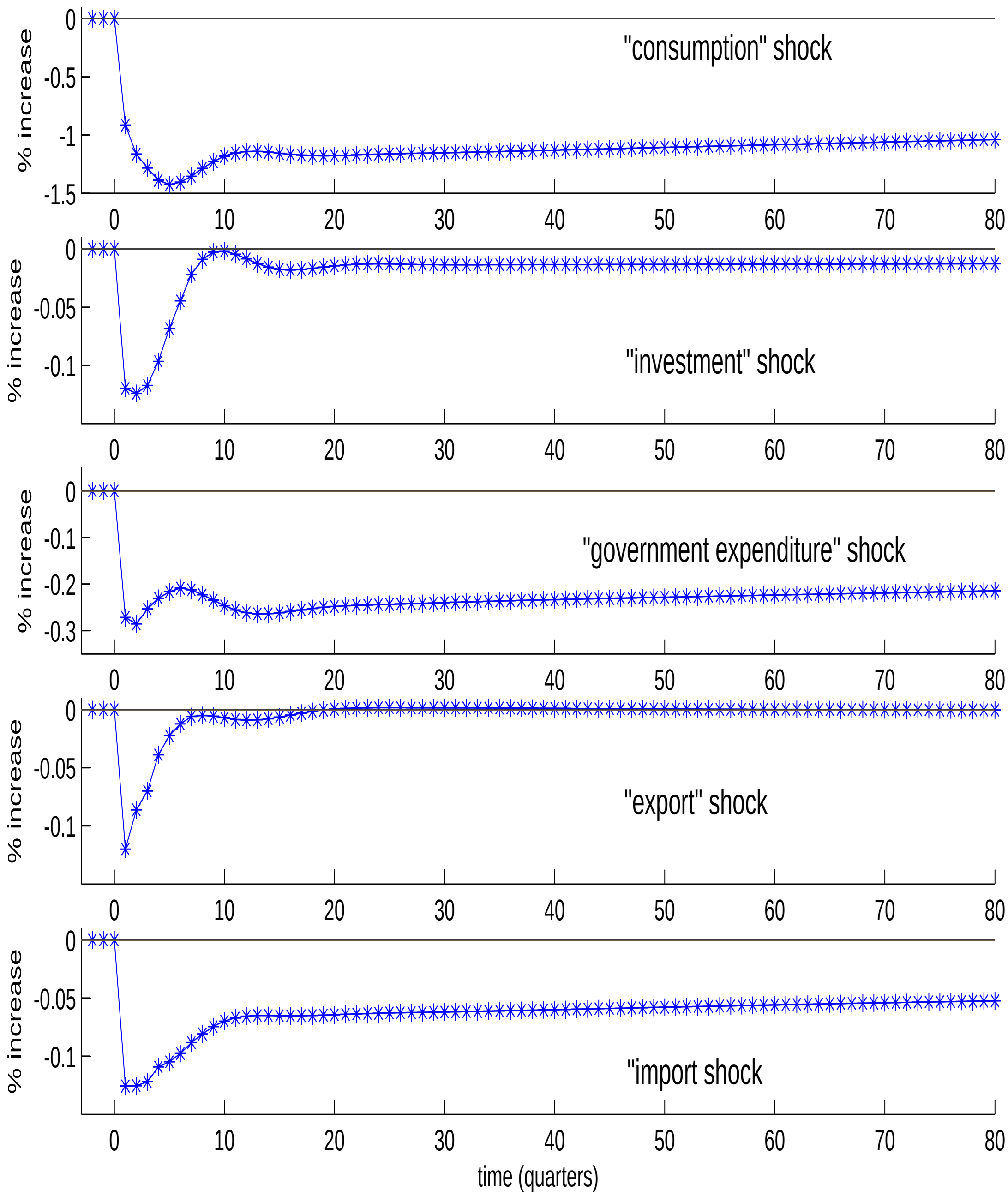
Notes: The graph plots the true responses of y_t to shocks in $e_{x,t}$ and $e_{z,t}$ and the response according to the misspecified one-type-shock model.

Figure 2: One-type-shock model and predicted impact of the shock



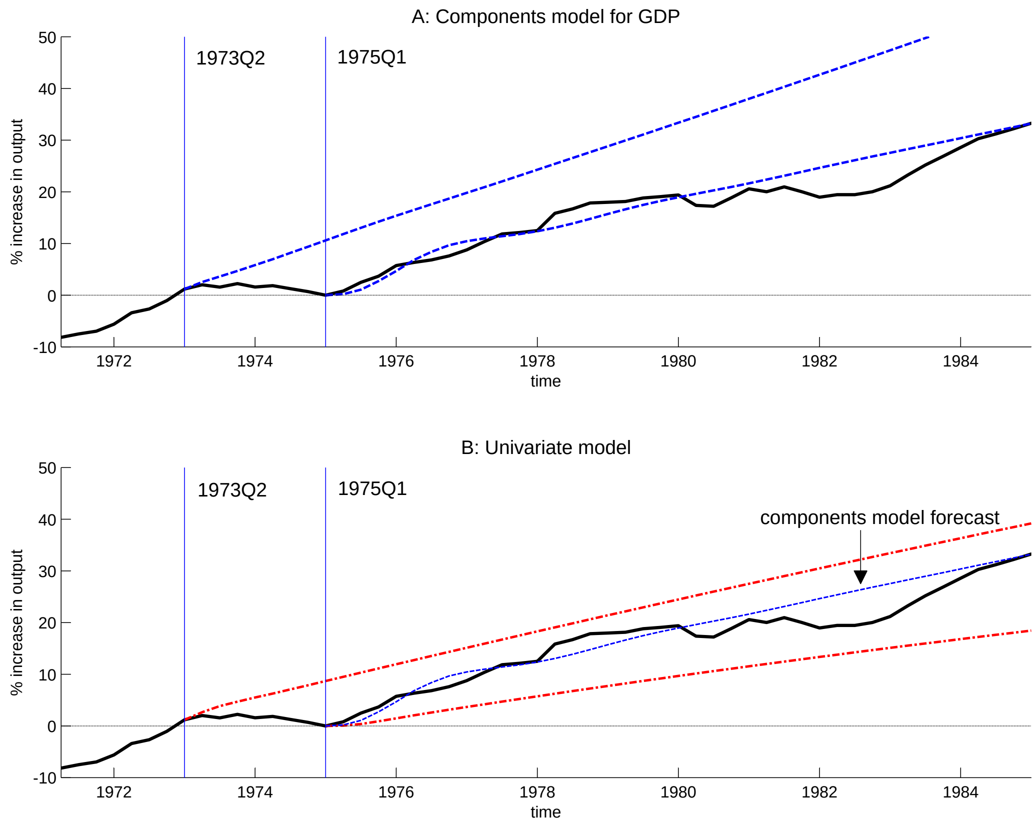
Notes: The graph plots the predicted responses of output following a one-standard-deviation negative shock.

Figure 3: Components model and predicted impacts of shocks



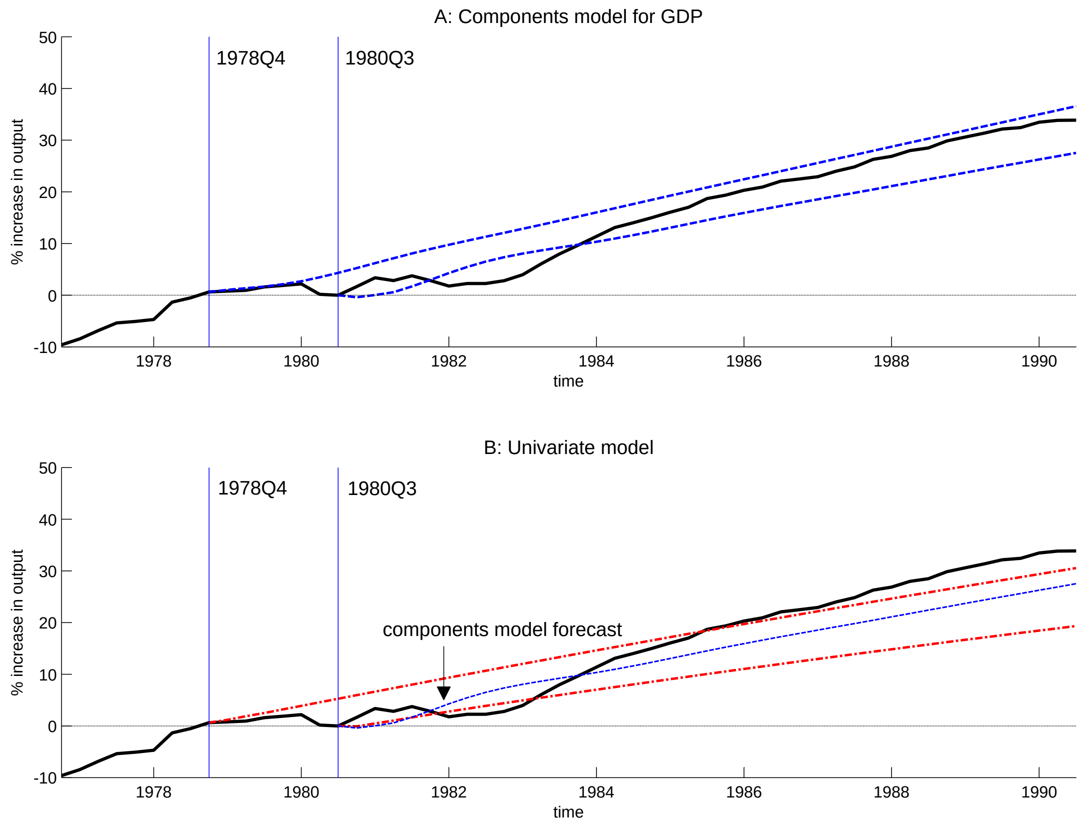
Notes: The graphs plots the predicted responses of output following a one-standard-deviation negative shock in the indicated component.

Figure 4: 1975Q1 (and 1973Q1) forecasts and realized GDP



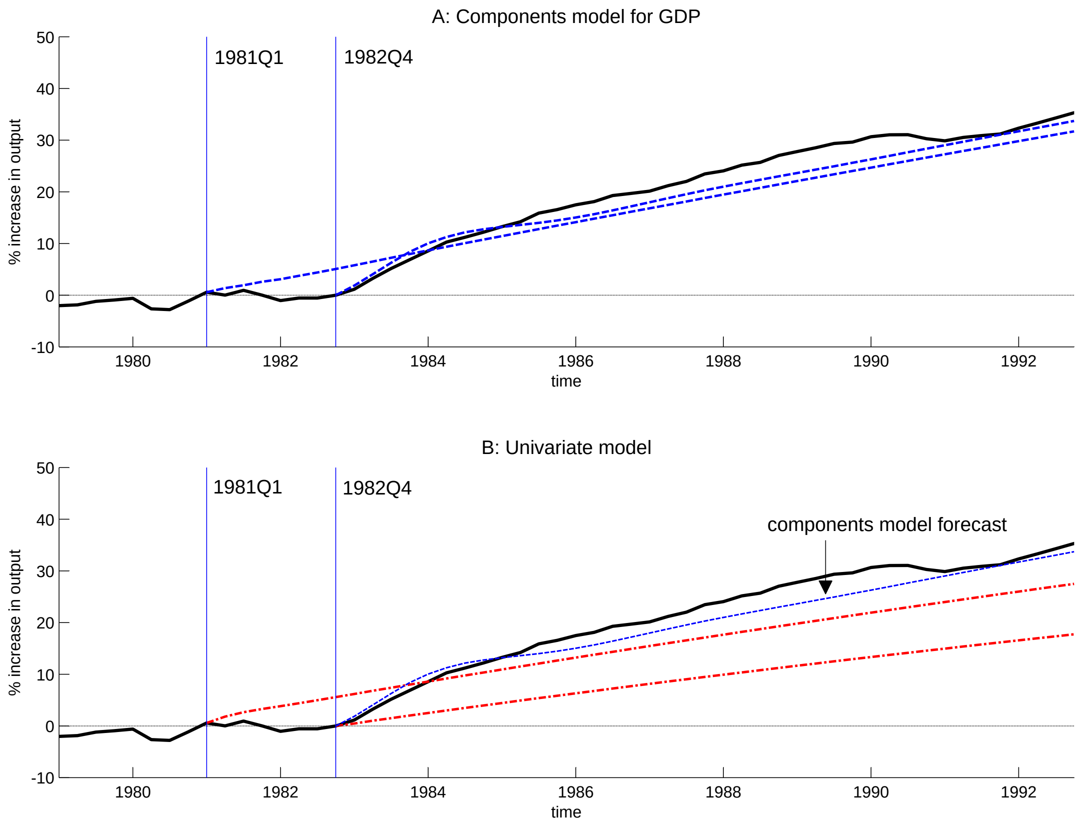
Notes: The solid line represents the true data and the vertical lines represent the forecasting points. In the top panel, the '- -' lines indicate the forecasts of the components model. The bottom line corresponds to the predictions made at the trough of the recession and the top line to the forecasts made at a recent quarter when predictions were most positive. In the bottom panel, the '. -' lines indicate the similar forecasts made with the univartiate model. In the bottom panel, we repeat the forecast from the top panel made at the trough of the recession. Again, it is indicated with '- -'.

Figure 5: 1980Q3 (and 1978Q4) forecasts and realized GDP



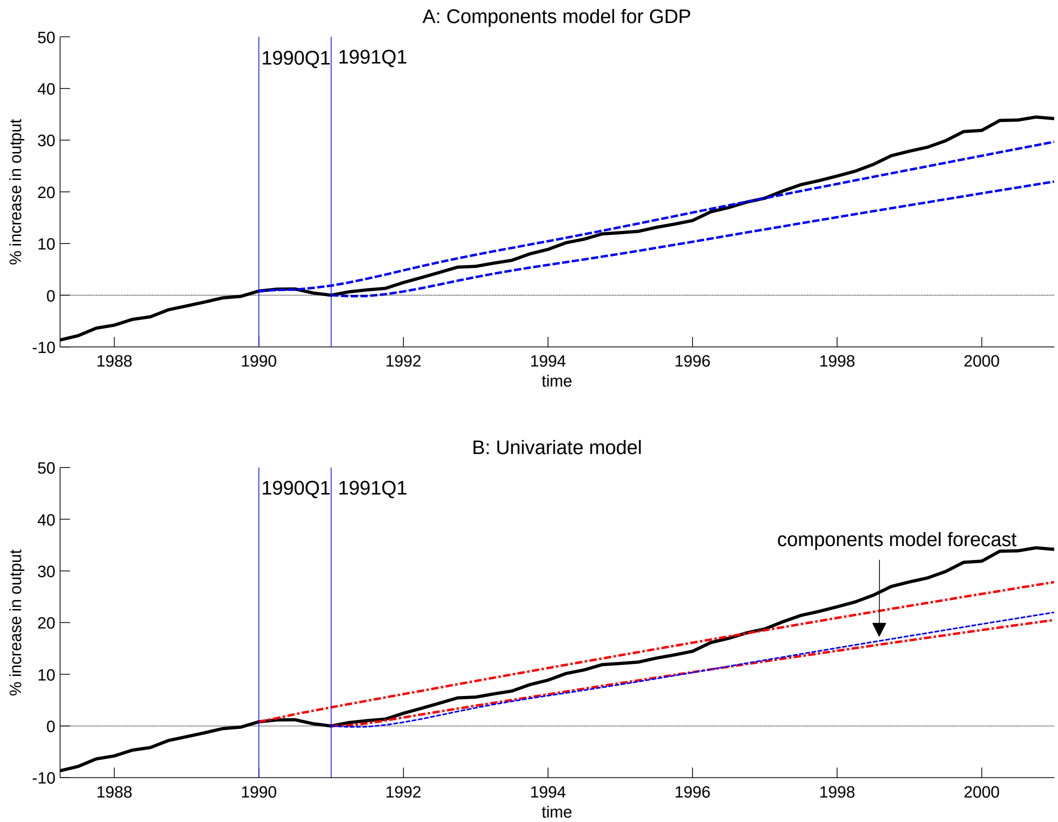
Notes: The solid line represents the true data and the vertical lines represent the forecasting points. In the top panel, the '- -' lines indicate the forecasts of the components model. The bottom line corresponds to the predictions made at the trough of the recession and the top line to the forecasts made at a recent quarter when predictions were most positive. In the bottom panel, the '. -' lines indicate the similar forecasts made with the univartiate model. In the bottom panel, we repeat the forecast from the top panel made at the trough of the recession. Again, it is indicated with '- -'.

Figure 6: 1982Q4 (and 1981Q1) forecasts and realized GDP



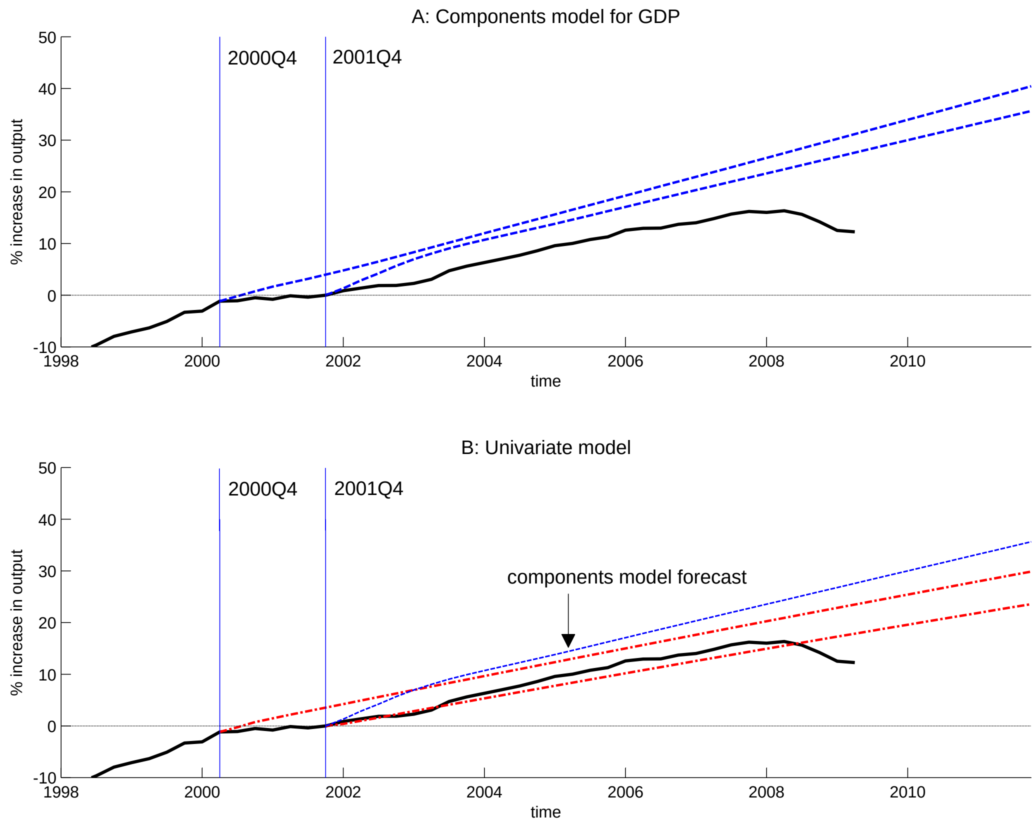
Notes: The solid line represents the true data and the vertical lines represent the forecasting points. In the top panel, the '- -' lines indicate the forecasts of the components model. The bottom line corresponds to the predictions made at the trough of the recession and the top line to the forecasts made at a recent quarter when predictions were most positive. In the bottom panel, the '. -' lines indicate the similar forecasts made with the univartiate model. In the bottom panel, we repeat the forecast from the top panel made at the trough of the recession. Again, it is indicated with '- -'.

Figure 7: 1991Q1 (and 1990Q1) forecasts and realized GDP



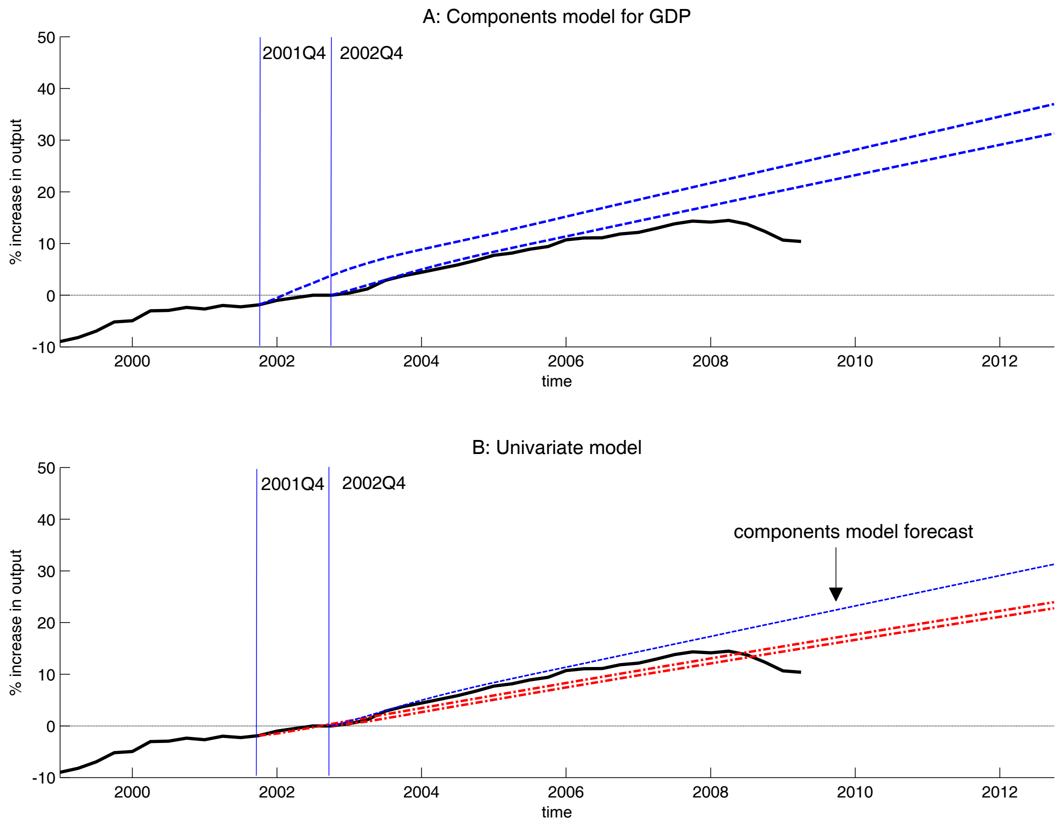
Notes: The solid line represents the true data and the vertical lines represent the forecasting points. In the top panel, the '- -' lines indicate the forecasts of the components model. The bottom line corresponds to the predictions made at the trough of the recession and the top line to the forecasts made at a recent quarter when predictions were most positive. In the bottom panel, the '. -' lines indicate the similar forecasts made with the univartiate model. In the bottom panel, we repeat the forecast from the top panel made at the trough of the recession. Again, it is indicated with '- -'.

Figure 8: 2001Q4 (and 2000Q4) forecasts and realized GDP



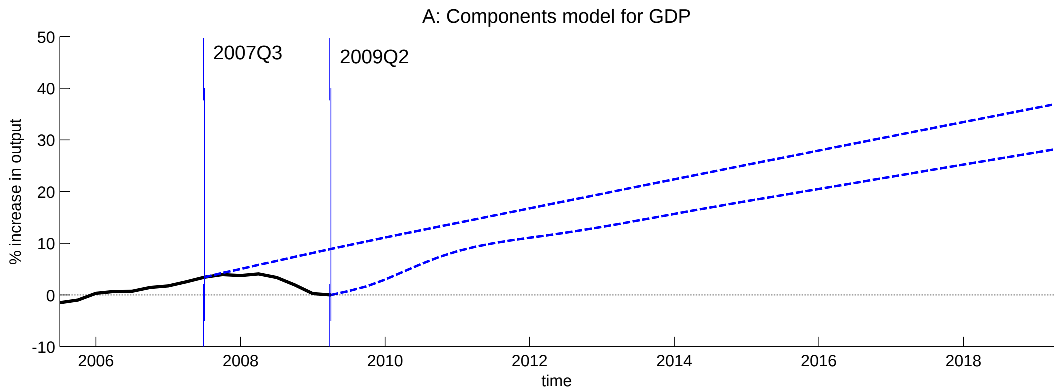
Notes: The solid line represents the true data and the vertical lines represent the forecasting points. In the top panel, the '- -' lines indicate the forecasts of the components model. The bottom line corresponds to the predictions made at the trough of the recession and the top line to the forecasts made at a recent quarter when predictions were most positive. In the bottom panel, the '. -' lines indicate the similar forecasts made with the univartiate model. In the bottom panel, we repeat the forecast from the top panel made at the trough of the recession. Again, it is indicated with '- -'.

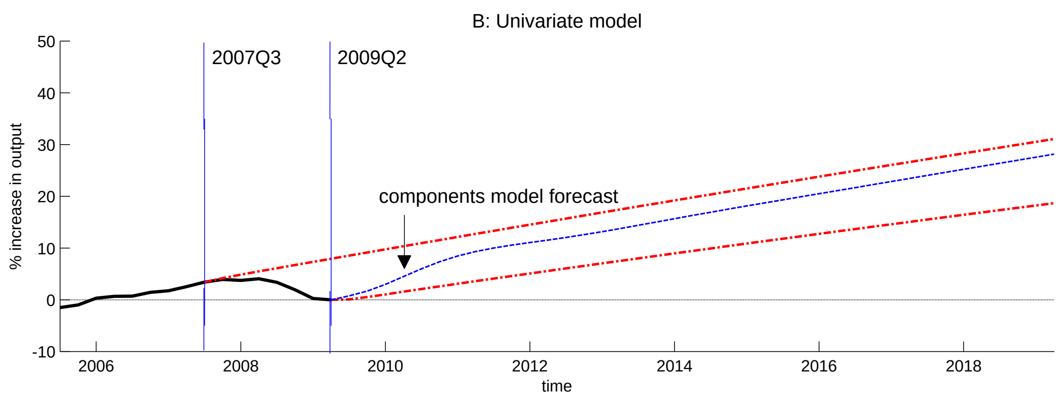
Figure 9: 2002Q4 (and 2000Q4) forecasts and realized GDP



Notes: The solid line represents the true data and the vertical lines represent the forecasting points. In the top panel, the '- -' lines indicate the forecasts of the components model. The bottom line corresponds to the predictions made at 2002Q4 one year after the trough of the recession and the top line to the forecasts made one year ealier. In the bottom panel, the '. -' lines indicate the similar forecasts made with the univartiate model. In the bottom panel, we repeat the forecast from the top panel made at 2002Q4. Again, it is indicated with '- -'.

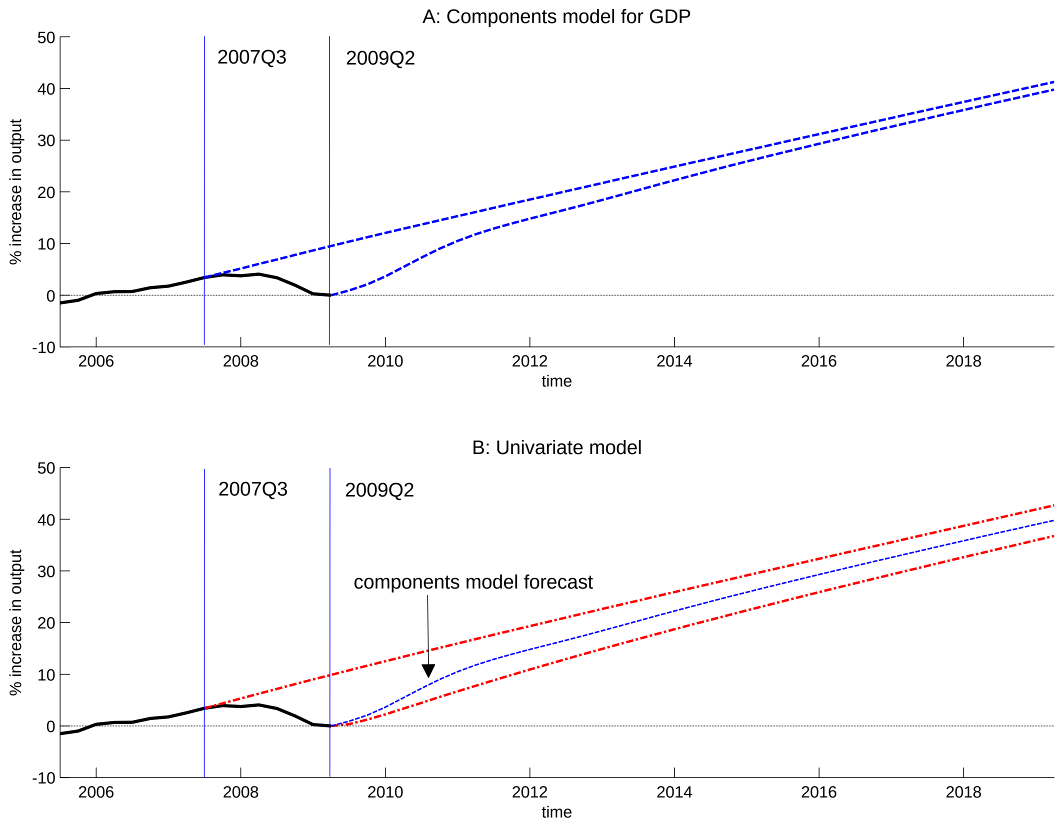
Figure 10: 2009Q2 (and 2007Q3) forecasts and realized GDP





Notes: The solid line represents the true data and the vertical lines represent the forecasting points. In the top panel, the '- -' lines indicate the forecasts of the components model. The bottom line corresponds to the predictions made at the end of 2009Q2 and the top line to the forecasts made at a recent quarter when predictions were most positive. In the bottom panel, the '. -' lines indicate the similar forecasts made with the univartiate model. In the bottom panel, we repeat the forecast from the top panel made at 2009Q2. Again, it is indicated with '- -'.

Figure 11: 2009Q2 (and 2007Q3) forecasts and realized GDP; deterministic trend added



Notes: The solid line represents the true data and the vertical lines represent the forecasting points. In the top panel, the '- -' lines indicate the forecasts of the components model. The bottom line corresponds to the predictions made at the end of 2009Q2 and the top line to the forecasts made at a recent quarter when predictions were most positive. In the bottom panel, the '. -' lines indicate the similar forecasts made with the univartiate model. In the bottom panel, we repeat the forecast from the top panel made at 2009Q2. Again, it is indicated with '- -'.