On Economic Forecasting and International Linkages
Three Empirical Factor Studies

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2007

vorgelegt
von

Diplom-Volkswirtin Sandra Eickmeier

aus

Heidelberg

Referent: Prof. Dr. Juergen B. Donges

Korreferent: Prof. Dr. Clemens Fuest

Tag der Promotion: 21.12.2007

Contents

Tables and Figures

I. Introduction*

In the three studies that comprise this dissertation, the first study (chapter II) investigates the forecast performance of large-scale factor models by means of a meta-analysis, while the second and third studies (chapters III and IV) estimate economic linkages at business-cycle and longer frequencies between Germany and the US and among the member states of the European Monetary Union (EMU), respectively. The present work is relevant for two types of readers: applied economists – particularly forecasters and international macroeconomists – and theoretical econometricians interested in factor models. Factor models form the link connecting the three studies. For this reason, this introduction embeds the present work in the existing factor literature and shows its contribution. It also motivates the use of factor models from a forecaster's, an international economist's and a monetary economist's point of view and explains recent advances in the factor literature. The individual studies are – in more detail – motivated in chapters II to IV. The chapters also highlight my contributions and discuss the relevant open issues.

Economists observed very early that macroeconomic variables move in parallel. Following this fundamental observation, Burns and Mitchell (1946) formulated the concept that these variables are driven by a few common factors or shocks. Factor models account for this type of interdependency between variables.

The variables under consideration are summarized in an N-dimensional vector $\mathrm{y}_t = [\mathrm{y}_{1t} \quad \cdots \quad \mathrm{y}_{Nt}]'$ (with $t = 1,...,T$), and each component, y_{it} is assumed to have the following factor structure: $\mathrm{y}_{it} = \mathrm{x}_{it} + \xi_{it} = \Lambda_i(L)'\mathrm{f}_t^d + \xi_{it} = \Lambda_i'\mathrm{f}_t + \xi_{it}$, where x_{it} and ξ_{it} denote the vectors of common and idiosyncratic (i.e. variable-specific) components, f_t^d the $q \times 1$-vector of common dynamic factors (which are mutually orthogonal and orthogonal to ξ_{it}), and $q \ll N$. The q factors and their s lags can affect the variables contained in y_t through the $q \times 1$ vector of factor loadings $\Lambda_i(L) = \Lambda_{i0} + \Lambda_{i1}L + ... + \Lambda_{is}L^s$. Each dynamic factor model also has a static representation, and f_t denotes the $r \times 1$ vector of common static factors which may contain the dynamic factors f_t^d and their lags, and Λ_i denotes the corresponding vector of factor loadings.

This model was first proposed by Sargent and Sims (1977) and Geweke (1977) and, at that time, designed for small datasets with a very few variables. Such small-scale factor models are labeled "strict factor models", since they assume that the idiosyncratic components are

* I thank Elke Baumann, Ben Craig, Juergen B. Donges and Malte Knüppel for helpful comments on this chapter. Small parts of this chapter were taken from Breitung and Eickmeier (2006), "Dynamic factor models", in: O. Hübler, and J. Frohn (eds.), Modern econometric analysis, chapter 3, Springer.

mutually uncorrelated. They typically are estimated by maximum likelihood. These strict factor models were employed over the following two decades in academic work.

At the same time, economists in central banks (and also in research institutions) in their daily work began to observe and monitor a larger number of variables than could be incorporated in such models, reflecting the central bankers' practice of "looking at everything" as emphasized by Bernanke and Boivin (2003). These variables reflect a variety of economic concepts, stem from a diversity of sources and are available in preliminary, revised and final versions. They are used to produce forecasts and to come to political decisions. The Federal Reserve Board communicates in its regular Monetary Policy Report to the Congress its assessment of the state of the economy by explaining the evolution of a large number of different indicators (see, for example, FRB, 2007). And also the ECB (2004) states that "[a] successful monetary policy [...] has to be broadly based, taking into account all relevant information in order to understand the factors driving economic developments [...]". It lists the variables subject to economic analysis, the first pillar of its monetary strategy: "developments in overall output; aggregate demand and its components; fiscal policy; capital and labour market conditions; a broad range of price and cost indicators; developments in the exchange rate, the global economy and the balance of payments; financial markets; and the balance sheet positions of euro-area sectors" and goes on to explain that "[all] these factors are helpful in assessing the dynamics of real activity and the likely development of prices from the perspective of the interplay between supply and demand in the goods, services and factor markets at shorter horizons".

In view of this gap between central bankers' practice and small-scale factor models employed by academics, the challenge for the latter was to develop models able to exploit information from a large number of variables. It was, furthermore, clear that it would be infeasible to estimate such large-scale models with maximum likelihood techniques. Instead, easy to apply estimation techniques would also need to be proposed to make it worthwhile for economists to deal with large, often costly constructed, datasets. The breakthrough came with the contribution of Stock and Watson (1998, 2002a), who proposed a model able to exploit information from large datasets (i.e. datasets having both a large cross-section dimension N and a large time dimensions T). "Large N" means that the datasets can contain hundreds or even thousands of variables. Factor models summarize the information contained in these variables in y_t, reducing the dimension of these large datasets. As pointed out by Stock and Watson (2002a), the large N assumption allows one to relax the assumption of mutual independence of idiosyncratic components characterizing "strict factor models". Instead, the idiosyncratic components can be only weakly cross- (and serially) correlated in the sense of Bai and Ng (2002), for which reason these models are labeled "approximate factor models". Cross-section correlation between idiosyncratic components is generally considered to be more realistic. Stock and Watson (1998, 2002a) suggest estimating the approximate factor model with a static principal component analysis which is extremely easy and is trivial to accomplish in Matlab. In the meanwhile, two other – somewhat more complex – methods to estimate large-dimensional factor models have been suggested. While Forni, Hallin, Lippi and Reichlin (2000) estimate the factor model with dynamic principal component analysis, Kapetanios and Marcellino (2004) rely on a state-space model and linear algebra techniques. For a detailed exposition of the three different factor estimation techniques, see chapter II.

Not least due to these advances in estimation techniques, large-scale factor models have become popular and widely applied in recent years in the academic world and among policy makers. The trade-off between gains from exploiting possibly valuable information from large datasets and costly collection and processing of these large datasets was further alleviated by databases which are now readily available from large companies such as Datastream, Reuters and Bloomberg. Moreover, once collected, the data can be processed easily and rapidly owing to the now widespread use of high-capacity computers and modern software with freely available Matlab codes that serve to estimate factor models and to carry out related tests.[1]

This – probably most important – ability of dynamic factor models to exploit information from large datasets is particularly useful for forecasting and for some fields of macroeconomic analysis. In what follows, we will explain the benefits of dynamic factor models to forecasters and monetary and international macroeconomists. The use of factor models in these fields is motivated with some examples. We will also briefly summarize existing applications (see Breitung and Eickmeier, 2006, for a more detailed overview) and explain the contributions of this thesis.

1. Forecasting with factor models

Macroeconomic forecasting is important for policymakers and private agents who base their decisions on the current but also on future states of the economy. It is therefore crucial that the forecasts are reliable. Unfortunately, forecasts are (almost) always erroneous, and depending on their size and persistence, the forecast errors grab the public's attention. This is often unfair. The sources of forecast errors may be twofold. Shocks, i.e. unexpected movements of macroeconomic variables, may lead to forecast errors. The shocks (and therefore also immediate movements of macroeconomic variables due to these shocks) are, by definition, unpredictable, even for the best forecasters, and forecast errors resulting from them are inevitable. However, forecast errors also arise when forecast models are misspecified and relevant information is not taken into account. In this case, forecasters are responsible for the resulting forecast errors. Choosing an adequate model and underlying information set presents a great challenge for forecasters. Factor models are particularly well suited for forecasting since they can potentially solve some of the issues surrounding model and variable selection.

[1] Particularly useful websites are those of Serena Ng (http://www-personal.umich.edu/~ngse/), Mark M. Watson (http://www.wws.princeton.edu/mwatson/) and Mario Forni (http://www.economia.unimore.it/forni_mario/).

Let us first briefly outline how forecasting with factor models compares to forecasting with commonly applied simple time series regression models. Forecasters who aim at predicting a certain macroeconomic variable $\tilde{y}_t$ usually estimate an equation such as $\tilde{y}_{t+h} = \tau_0 + \tau_1(L)'\mathrm{g}_t + \tau_2(L)\tilde{y}_t + e_{t+h}$, where h is the forecast horizon, g_t is a vector of a few selected observable indicators and $\tau_i(L) = \tau_{i0} + \tau_{i1}L + \tau_{i2}L^2 + ... + \tau_{ip^{\tau_i}}L^{p^{\tau_i}}, i = \{1,2\}$ denote lag polynomials. This framework is fairly general and embeds pure autoregressive models (for $\tau_{1j} = 0, j = \{0,..., p^{\tau_1}\}$) which are among the simplest univariate models used in practice. One problem with this framework is, however, that the target variable may be influenced by movements of not just a few, but of a large number of indicators, and forecasters may want to take them all into account. In other words, the true dimension of g_t may not be small, but large. A regression of $\tilde{y}_t$ on a large number of indicators is, however, infeasible, since the number of degrees of freedom is limited in such regression models. Moreover, parameter uncertainty increases with the number of regressors, and forecasts will be less precise. The idea of forecasting with factor models is to point out that if the indicators useful to predict $\tilde{y}_t$ comove, then they can be summarized into few common factors. The factors instead of the large number of indicators themselves can then be included into the forecasting equation. This again renders the estimation feasible. Factor forecasting thus proceeds in two steps. In a first step, a $r \times 1$ vector of unobserved common factors f_t is estimated from a large dataset y_t. In a second step, the estimated factors or a subset thereof (denoted here by the $\underline{r} \times 1$ vector $\hat{\mathrm{f}}_t^0$) replace g_t in the forecasting equation $\tilde{y}_{t+h} = \alpha_0 + \alpha_1(L)'\hat{\mathrm{f}}_t^0 + \alpha_2(L)\tilde{y}_t + \varepsilon_{t+h}$, where $\alpha_i(L) = \alpha_{i0} + \alpha_{i1}L + \alpha_{i2}L^2 + ... + \alpha_{ip^{\alpha_i}}L^{p^{\alpha_i}}, i = \{1,2\}$ denotes $\underline{r}$-dimensional ($i = 1$) and scalar ($i = 2$) lag polynomials, respectively.

The main difficulty is to decide upon the variables to be included in the large dataset y_t. In theory, the precision of factor estimates and hence also of the factor forecasts should improve with N, and the forecaster should simply use the largest possible dataset s/he has available. Typical datasets used to predict output growth with factor models contain real variables such as GDP, demand components, industrial production, labor productivity, variables capturing labor market developments, retail trade, and expectations taken from surveys as well as nominal variables such as consumer and producer prices, interest rates of different maturities, asset prices, monetary aggregates and credits (cf., for example, Schumacher, 2007; Den Reijer, 2005). These datasets also roughly correspond to the set of variables the ECB deems relevant to govern its first pillar. Some factor forecasters argue that international variables have predictive content for domestic output growth and consider them as well (cf. Cheung and Demers, 2007; Banerjee et al., 2006a). Inflation is in most cases predicted with similar datasets (cf. Cheung and Demers, 2007; Gavin and Kliesen, 2006; Stock and Watson, 1998), while some forecasters only include aggregated and disaggregated price measures to forecast inflation (cf. Camba-Méndez and Kapetanios, 2005).

Enabling forecasters to exploit information from large datasets is not the only advantage of factor models. A second advantage is that factor models are particularly well suited to cope with structural breaks. In the case of an obvious structural break, forecasters (just like macroeconomic analysts) have two options. They will either start the estimation of their model after the break and disregard the data before the break or, if they consider the data before the break relevant, they will exploit the information before the break as well. An example is the foundation of the European Monetary Union (EMU). Some empirical studies investigating issues regarding euro-area economic developments start in 1999, while others exploit information from longer time series and account for possible breaks in the series by including a dummy variable into a regression model or by removing breaks from the series in a rather *ad hoc* manner. Statistics may also simply be not available backwards from a certain point. Prominent examples are the German economy with key data lacking for East Germany before unification and central and east European countries with data being generally available only from the beginning of the 1990s onwards. Forecasting with factor models can be an advantage no matter whether forecasters decide (or are obliged) to rely only on data after the structural break or to consider information from the period before the break as well. In the first case, a large cross-section dimension ensures precise factor estimates even if the time dimension is relatively small. Banerjee et al. (2006b) provide forecasting and Eickmeier and Breitung (2006) provide analytical illustrations using central and east European countries. In the second case (when forecasters decide to use long time series), Stock and Watson (1998, 2007) have shown that breaks in the factor loadings generally represent a minor problem. Even if there is mild time variation in the factor loadings, factors can be estimated consistently if $N > T$ and if the instability is sufficiently different across series. Intuitively, under these conditions, using many series to estimate the factors leads to an "averaging out" of the breaks, and full sample estimates of the factors can be used for forecasting.[2]

A third advantage is that factor models are particularly useful in real-time environments. Statistical offices publish preliminary estimates of national account and other data, followed by revised and then final versions of these estimates. Factor models separate true realizations from measurement error (and other variable-specific movement) by decomposing each series into a common part and an idiosyncratic part which is likely to contain the measurement error. Since the common part of the variables, the factors, obtain a relatively large weight in the estimation, factor-based forecasts will be more robust against measurement error than forecasts based on small-scale models. Not surprisingly, Schumacher and Breitung (2006), who compare factor forecasts of German GDP based on a real-time dataset with those based on a final dataset, find that revisions barely affect factor forecasts.

Real-time environments are also characterized by missing values. Some data are published with a delay and the last (or the last few) observation(s) may not be available yet when the forecast is made. Another possibility is that some statistics start later than others. Moreover,

[2] Stock and Watson (2007) further show that breaks in the dynamics of the factors may be more of an issue and that forecasters may be better off using subsample or time-varying estimates of coefficients in the forecasting equation.

some series may be discontinued due to a general lack of quality of data provided by statistical offices or simply because statistics were not reported to them for some reason. Data with missing observations are of little use for small regression-based approaches without interpolating them and making rather strong *ad hoc* assumptions. By contrast, they can be exploited with factor models. Stock and Watson (2002a) have suggested the use of an expectation-maximization (EM) algorithm which can handle unbalanced panels, i.e., panels containing time series with missing observations. This algorithm involves, in a first step, extracting factors from the full dataset which includes all times series with missing values where the latter are replaced with initial estimates.[3] These factors are then used in a second step to obtain improved estimates of the missing observations. The estimation procedure is repeated until convergence. The EM algorithm can also be applied to a panel with mixed frequencies. Monthly indicators may contain valuable information for quarterly series and forecasters may want to exploit information incorporated in these indicators as well. Schumacher and Breitung (2006) apply a large-scale factor model and the EM algorithm to a panel of 13 quarterly and 39 monthly indicators to predict quarterly GDP and find significant improvements of estimations based on mixed-frequency data compared to the quarterly data only.

A fourth advantage of factor models is that, at least in theory, they should be able to make more accurate predictions over longer horizons than small-scale time series models. Forecasts based on the latter models only have a predictive content over very short horizons (often barely longer than ½ year). Policy makers may, however, be also interested in longer horizon predictions. For example, the ECB targets medium-term inflation. Transmission lags make it impossible in the short run for monetary policy to offset unanticipated shocks, and the medium-term orientation should help to avoid introducing unnecessary volatility (ECB, 2004). The main reason why factor models could deliver improved longer-horizon forecasts is that the estimated common factors which enter the factor forecast equation tend to be more persistent than the idiosyncratic components, as shown by Giannone et al. (2002) for US and Altissimo et al. (2001) for euro-area output growth, and therefore better predictable.

Due to these advantages, large-scale dynamic factor models have become increasingly popular over the past few years to forecast macroeconomic variables. Many studies have evaluated the forecast performance of factor models for different variables and different countries (cf. Stock and Watson, 2002a and 2005, and Banerjee and Marcellino, 2006, for output and inflation in the US; Artis et al., 2005, for output and inflation and Kapetanios et al., 2005, for inflation in the UK; Schumacher, 2006, for output in Germany; Brisson et al., 2003, and Cheung and Demers, 2007, for output and inflation in Canada; Bruneau et al., 2007, for inflation in France; Favero et al., 2004, for inflation in Italy; Den Reijer, 2005, for

[3] Schumacher and Breitung (2006) use as initial values for the missing observations the unconditional means of the series.

Dutch output; Schneider and Spitzer, 2004, for output and Moser et al., 2007, for inflation in Austria; Van Nieuwenhuyze, 2006, for output in Belgium; Banerjee et al. (2006a) for the central and east European countries; for further references see chapter II), and policy makers and research institutions are on the verge of integrating factor models into their regular forecasting process.

Results from early studies were promising in this respect. However, more recent studies have found that factor models are only slightly, if at all, better than other models. In view of the arguments in favor or factor models, this has come as a surprise to many researchers and practitioners. Theoretically, various determinants are able to affect the forecast performance of factor models. Most determinants can be manipulated directly by the forecasters, thus giving them a great deal of responsibility. Factor models *can* deliver improved forecasts. However, factor models will fully unfold their abilities, only if forecasters respect some basic principles. For instance, the datasets from which the factors are estimated need to be constructed carefully. Boivin und Ng (2006) have shown that not only the pure amount of data (i.e. the size of the dataset, measured by N and T), but also its quality matters for factor forecasts. Only if adding more variables or observations leads to tighter relationships between the components of y_t on the one hand and between y_t and the variable to be forecast on the other hand can a reduction of the prediction error be achieved *ceteris paribus*. Chapter II provides a more elaborate discussion on the determinants and risks. It employs a meta-analysis – an empirical tool which has only recently been applied in macroeconomics – to systematically summarize existing studies on the forecast performance of factor models and to examine which determinants have improved factor forecasts and which have made them worse.

2. Macroeconomic analysis with factor models

Fields of macroeconomic analysis where factor models may be particularly useful are monetary economics and international macroeconomics. As for forecasting, the most important advantage is that a large number of variables can be exploited in these analyses.

We have already pointed out that central bankers observe lots of data and use them when making their assessments. Some empirical applications estimate monetary vector autoregressive (VAR) models with common factors extracted from large macroeconomic datasets. These factors capture all possible information used by the central bank and not already captured by the observable indicators in the VAR. This approach is labeled factor-augmented VAR (FAVAR) and has been suggested by Bernanke et al. (2005) and further developed by Stock and Watson (2005c). Applications by Bernanke et al. (2005) for the US and Favero et al. (2005) for Europe suggest that including factors in the VAR helps avoid the omitted variable bias and eliminates the well-known price puzzle. The price puzzle denotes a decrease rather than an increase of prices after an expansionary monetary policy shock and is usually present in small-scale VAR models.

Factor models are also particularly well suited to capture comovements representing both intranational and international economic interdependencies between variables that can be complex. Agents from different economic sectors and different countries interact with each other through various channels. An attractive feature for macroeconomic and especially for international analyses is that factor modelers can exploit all information possibly necessary to accurately model the interdependencies and to estimate the common variation of the variables more precisely.

A second important feature is that the common factors or common shocks may be of intrinsic interest in macroeconomic applications. The factor literature has made some progress in this direction. Policy makers may, for example, find it useful to know the dimension of (i.e. the number of factors or shocks driving) the domestic, European and/or world economies. Informal tests based on the variance share explained by common factors (Forni et al., 2000) and formal information criteria (Bai and Ng, 2002, 2007b; Breitung and Kretschmer, 2007; Amengual and Watson, 2006) have been suggested to answer this question. In addition, the patterns of national, European or world business cycles are of interest. The most prominent examples of national/regional business cycles are the *Chicago Fed National Activity Index* (*CFNAI*)[4] for the US and *EuroCOIN*[5] for the euro area. The *CFNAI*, which dates back to 1967, is simply the first principal component of a large macroeconomic dataset. It is the most direct successor to indicators which were first developed by Stock and Watson. *EuroCOIN* is estimated as the (smoothed) common component of euro-area GDP based on dynamic principal component analysis. It was developed by Altissimo et al. (2001) and is made available from 1987 onwards by the CEPR. Similarly, factor models are also used to estimate core inflation (cf. Cristadoro et al., 2001, for the euro area and Kapetanios, 2004, for the UK).

The economic interpretation of the common factors provides an important challenge. In dynamic factor models, the common factors are unobservable. Estimation based on any of the above-mentioned methods, however, yields factors which are not uniquely identified, i.e. not f_t, but a rotation of f_t, Rf_t, is estimated, with all elements of Rf_t being orthogonal. This issue is not relevant for forecasters, since the prediction is unchanged whether Rf_t or f_t is included in the forecasting equation. In macroeconomic analyses, however, individual factors (or underlying individual shocks) may be of interest. This was long considered a major drawback of factor models and delayed their use for macroeconomic analyses for quite some time. Practitioners apparently have a strong preference to deal with measured variables, and view the factors, i.e. constructs without a clear economic meaning, with suspicion. In the meanwhile, however, progress has been made in this regard, and a number of methods which

[4] See http://www.chicagofed.org/economic_research_and_data/cfnai.cfm.
[5] See http://www.cepr.org/data/eurocoin/.

allow one to interpret the factors have been developed. The estimated factors themselves are rotated and the rotation matrix can be determined by means of statistical or economic criteria; see the tests developed by Bai und Ng (2006) and the applications of Gilbert und Meijer (2006) and Eickmeier (2005). The alternative upon which this dissertation is relying is the structural factor model suggested by Forni and Reichlin (1998). This model assumes that the vector of dynamic factors f_t^d has a VAR representation: $A(L)f_t^d = Qw_t$, where w_t denotes the q-dimensional vector of common structural shocks and $A(L) = A_0 + A_1 L + ... + A_p L^p$. As in the structural VAR literature, the shocks w_t are identified by imposing restrictions on Q. More recent approaches directly restrict the factor loadings $\Lambda_i(L)$ to provide the factor with economic meaning. Doz et al. (2006), for instance, employ an approach based on quasi-maximum likelihood. Kose et al. (2003a) impose zero restrictions on the factor loadings, so that some factors only affect certain variables or groups of variables. That approach is preferable in an international macro context if not only global factors (i.e. factors driving all variables in the panel), but also regional factors are of interest. Giannoni and Boivin (2006) exploit restrictions obtained from general equilibrium models. All these more recent approaches have the drawback of being more complex than the simple principal component analysis. The work by Giannoni and Boivin (2006) highlights an interesting recent tendency to re-impose more structure in the originally agnostic factor world.

The fact that factor modelers can remain relatively agnostic about the structure of the economy provides a third distinctive feature to these estimates, in spite of these recent tendencies. Factor models do not need to rely on overly tight assumptions as is sometimes the case with structural models. It also represents an advantage over structural vector-autoregressive (VAR) models where the researcher has to take a stance on which variables to include, thereby influencing the outcome. Moreover, in these models the number of variables determines the number of shocks, while the latter can be freely estimated in factor models. Obviously, structural models cannot be replaced by factor models, when carrying out counterfactual simulations or when trying to understand the transmission of individual shocks through the system and the role of structural parameters. Factor models have only a reduced-form representation, and estimated factors are not uniquely identified when estimated for instance with principal components.

A fourth feature is that recent technical advances allow a focus not only on common stationary factors, but also on non-stationary factors or stochastic trends. Approximate factor models were originally designed for stationary variables. To employ these models, non-stationary variables were first made stationary through differencing, and then stationary factors and stationary idiosyncratic components were estimated from a fully stationary dataset. Recent estimation methods proposed by Bai (2004) and Bai and Ng (2004) allow one to estimate non-stationary factors and stationary idiosyncratic components (Bai) and non-stationary factors and non-stationary idiosyncratic components (Bai and Ng) from a non-stationary dataset. Common non-stationary factors or trends are of particular interest for international macroeconomists and subsequent paragraphs will clarify why factor models provide a particularly useful approach.

Fifth, current advances account for time-variation of parameters in factor models. These advances can be motivated by the ongoing global integration of goods and financial markets and institutional changes such as the formation of EMU which may alter the international comovements. Del Negro and Otrok (2005) have developed a factor model with time-varying factor loadings, where $\Lambda(L) = \Lambda_t(L)$. Mumtaz and Surico (2006) model time-variation of the coefficients associated with the VAR model fitted to the factors. These models are, however, still much more difficult to estimate than models with time-invariant parameters. In a recent paper, Breitung and Eickmeier (2007) propose a Chow-type test to detect a break in the factor loadings. Although factor estimates are barely affected by structural breaks in the loadings as discussed in the subsection on forecasting with factor models, the latter may themselves be of interest to the researcher who conducts a macroeconomic analysis. If a break is found, it may be better to carry out the analysis for subsamples separately.

Factor models clearly are particularly well suited to application in a world where economic variables comove. In what follows, we present two recent examples where comovements are discussed intensively. These examples demonstrate the usefulness of dynamic factor models and also motivate chapters III and IV of the thesis.

Example 1: International shock transmission from the US

The first example is the international transmission of US shocks. Cyclical fluctuations in the US, the world's largest economy, affect the economies in Europe, including Germany. The question of whether and to what extent business cycle fluctuations are transmitted internationally from the US becomes an issue for policymakers and the public whenever fluctuations observed in the US go beyond "normal" cyclical fluctuations. This was the case in the second half of the 1990s, when the US economy experienced extraordinary productivity gains. These gains were due to technological advances mainly in the information and communications industry and to the wide-spread use of these technologies in many sectors of the US economy. They were thought to be long lasting, which stimulated global demand and gave rise to a stock market bubble in the US, as well as in many other countries. The subsequent bursting of the bubble initiated the world wide downturn in 2001. Its strength, speed and synchronicity in the industrial countries were surprising and could not be explained by a transmission through the trade channel alone. At the same time, financial markets and confidence indicators were negatively affected, and this gave rise to a renewed interest in international business cycle comovements. Besides the traditional trade channel, "new transmission channels" (financial markets and the confidence channel) were suspected to have become more important in the course of the globalization process.

The downturn in the US inspired much research in this area. Business cycle linkages between the US and Germany have been studied by Artis et al. (2006, 2007) and the German Council of Economic Experts (2001, GCEE) who employ small-scale VAR models. In addition, the latter study uses the Deutsche Bundesbank's large-dimensional macroeconometric multicountry model.

Chapter III identifies US supply and demand shocks in a structural dynamic factor model and assesses their propagation to the German economy. This framework allows one to assess not only the response of German GDP to these shocks, but also the responses of prices, interest rates, labor market and other variables contained in the large underlying dataset. A particular emphasis is put on the transmission channels. The advantage of the factor model is that all transmission channels can be accounted for simultaneously; multicollinearity is not a problem in factor models unlike in small-scale regression models. This is also advantageous compared to fully structural models, since it is still unclear how to model the different channels, in particular the "new" channels, in structural models, and factor models provide a reduced form solution.

During the last months, we have again observed extraordinary economic movements in the US, and many fear that the ongoing burst of the housing price bubble and the financial turmoil in the US will have serious consequences for the rest of the world as well. In the concluding chapter of this dissertation, chapter V, we will explain to what extent our analysis from chapter III can help to (tentatively) predict the impact of the current US crisis on Germany.

Example 2: EMU

Comovements and heterogeneity (the other side of the coin) have been extensively discussed in Europe before and after formation of EMU. There is a broad consensus that a number of criteria need to be satisfied before countries can form a monetary union (or new countries can enter an existing monetary union), among them a sufficiently high degree of business cycle synchronization. If business cycles are not synchronized among countries, possibly as a result of asymmetric shocks or differences in the transmission of common shocks due to differences in economic structures and policies, monetary union could be costly. Giving up national monetary policy instruments means that new members could lose important stabilization tools for responding to asymmetric shocks or to an asymmetric transmission of common shocks. A common monetary policy whose mandate is to stabilize aggregate inflation would barely be willing to respond to sustained deviations of an individual member state's inflation rate from equilibrium as long as the aggregate remains unaffected. If analyses had consistently concluded that the costs of a monetary union would outweigh the benefits, i.e. a reduction in transaction costs and uncertainty along with more transparency in the price determination mechanisms, this would have been a strong argument against EMU. At the end, however, business cycles were considered sufficiently synchronized, and other criteria were satisfied as well, and EMU was formed in 1999.

Recently, persistent output growth differentials between the large euro-area economies since the mid-1990s and an increase in inflation differentials observed since 2000 have attracted the policymakers' and academics' attention. Heterogeneity in the euro area is not necessarily harmful and does not automatically call for policy intervention. Output and inflation differentials may partly reflect the catching-up process, in the course of which countries which lower initial incomes experience higher output growth and inflation. In addition, adjustments in individual countries to shocks naturally trigger temporary inflation dispersion. If, however, such adjustments are slow due to nominal rigidities and imperfect factor mobility, this may lead to long-lasting undesirable output and inflation differentials. Heterogeneity may also reflect inappropriate national economic policies or other unwarranted domestic developments. In these cases, if not counteracted by economic policies, heterogeneity may persist and result in large welfare losses for individual countries.

One important aspect of EMU is also that a common monetary policy shock itself may be transmitted in a heterogeneous way to individual member states. The ECB puts a great emphasis on the transmission of monetary policy in Europe. The first research network launched by the ECB shortly after its formation was the Monetary Transmission Network (MTN). This network aimed at bringing together empirical researchers at the ECB and the national central banks. Output of this network emphasized structural and VAR macro-models for the euro area and national economies, panel micro data analyses of the investment behaviour of non-financial firms and panel micro data analyses of the behaviour of commercial banks (as summarized in Angeloni et al., 2003).

Only at a later stage, when a better understanding of dynamic factor models was achieved and when these models gained popularity, were they also used to examine the monetary policy transmission in the euro area. Sala (2003) was the first to apply a structural dynamic factor model to investigate the transmission of a common euro-area monetary policy shock (approximated with a shock to the German short-term interest rate) to eight euro-area countries' key macroeconomic variables between 1985 and 1998. Eickmeier and Breitung (2006) extend his work to the central and east European countries and investigate, how a euro-area monetary policy shock and other aggregate euro-area supply and demand shocks affect them in comparison with the current members in view of the coming EMU enlargement. Chapter IV of this thesis examines the propagation of an even richer set of shocks on current EMU members. That chapter, in addition, uses the Bai and Ng (2004) framework to allow for non-stationarity in common and idiosyncratic components. One important goal is to study the degree of persistence and the sources of heterogeneity. It has become clear from the previous discussion, that this is of great political interest.

3. Organization of the dissertation

The rest of the dissertation is organized as follows. Chapter II studies the forecast performance of factor models by means of a meta-analysis. This chapter summarizes results from existing studies (with – at first glance – ambiguous results) on the forecast performance of factor models and its determinants. Chapter III investigates the business cycle transmission from the US to Germany, with a particular focus on the transmission channels – trade, financial markets and confidence – and specific periods – the economic boom in the US in the second half of the 1990s and the recession in 2001. Chapter IV establishes stylized facts about the cyclical and longer-term comovements and dispersion in the euro area between 1981 and 2003. It then decomposes heterogeneity into its components: idiosyncratic shocks and adjustments to these shocks and the asymmetric transmission of common shocks. Chapter V summarizes the thesis and gives an outlook for future research. Lastly, a technical remark may be appropriate. Chapters II, III and IV can be read in isolation.

II. How Successful are Dynamic Factor Models at Forecasting Output and Inflation? A Meta-Analytic Approach*

1. Introduction

Policymakers and economic research institutions are increasingly turning to large-dimensional dynamic factor models to forecast key macroeconomic variables such as real output and inflation. This is due partly to the ready availability these days of many time series, and modern computers and software allow us to efficiently summarize the information contained in large datasets. The efficacy of dynamic factor models has been further improved by recent advances in estimation techniques proposed by Stock and Watson (2002a; henceforth SW), Forni, Hallin, Lippi and Reichlin (2005b; henceforth FHLR) and Kapetanios and Marcellino (2004; henceforth KM). These techniques allow forecasters to easily summarize the information contained in large datasets and extract a few common factors. The estimated factors are then entered into fairly simple regression models to predict key macroeconomic variables.

Exploiting information from large panels should normally help to improve forecasts. Initial results were very promising (cf. SW, FHLR). However, more recent applications have detected either minimal improvement or none whatsoever (cf. Schumacher, 2006; Schumacher and Dreger, 2004; Gosselin and Tkacz, 2001; Banerjee et al., 2004; Angelini et al., 2001). These conflicting results have launched a lively discussion on whether large-scale factor models are really as useful for forecasting practice as initially expected. In fact, some researchers have speculated about the conditions under which factor models actually perform well in forecasting.

The motivation for this chapter is twofold. First, numerous papers investigating the forecast performance of factor models have now been written, and forecasters from policy institutions are on the verge of integrating factor models into the regular forecasting process. We therefore believe that it is time to systematically summarize this literature. Second, we seek to contribute to the discussion on the determinants of the relative forecast performance of factor models. It is unclear *a priori* how some of the potential determinants affect the factor forecast performance. This question requires an empirical solution, and we seek to identify the

* This chapter is based on Eickmeier, S., C. Ziegler, „How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach“, *Journal of Forecasting*, forthcoming. It benefited from helpful comments by Jörg Breitung, Ben Craig, Ard den Reijer, Heinz Herrmann, Robinson Kruse, Massimiliano Marcellino (the editor), Christian Offermanns, Christian Schumacher and two anonymous referees. This paper was presented at a seminar at the Deutsche Bundesbank, at macroeconometric workshops in Berlin and Halle and the 13th International Conference on Computing in Economics and Finance 2007 in Montréal.

relevant determinants and indicate ways in which factor model forecasts can be further improved.

Our study surveys studies which have used large-scale dynamic factor models to predict real economic activity and inflation. For this purpose, we carry out a meta-analysis. Meta-analyses were initially applied in health, educational and psychological sciences for a while, and have recently found favor in macroeconomics as well (cf. Stanley, 2001, articles in the special edition of the *Journal of Economic Surveys* 2005, Vol. 19(3) and Weichselbaumer and Winter-Ebmer, 2005, the methodology in which is closely related to our analysis). They are powerful tools, based on formal statistical and econometric techniques, with which findings from previous studies can be summarized. The idea is to collect existing studies on a certain issue of interest, extract the appropriate statistics from these studies, examine empirical distributions and regress these statistics on a number of determining characteristics. Our statistic of interest (or meta-dependent variable) measures the relative factor forecast performance, and we take the root mean squared error (RMSE) of a forecast based on a large-scale dynamic factor model relative to the RMSE of a forecast based on a benchmark model. Overall, we collect 50,520 relative RMSEs from a total of 52 studies. We provide some descriptive statistics and test whether factor model forecasts perform significantly better or worse than other models' forecasts. Theory gives us some guidance on possible determinants of the factor forecast performance. We record these determinants for each observation and regress relative RMSEs on them to find out which forecast environments and designs lend themselves to factor forecasts and which do not.

Our study is related to other papers which survey factor forecast applications (among other things) such as Reichlin (2003), Stock and Watson (2006) and Breitung and Eickmeier (2006). Those papers all adopt a narrative approach, which is more prone to subjectivity regarding the choice of papers and results. This chapter is also related to a strand of literature which concentrates on certain aspects of factor forecasting. Examples include Kapetanios and Marcellino (2004) and Schumacher (2007), who explicitly compare factor estimation techniques; Boivin and Ng (2005) and D'Agostino and Giannone (2006), who also concentrate on the implementation of estimated factors in the forecasting equation; and Boivin and Ng (2006), who look at the composition of the dataset from which the factors are extracted. The advantage of our meta-analytic approach is that many possible determinants can be considered simultaneously, not just a few. It has a very broad scope, which allows us to reconcile and explain differences in findings across individual studies.

The chapter is organized as follows. Section 2 presents the approximate dynamic factor model and explains how it is used for forecasting in macroeconomics. It further discusses determinants of the factor forecast performance. Section 3 describes the preparatory work in the run-up to the meta-analysis, including the collection of relevant papers and the construction of the dataset. Section 4 presents some descriptive statistics, the meta-analytic design and the results. Section 5 concludes.

2. Forecasting with dynamic factor models

We consider a situation where a forecaster is interested in predicting a certain macroeconomic (target) variable y_t. S/he may do this by fitting small-scale time series models such as AR or VAR models. These simple models have been shown to perform fairly well in the past. Nowadays, however, lots of data are available to forecasters. Those data may contain information which is useful for predicting y_t.

It is, however, not feasible to include each potentially relevant variable simultaneously in a forecasting equation. And this is where factor models come into play. The idea underlying factor models is that the bulk of variation of many variables can be explained by a small number of common factors or shocks (cf. Burns and Mitchell, 1946). Factor models exploit the variables' comovement and efficiently reduce, in a first step, the dimension of the dataset to just a few underlying factors. In a second step, these factors are entered into a rather small forecasting equation to predict y_t, and only a few parameters need to be estimated. Let us explain these two steps in some detail.

2.1. A two-step forecasting approach

It is assumed that a large number of variables y_{it}, $i=1,...,N$, collected in $\mathrm{Y}_t = \begin{bmatrix} y_{1t} & \cdots & y_{Nt} \end{bmatrix}'$, are driven by few ($q \ll N$) unobserved common factors, summarized in $\mathrm{f}_t^d = \begin{bmatrix} f_{1t}^d & \cdots & f_{qt}^d \end{bmatrix}'$. Accordingly, dynamic factor models express the variables $y_{it}, \forall i$ as the sum of a common component x_{it} and an idiosyncratic component ξ_{it}. The common component is the product of the $q \times 1$ vector of dynamic factors which are common to all variables in the set, f_t^d, (and possibly their lags) and the factor loadings $\Lambda_i(L) = \Lambda_{i0} + \Lambda_{i1}L + ... + \Lambda_{is}L^s$:

$$y_{it} = x_{it} + \xi_{it} = \Lambda_i(L)'\mathrm{f}_t^d + \xi_{it}. \tag{1}$$

The dynamic model has a static representation:

$$y_{it} = \Lambda_i'\mathrm{f}_t + \xi_{it}, \tag{2}$$

where f_t is a vector of $r \geq q$ static factors that comprises the dynamic factors f_t^d and all lags of the factors which enter with at least one non-zero weight in the factor representation. The $r \times 1$ vector Λ_i comprises all non-zero columns of $(\Lambda_{i0},...,\Lambda_{is})$. The factors are orthogonal to each other, and the idiosyncratic components are allowed to be weakly serially and cross-correlated in the sense of Chamberlain and Rothschild (1983) and Bai and Ng (2002).[6]

Three key methods of estimating the factors f_t from a large dataset are known in the literature: those proposed by SW, FHLR and KM. We will explain the different estimation methods in detail below. For the moment, however, let us simply assume that $\hat{\mathrm{f}}_t$ denotes the vector of estimated factors by one of the three methods.[7]

In a second step, the estimated factors or a subset thereof – let us denote them by the $\underline{r}\times 1$ vector $\hat{\mathrm{f}}_t^0$ – are included in a forecasting equation to predict $\tilde{y}_t$ which may or may not be included in Y_t.[8] The equation is usually given by

$$\tilde{y}_{t+h} = \alpha_0 + \alpha_1(L)'\hat{\mathrm{f}}_t^0 + \alpha_2(L)\tilde{y}_t + \varepsilon_{t+h}, \tag{3}$$

where h is the forecast horizon and $\alpha_i(L) = \alpha_{i0} + \alpha_{i1}L + \alpha_{i2}L^2 + ... + \alpha_{ip^{\alpha_i}}L^{p^{\alpha_i}}, i = \{1,2\}$ denotes $\underline{r}$-dimensional ($i = 1$) and scalar ($i = 2$) lag polynomials, respectively.

Papers using dynamic factor models to predict macroeconomic variables differ in various respects and, not surprisingly, come to different conclusions regarding the relative forecasting performance of factor models. Theory provides some guidance as to which conditions and what forecast design should lead to good outcomes.

2.2. Determinants of factor forecasts

From equation (3), it is apparent that AR(1) forecasts can be regarded as a special case of factor forecasts (cf. Boivin and Ng, 2005). Moreover, factor forecasts deliver smaller mean squared errors than AR(1) models as long as the factors and parameters are known, the target variable is correlated with the factors, i.e. $\alpha_1(L) \neq 0$, and the forecasting equation is correctly specified. In practice, however, factors are unknown. They need to be estimated, and it is these factor estimates, and not the true factors, which are included in the forecasting equation. Moreover, the target variable is not necessarily correlated with the factors included in the forecasting equation; introducing uninformative factors in the forecasting equation would only increase sampling variability. Finally, researchers do not know the models' parameters and are not immune to misspecifications in the forecasting equations.

Whether or not forecasters face these problems will depend on the specific forecast environment and the forecast design. While they generally have to take the former as given,

---

[6] For this reason, the model is labelled "approximate factor model". In "strict factor models", idiosyncratic errors are independent.

[7] Here and in the following, a '^' stands for an estimate.

[8] It may be that $\hat{\mathrm{f}}_t^0 = \hat{\mathrm{f}}_t$, i.e. that the factors underlying Y_t are identical to the factors included in the forecasting equation. However, some papers include $\hat{\mathrm{f}}_t^0 \neq \hat{\mathrm{f}}_t$ in the forecasting equation, where $\hat{\mathrm{f}}_t^0$ is a subset of $\hat{\mathrm{f}}_t$ (in this case, $\underline{r} \leq r$). And sometimes r is not estimated and a large number of factors are included in equation (3).

they can choose the latter and, in this respect, have some influence on the outcomes. In the following, we identify the determinants of the relative factor forecast performance and classify them into determinants capturing the forecast environment and those affecting the forecast design. We then discuss implications for the precision of factor estimates, the commonality of the target variable and the specification of the forecasting equation.

2.2.1. Forecast environment

The degree of commonality obviously differs across variables, with some variables linked more closely to overall economic development than others. We will distinguish forecasts of output and inflation in different countries or regions below.

In addition, the relative forecast performance of factor models should vary with the benchmark model, with some models accounting more for the commonality of the target variable than others. Univariate models such as random walks and ARIMA models do not consider the cross-variation between the variables at all. By contrast, other popular benchmarks such as VAR models or single equation models with one or a very few observable indicators do. Moreover, the issue of whether the target variable is more closely related to the factors or to the observable indicators will govern the relative forecast performance. Recently, alternative methods which are able to exploit data-rich environments, such as bagging, ridge regression, shrinkage and partial least squares have been proposed, as well as forecast pooling or model averaging (including Bayesian model averaging) (cf. Stock and Watson, 2004 and 2006; Lin and Tsay, 2005), and we will consider these methods below as well.

Also, the forecast horizon may govern the relative forecast performance of factor models. The greater the predictability of the common component relative to the idiosyncratic component at larger horizons (which is positively related to the relative persistence of the two components), the better the forecast performance of factor models should be relative to small-scale models. Moreover, if Y_t contains leading indicators of $\tilde{y}_t$, which is typically the case, we would expect factor models to be more successful than univariate models at predicting $\tilde{y}_t$ at longer horizons.

2.2.2. Forecast design

Choosing the forecast design mainly means deciding upon the construction of the dataset from which the factors are extracted and the way the factors and the forecasting equation are estimated.

Regarding the dataset, it is well-known from the factor literature that greater precision of the factor estimates is one of the benefits of using information contained in large datasets. Stock and Watson (2002a) and Bai and Ng (2002), for example, show that the uncertainty associated with the factor estimation becomes negligible and factors can be treated as known if the cross-section dimension of Y_t, N, and the number of observations, T, tend to infinity. This result suggests that the forecast performance of factor models should improve with increasing N and T.

However, various studies demonstrate that it is not the pure size of the dataset[9] but also its characteristics which matter for forecasting. Forecasters should make sure that variables which are highly correlated among each other are included in the dataset; this should improve the precision of the factor estimation. Moreover, the dataset should contain variables which are highly correlated with the target variable. Forecasters often face a trade-off between including many time series and/or many observations and time series which satisfy these two requirements. This trade-off is reflected in the following determinants.

One issue is whether it is well suited to extract factors from a balanced or an unbalanced panel. Forecasters relying on an unbalanced panel argue that additional information from time series with missing observations can be exploited. Improved forecasting performance can, however, only be expected if these additional time series comove with the other variables contained in the panel and/or with the target variable.

Another issue is raised by Boivin and Ng (2006). They show that the inclusion of variables with errors which have large variances and/or are cross-correlated should worsen the precision of factor estimates. The forecast performance of factor models may also be worsened if variables that are irrelevant for $\tilde{y}_t$ are included in Y_t – this is referred to as the oversampling problem. Boivin and Ng (2006) suggest pre-selecting the variables to be included in Y_t and removing variables with correlated and/or large errors and/or variables which are irrelevant for the target variable prior to estimating the factors.

A third issue which is relevant in this context is whether to use a recursive or a rolling factor estimation (and forecasting) scheme. A forecast based on a recursive scheme relies on an estimation period of increasing length, where the starting point remains fixed, whereas a rolling scheme relies on a fixed-length window which is shifted every period. It is unclear *a priori* which scheme yields more precise factor estimates. On the one hand, the recursive scheme allows us to exploit more information since estimation tends to be based on a larger T. On the other hand, if factors (and/or factor loadings) are subject to structural breaks which occur during the estimation period, a rolling scheme which gives observations in the past a lower weight compared to a recursive scheme should deliver better forecasts. At the same time, it is often argued that factor models are relatively immune to structural breaks. Stock and Watson (1998) argue that, even if there is mild time variation in the loadings, the factors

---

[9] Watson (2003) and Bai and Ng (2002) show in real-time experiments and in simulations that there are basically no gains from increasing N beyond 50 or 40, respectively. Boivin and Ng (2006) demonstrate that increasing N beyond a certain number can even be harmful and may result in efficiency losses.

are still estimated consistently from large datasets. Whether or not it is better for a factor forecaster to choose a rolling or a recursive scheme is thus an empirical issue.

A fourth issue concerns the frequency. Quarterly time series correspond to averages of the monthly series. If idiosyncratic noise is averaged away rather than the common part of the variables, as one might expect, commonality will be greater at quarterly frequency than at monthly frequency, although monthly data potentially contain more information.

Besides the size and the composition of the dataset, estimation techniques also matter for factor forecasts. The technique used to estimate the factors should affect the precision of the estimates. As already pointed out above, there are three different methods which dominate in the literature: those proposed by SW and FHLR and the more recent, yet less frequently applied KM method. Let us briefly explain them.

SW propose estimating F_t with static PCA applied to Y_t. The factor estimates are simply the first r principal components of Y_t, $\hat{\mathrm{f}}_{\mathrm{t}}^{\mathrm{SW}} = \hat{\Lambda}'\mathrm{Y}_t$, where $\hat{\Lambda}$ is the $N \times r$ matrix of the eigenvectors corresponding to the r largest eigenvalues of the sample covariance matrix $\hat{\Sigma}$.

FHLR propose a weighted version of the principal components estimator suggested by SW, where time series are weighted according to their signal-to-noise ratio and the latter is estimated in the frequency domain. The authors proceed in two steps. First, the covariance matrices of common and idiosyncratic components of Y_t are estimated with dynamic PCA. This involves estimating the spectral density matrix of Y_t, $\Sigma(\omega)$. For each frequency ω, the largest q eigenvalues and the corresponding eigenvectors of $\Sigma(\omega)$ are computed, and the spectral density matrix of the common components $\Sigma_\chi(\omega)$ is estimated. The spectral density matrix of the idiosyncratic components is given by $\hat{\Sigma}_\xi(\omega) = \hat{\Sigma}(\omega) - \hat{\Sigma}_\chi(\omega)$. Inverse Fourier transform provides the time-domain autocovariances of the common and the idiosyncratic components $\hat{\Gamma}_\chi(k)$ and $\hat{\Gamma}_\xi(k)$ for lag k. Since dynamic PCA corresponds to a two-sided filter of the time series, this approach alone is not suited for forecasting. Therefore, in a second step, FHLR compute the r linear combinations of Y_t that maximize the contemporaneous covariance explained by the common factors $\hat{Z}_j{}'\hat{\Gamma}_\chi(0)\hat{Z}_j$, $j = 1,...,r$. This optimization problem is subject to the normalization $\hat{Z}_j{}'\hat{\Gamma}_\xi(0)\hat{Z}_i = 1$ for $i = j$ and 0 for $i \neq j$. It can be reformulated as the generalized eigenvalue problem $\hat{\Gamma}_\chi(0)\hat{Z}_j = \hat{\mu}_j\hat{\Gamma}_\xi(0)\hat{Z}_j$, where $\hat{\mu}_j$ denotes the j-th generalized eigenvalue and $\hat{Z}_j$ its $N \times 1$ corresponding eigenvector. The factor estimates are obtained as $\hat{\mathrm{f}}_t^{FHLR} = \hat{Z}'\mathrm{Y}_t$ with $\hat{Z} = \begin{bmatrix} \hat{Z}_1 & \cdots & \hat{Z}_r \end{bmatrix}$.

KM propose a state-space framework to estimate the factors. The starting point is the prediction error representation $\mathrm{Y}_t = \Lambda \mathrm{f}_t + \Xi_t$, $\mathrm{f}_{t+1} = \mathrm{A}\mathrm{f}_t + \mathrm{L}\Xi_t$, where Ξ_t is the vector of innovations. KM apply a subspace algorithm which allows the factors to be estimated without specifying and identifying the full state space model. The model can be written as a vector equation $\mathrm{Y}_t^f = OK\mathrm{Y}_t^p + \mathrm{E}\Xi_t^f$. $\mathrm{Y}_t^f = \begin{bmatrix} \mathrm{Y}_t' & \mathrm{Y}_{t+1}' & \cdots \end{bmatrix}'$, $\mathrm{Y}_t^p = \begin{bmatrix} \mathrm{Y}_{t-1}' & \mathrm{Y}_{t-2}' & \cdots \end{bmatrix}'$, $\Xi_t^f = \begin{bmatrix} \Xi_t' & \Xi_{t+1}' & \cdots \end{bmatrix}'$, where, in practice, leads and lags need to be truncated, and the matrices O, K and E are complicated functions of the parameters in the prediction error representation. OK is estimated as $(\mathrm{Y}_t^p{}'\mathrm{Y}_t^p)^{+}\mathrm{Y}_t^p{}'\mathrm{Y}_t^f$, where B^{+} denotes the Moore-Penrose inverse of a matrix B. The coefficient can be decomposed by a singular value decomposition $(\mathrm{Y}_t^p{}'\mathrm{Y}_t^p)^{+}\mathrm{Y}_t^p{}'\mathrm{Y}_t^f = \hat{U}\hat{S}\hat{V}'$. The factor estimates are given by $\hat{\mathrm{f}}_t^{KM} = \hat{K}\mathrm{Y}_t^p$, where $\hat{K}_t = \hat{U}_r\hat{S}_r^{1/2}$, and $\hat{U}_r$ denotes the first r columns of the left singular value matrix $\hat{U}$, and $\hat{S}_r^{1/2}$ is the $r \times r$ upper left square matrix of the square root of the singular value matrix $\hat{S}$ containing the largest singular values in descending order.

It is not clear *a priori* which of the three methods will perform best in practice. Weighting the time series according to their signal-to-noise ratios, as is done in FHLR, should deliver efficiency gains compared to the unweighted SW and KM versions. Efficiency gains should also be obtained because the FHLR and KM methods allow for richer dynamics: factors are estimated as linear combinations of contemporaneous time series and their leads and lags, whereas only contemporaneous comovement between variables is accounted for in the approach originally proposed by SW.[10] The SW approach, by contrast, has the advantage of only requiring the estimation of a single auxiliary parameter (r), whereas more unknown parameters have to be set in KM and FHLR,[11] rendering those two approaches more vulnerable to misspecification. Also, if no lagged relationship between y_{it} and f_t exists in the data, unnecessary estimation of the spectral density matrix for the FHLR approach could induce efficiency losses. The KM approach clearly gives more structure to the data than SW and accomplishes this – at least in existing practical applications – by rather restrictive processes assumed for the innovations and the factors.[12] This may be more efficient if the data are well described by this structure. However, overly tight restrictions will lead to less precise KM factor estimates.

Recently, Boivin and Ng (2005) have pointed out, and D'Agostino and Giannone (2006) have discussed, another difference between the approaches of SW and FHLR (the KM approach is disregarded) besides the way in which the factors are estimated. FHLR impose the restrictions implied by the factor model (equation (1)) in the forecasting equation (3), i.e. $\alpha_1(L)$ is a function of the loadings associated with $\tilde{y}_t$ which, in this case, needs to be included in Y_t and the dynamics of the factors and the idiosyncratic components and $\alpha_2(L)$ depend on the latter dynamics.[13] By contrast, SW propose estimating the forecasting equation (3) unrestrictedly with OLS. Again, the impact of imposing or not imposing the restrictions implied by the

[10] An exception is the "stacked" version of the SW method, where factors are estimated as linear combinations of Y_t and its lags. This approach is also used by Grenouilleau (2004).

[11] Besides r, the number of dynamic factors q, the truncation lag parameters for spectral estimation as well as the number of frequency grids needs to be chosen in FHLR. The KM model requires to set r as well as the truncation leads and lags for the subspace algorithm.

[12] Innovations are assumed to be serially uncorrelated. Factors are generally assumed to follow a VAR(1) process.

[13] The nonparametric forecast of FHLR involves predicting the common component as $\hat{\mathrm{x}}_{t+h} = \hat{\Gamma}_\chi(\omega)\hat{Z}(\hat{Z}'\hat{\Sigma}\hat{Z})^{-1}\hat{Z}'\mathrm{Y}_t = \hat{\Gamma}_\chi(\omega)\hat{Z}(\hat{Z}'\hat{\Sigma}\hat{Z})^{-1}\hat{\mathrm{f}}_t^{FHLR}$ and, thus, takes the restrictions implied by the factor model into account.

factor model is unclear and depends strongly on whether or not the factor model is correctly specified and the associated parameters are precisely estimated.

The forecast performance of factor models could also be affected by whether direct or iterated multi-step forecasts are made. Iterated forecasts use a one-period ahead model, iterated forward for h periods, whereas direct forecasts use a horizon-specific estimated model where the dependent variable is the multi-period ahead value being forecasted (Marcellino et al., 2006). Theoretically, iterated forecasts are more efficient if the one-period model is correctly specified, but direct forecasts are more robust to model misspecification. In our context of factor forecasts, iterated forecasts require specifying a dynamic process for the factors (generally a VAR process). They will yield improvements over direct forecasts in terms of smaller relative forecast error losses of factor models if this process is correctly specified.[14]

The commonality of the target variable, the precision of factor estimates and, hence, the relative forecast performance of factor models may also be altered by the specific estimation period. Variables comove more closely in certain periods (economic downturns, for example) than in others, and we would therefore expect factor models to be better suited for forecasting in periods of closer comovement. Likewise, periods which are characterized by lower economic volatility such as the great moderation period have been shown (for instance, by D'Agostino and Giannone, 2006) to be less well described by a factor model. Moreover, as already discussed above, factor models are often held to be more robust against structural breaks. From this perspective, we would therefore expect these models to perform better than smaller models such as AR or VAR models, in periods of structural breaks, something which will, to some extent, already be captured by distinguishing between rolling and recursive forecasts. There are no papers that focus on periods which are clearly characterized as periods with high or low comovement and/or periods with or without structural breaks, and we therefore do not explicitly capture these aspects here.[15]

Factor forecast applications also differ in other respects concerning the specification of the forecasting equation (3), especially in their choice of factors and lags of the factors and dependent variables.[16] There is, however, much heterogeneity among the papers with respect

[14] Some forecasters, in addition, consider forecasts of the idiosyncratic components (cf. Den Reijer, 2005). This also requires specifying a dynamic process for the idiosyncratic components (generally an AR process). And in this case, it will also matter if the process for the idiosyncratic component is correctly specified.

[15] To nevertheless investigate whether the factor forecast performance has changed over time, we included a dummy variable in the regression equation to be described further below which equals 1 if the estimation starts no later than 1990 and 0 otherwise. Results which are available upon request are inconclusive for output and inflation. The coefficient is, in our baseline specification, significantly positive for output, i.e. the forecast performance of factor models tend to have worsened over time, and (insignificantly) negative for inflation.

[16] Some determine r based on formal information criteria (i. e. those proposed by Bai and Ng, 2002) or on other, rather informal, criteria (i. e. the variance shares explained by r factors) and include all estimated factors ($\underline{r} = r$) in equation (3) (cf. Schneider and Spitzer, 2004; Schumacher, 2007; Van Nieuwenhuyze, 2006). Others include only the first factor in the forecasting equation, which is often seen as a measure of the business cycle (see, for example Watson, 2003, or the CFNAI constructed by the Chicago Fed) or core inflation (cf. Camba-

to the factors and lags included in the forecasting equation. Since it would be difficult to classify results meaningfully, we do not address this issue here.

3. Preparing the meta-analysis

Before the meta-analysis can be carried out, much preparatory work needs to be done. We have to collect the relevant papers; choose the independent and dependent meta-variables covering the relative forecast performance of dynamic factor models and its determinants, respectively; and, based on this decision, construct the dataset. Replicability and completeness are important principles in meta-analysis, which we therefore try to follow (Lipsey and Wilson, 2000; Stanley, 2001).

3.1. Collecting relevant papers

We start with an extensive computer search in the EconLit, Google scholar and IDEAS databases and search for empirical studies on macroeconomic forecasting with factor models. The keyword is “forecast” combined with “factor models”, “dynamic factors” “principal components” or “diffusion index”. We also search the working papers series of central banks, the Bank for International Settlements, and the International Monetary Fund, and look at the websites of researchers who are known in the research community as specialists in the field of dynamic factor modeling and forecasting. Some papers’ main focus is the forecast performance of factor models, while others only use them as benchmarks. We concentrate here on studies that forecast real economic activity and inflation.

We include published as well as unpublished papers which comprise working papers and manuscripts in our sample, which enables us to consider as many results as possible. Forecasting with factor models is a relatively new field of research. This is reflected in the fact that only 46% of the papers we consider are already published (or forthcoming) and that most unpublished papers were written up to two or three years ago. Unpublished paper versions of the published papers are generally also available to use. In one case, the unpublished version provides more results than the published version, probably because it has been shortened for publication. In this case, we consider all results reported in the published

---

Méndez and Kapetanios, 2005). Some also consider the first $\underline{r} > 1$ factors, where $\underline{r}$ is chosen somewhat *ad hoc* (Lin and Tsay, 2005; Stavrev, 2006). Bruneau et al. (2007) include the first, the second, etc., each at a time, in equation (3) to assess the marginal contribution of each of the factors to the forecast. Most papers, however, set a maximum number of factors and lags (of the dependent variable and/or the factors) and determine the factors and the lags to be included in the forecasting equation simultaneously using Akaike or Bayesian information criteria (cf. Matheson, 2006; Stock and Watson, 1999, 2003; Banerjee et al., 2005, 2006a; Artis et al., 2004; Jeon, 2004) or performance-based measures such as the mean squared error (cf. Schumacher, 2007; Forni et al., 2001, 2003). Others do not consider lags of the factors (Schumacher, 2007; Giacomini and White, 2006; Forni et al., 2001; Boivin and Ng, 2006) and/or autoregressive terms from the outset (cf. Liu, 2004; Stock and Watson, 1998; Tatiwa Ferreira et al., 2005).

Table 1: Studies included in the meta-analysis

Paper	# obs	Variable	Country	Factor estim. method
Aguirre, Céspedes (2004)	24	Output, inflation	CHL	SW
Angelini, Henry, Mestre (2001)	4050	Inflation	EA	SW
Artis, Banerjee, Marcellino (2005)	839	Output, inflation	UK	SW
Banerjee, Marcellino (2006)	2470	Output, inflation	US	SW
Banerjee, Marcellino, Masten (2005)	120	Output, inflation	EA	SW
Banerjee, Marcellino, Masten (2006a)	2805	Output, inflation	CZ, HU, PL, SK, SI	SW
Banerjee, Marcellino, Masten (2006b)	1719	Output, inflation	EA, SI	SW
Boivin, Ng (2005)	240	Output, inflation	US	SW, FHLR
Boivin, Ng (2006)	144	Output, inflation	US	SW
Brisson, Campell, Galbraith (2003)	159	Output, inflation	CAN	SW
Bruneau, de Bandt, Flageollet (2003)	124	Inflation	EA	SW
Bruneau, de Bandt, Flageollet, Michaux (2007)	471	Inflation	FRA	SW
Camacho, Sancho (2003)	24	Output, inflation	ESP	SW
Camba-Méndez, Kapetanios (2005)	768	Inflation	EA, DEU, ESP, FRA, ITA, NLD	KM
Cheung, Demers (2007)	1224	Output, inflation	CAN	SW, FHLR
Cristadoro, Forni, Reichlin, Veronese (2005)	84	Inflation	EA	FHLR
D´Agostino, Giannone (2006)	558	Output, inflation	US	SW, FHLR
De Mol, Giannone, Reichlin (2006)	714	Output, Inflation	US	SW
Den Reijer (2005)	48	Output	NLD	SW, FHLR
Favero, Ricchi, Tegami (2004)	156	Inflation	ITA	SW
Forni, Hallin, Lippi, Reichlin (2002)	16	Output, inflation	EA	FHLR
Forni, Hallin, Lippi, Reichlin (2003)	96	Output, inflation	EA	SW, FHLR
Giannone, Matheson (2006)	54	Inflation	NZL	SW, FHLR
Gavin, Kliesen (2006)	120	Output, inflation	US	SW
Giaccomi, White (2006)	36	Output, inflation	US	SW
Gosselin, Tkacz (2001)	12	Inflation	CAN	SW
Grenouilleau (2004)	3	Output	EA	SW
Hofmann (2006)	1701	Inflation	EA	SW
Inoue, Kilian (2007)	521	Inflation	US	SW
Jeon (2004)	840	Output, inflation	CAN, FRA, DEU, JPN, UK, US	SW
Kabundi (2004)	2	Output	FRA	FHLR
Kapetanios, Labhard, Price (2005)	864	Inflation	UK	SW
Kapetanios (2004)	16	Inflation	UK	KM
Lin, Tsay (2005)	9360	Output	US	SW

Table 1: Studies included in the meta-analysis cont.

Paper	# obs	Variable	Country	Factor estim. method
Liu (2004)	480	Output, inflation	US	SW
Marcellino, Stock, Watson (2001, 2003)	1767	Output, inflation	EA, AUT, BEL, FIN, FRA, DEU, IRE, ITA, LUX, NLD, PRT,	SW
Matheson (2006)	4928	Output, inflation	NZL	SW
Moser, Rumler, Scharler (2007)	12	Inflation	AUT	SW
Schneider, Spitzer (2004)	6	Output	AUT	FHLR
Schumacher (2007)	188	Output	DEU	SW, FHLR, KM
Schumacher, Dreger (2004)	48	Output	DEU	SW
Stavrev (2006)	2778	Inflation	EA	FHLR
Stock, Watson (1998)	410	Output, inflation	US	SW
Stock, Watson (1999)	6184	Inflation	US	SW
Stock, Watson (2002a)	864	Output, inflation	US	SW
Stock, Watson (2002b)	18	Output	US	SW
Stock, Watson (2004)	2028	Output	CAN, FRA, DEU, ITA, JPN, UK, US	SW
Stock, Watson (2005a)	336	Output, inflation	US	SW, FHLR
Stock, Watson (2006)	48	Output	US	SW, FHLR
Tatiwa Ferreira, Bierens, Castelar (2005)	20	Output	BRA	SW
Van Nieuwenhuyze (2006)	6	Output	BEL	FHLR
Watson (2003)	20	Output, inflation	US	SW

Notes: # obs refers to the number of observations before outlier removal. We refer to the unpublished version of Forni, Hallin, Lippi, Reichlin (2002) since no forecast performance measures were reprted in the published version. Abbreviations are EA: euro area, AUT: Austria, BEL: Belgium, FIN: Finland, FRA: France, DEU: Germany, ITA: Italy, IRE: Ireland, LUX: Luembourg, NLD: Netherlands, PRT: Portugal, ESP: Spain, NZL: New Zealand, CHL: Chile, BRA: Brazil, CAN: Canada, JPN: Japan, US: United States, UK: United Kingdom, CZ: Czech Republic, HU: Hungary, SI: Slovenia, PL: Poland, SK: Slovakia.

version plus those in the unpublished version which have not already been taken from the published version. We consult a total of 52 studies (listed in Table 1), with published and unpublished versions being counted as one paper.

3.2. Meta-dependent variable

An important decision to be made is on the dependent variable of our meta-regression. This variable is supposed to measure the forecast performance of a factor model relative to a benchmark model. Such a measure is ideally contained in all collected papers. Most studies report forecast error losses such as mean squared errors (MSE) or root mean squared errors (RMSE) of models used in these studies to predict a certain target variable or forecast error losses of more complex models relative to simple benchmark models such as random walks or AR models. We decide to focus on the RMSE of factor models (*DFM*) relative to the RMSE of a certain benchmark model (*Bench*). Let us denote this ratio associated to observation $j = 1,...,\widetilde{n}$ as ψ_j;

$$\psi_j = \left(\frac{RMSE^{DFM}}{RMSE^{Bench}} \right)_j , \tag{4}$$

where an observation refers to a result in a study and benchmark models can differ across papers. Let us summarize all observations in $\widetilde{\Psi} = [\psi_1 \quad ... \quad \psi_j \quad ... \quad \psi_{\tilde{n}}]$, where $\tilde{n}$ equals the total number – 50,520 – of observations. Results in the papers that were not already defined as in equation (4) were converted. To avoid any selection bias, we tried to include all results.

$\widetilde{\Psi}$ contains some large outliers. As shown by Rousseuw and Leroy (1987), outliers can distort parameter estimates. We therefore remove observations outside 1.5 times the interquartile range. As we will show, most results are robust with respect to other outlier correction methods. After outlier removal, we are left with $n = 43,389$ observations, 20,540 of which refer to output and 22,849 to inflation. Let us denote the outlier-adjusted $n \times 1$ vector of relative RMSEs by Ψ.

3.3. Meta-independent variables

As discussed in the previous section, theory gives us some guidance as to which variables may determine the forecast performance of factor models. In addition to these variables, we consider variables which capture the publication strategy. In the following, we list and briefly explain the meta-independent variables. Some of them are continuous variables, others discrete. The latter can be divided into certain cases which are given in parentheses.

- VARIABLE (OUTPUT, INFL) captures the economic meaning of the target variable $\tilde{y}_t$. We distinguish between variables of real economic activity (OUTPUT) and those of inflation (INFL). Real economic activity includes GDP, industrial production, employed persons, hours and unemployment, retail sales, real personal income, real manufacturing trade and sales, consumption, investment, inventories and orders (in levels or first differences). Inflation measures are consumer prices, producer prices, retail prices and other sub-aggregates, wages and the GDP deflator (in first or second differences) and measures of core inflation.
- COUNTRY (US, UK, EA, OTHER) is divided into four groups of predictions: those associated with the US, the UK, the euro area as an aggregate plus individual euro-area countries (EA) and other countries (OTHER). The latter group contains results for Canada, New Zealand, Japan, Brazil, Chile and central and east European countries.
- BENCH (ARIMA, RW, VAR, INDIC, LARGE) captures the benchmark models to which large factor models are compared. We distinguish between random walks (RW), ARIMA models where most often AR models are employed, VAR models and single equation models with indicators (INDIC) which are similar to equation (3), where $\hat{\mathrm{f}}_t^0$ is replaced with one or few measurable indicators. Notice that INDIC also comprises some structural

models such as the Phillips curve, which is often used to predict inflation.[17] More recently, large-scale factor models, which pool data, have also been compared with the pooling of forecasts or model averaging (FOREC POOL) as well as with other methods suited to exploiting data-rich environments such as ridge regression, partial least squares and shrinkage. We summarize results from the latter methods in OTH LARGE. Very recent methods, such as factor forecast pooling, have been shown to deliver good forecasts, for instance by Koop and Potter (2004) and Stock and Watson (2006), but are beyond the scope of our analysis.

- HORIZON refers to the forecast horizon. Months were converted to quarters.
- N, the dimension of the cross-section (in logs).
- T, the time dimension of the sample on which the estimation is based (in logs). Note that, when a recursive forecasting scheme is applied, T varies over time. In this case, we compute the average T, given by $\min\{T\} + (\max\{T\} - \min\{T\})/2$.
- BAL (YES, NO) reflects whether the factors are estimated from a balanced (YES) or an unbalanced panel (NO).
- PRESEL (NO, 1, 2). This variable distinguishes whether authors use all the data they have collected to extract the factors (NO) or make a pre-selection, either by removing data with correlated errors and/or errors that have a large variance (1)[18] or by removing variables which they think are irrelevant for the target variable or including potentially relevant subgroups of variables which are formed based on economic considerations (2)[19]. Some of the papers which fall under case (2) not only aim to solve the oversampling problem and remove variables which are irrelevant for the target variables, but also focus on the predictive content of certain groups of variables irrespective of the overall forecast performance of the factor model. Unfortunately, these two aspects cannot be regarded separately. This needs to be kept in mind when interpreting results. Note also that, in a

[17] Other structural models include inflation indicators derived from an SVAR and a Blanchard-Quah decomposition and from a P* model into the forecasting equation (Stavrev, 2006).

[18] This case (1) contains observations taken from papers such as Boivin and Ng (2006), Banerjee et al. (2006b), Schneider and Spitzer (2004), Den Reijer (2005), Van Nieuwenhuyze (2006), Matheson (2006), and Bruneau et al. (2007), with the first five papers focusing on output and the first two and the last two papers on inflation. Boivin and Ng (2006) and Banerjee et al. (2006b) exclude series with large and correlated errors from the dataset. Schneider and Spitzer (2004), Den Reijer (2005) and Matheson (2006) order the variables according to their correlation with the target variables; Schneider and Spitzer (2004) and Den Reijer (2005) then sequentially include them into the dataset to minimize the forecast error loss, and Matheson (2006) adopts a *ad hoc* approach by including the first 5, 10 and 50% in Y_t. An *ad hoc* approach is also taken by Van Nieuwenhuyze (2006) who considers only the 75% of the variables with the highest commonality ratio. As an alternative, he also selects a dataset which maximizes the commonality ratio of the target variable. Bruneau et al. (2007) estimate their factor model based on a dataset which includes those indicators which, individually, delivered MSEs relative to AR-MSEs of significantly below 1.

[19] This case (2) includes observations from papers such as Angelini et al. (2001) who extract factors from a nominal and a non-nominal dataset separately to forecast inflation, Bruneau et al. (2003) who use purely French and Belgian factors to predict euro-area inflation, FHLR who investigate whether financial factors help to predict output and inflation, Inoue and Kilian (2007) who assess whether real variables help to predict inflation, etc.

way, some (at least implicit) form of pre-selection always occurs when forecasters construct their large datasets. However, only when variables are included or excluded from the original dataset after applying some formal or explicitly specified criteria do we attribute the resulting observations to cases 1 or 2; otherwise, we assign them to NO.

- ROLREC (ROL, REC) captures whether a rolling (ROL) or a recursive (REC) forecasting scheme is adopted.
- FREQ (Q, M) captures whether the forecast is made on a monthly (M) or a quarterly (Q) basis.[20]
- FACTOR (SW, FHLR, KM) distinguishes between the different factor estimation techniques.
- RESTR (YES, NO) captures whether restrictions implied by the factor model are imposed on the forecasting equation (YES) or not (NO).
- ITDIR (IT, DIR, 1_STEP) states whether a direct (DIR), an iterated (IT) multi-step forecast or a one-step ahead forecast (1_STEP) is made.
- PUBL (YES, NO) reflects whether a paper has already been published or is still a working paper or a manuscript. In meta-analyses, this variable generally captures possible publication bias, where journals' editors have a tendency to publish significant results. It is, however, not clear how this translates to our context. Some meta-studies also use this variable to capture differences in quality and weight observations accordingly (cf. Weichselbaumer and Winter-Ebmer, 2005). These studies presume that published studies which have gone through a rigorous referee process are qualitatively better in the sense that errors are eliminated and only accurate and robust results survive. However, we have our doubts as to whether this argument applies to our context. As pointed out above, factor forecasting is a relatively new field of research and many papers have simply not yet been published due to long publication lags. Although the interpretation of PUBL is not fully clear, we will keep this variable in our regression.
- AUTHOR (YES, NO) captures whether one (or more) of the authors of the particular study was (were) among the developers of (one of) the dynamic factor model(s) used in that particular study, namely Stock, Watson, Forni, Hallin, Lippi, Reichlin, Kapetanios and Marcellino. The hypothesis we test is whether results are biased in favor of factor models when produced by the developers of the model, who may be interested in seeing their models widely applied. Whenever authors not only focus on one factor estimation technique but compare different factor estimation techniques, and developers of one of the applied models are among the authors, only observations that refer to the model which

[20] SW also show how datasets with mixed frequencies can be exploited. Schumacher and Breitung (2006) use this method to predict German output. Their results are not considered here, however.

was also developed by (one of) the author(s) are attributed to YES. Observations associated with other factor estimation techniques are assigned to NO.

- FOCUS_DFM (YES, NO) captures another possible source of bias.[21] Some papers may provide only the best forecasting results for a particular method since they wish to show its potential advantages, whereas other papers show a wide range of results, including the worst ones, in order to check for robustness against parameter misspecifications or compare different methods over a wide parameter grid. In the latter case, average results may be rather poor even if the method performs well. We divide the papers into those focusing on factor models (YES) (which may have an interest to show the good performance of factor models) and those not focusing on factor models (NO). Some of the latter papers focus on other models and use factor models only as a benchmark; others primarily aim at producing a good forecast and use various models, including factor models to achieve this goal.

Two remarks are in order. First, there is certainly some overlap between the meta-independent variables. For example, FACTOR_FHLR implies a weighting of time series where weights are inversely related to the variance of idiosyncratic components. This idea is also captured by PRESEL_1, where data with important idiosyncratic components are either downweighted or dropped (cf. Stock and Watson, 2006; D'Agostino and Giannone, 2006). In order to disentangle the two variables, we include in PRESEL_1 only cases where data are completely eliminated from the dataset, leading thus to weights of 0 or 1. Another difficulty arises, for example, when disentangling different benchmark models. In most cases, deciding on which indicators to include in BENCH_INDIC involves some pre-testing. If indicators are chosen out of a very large set of variables, this also means that information from a data-rich environment is exploited. In fact, Stock and Watson (2005a) and Lin and Tsay (2005) attribute single equation models with indicators where the latter are selected from a large dataset to the class of large predictors. We could have included parts of these models in BENCH_LARGE as well, and acknowledge that our choice is somewhat *ad hoc*. A final example is overlap between HORIZON and ITDIR where we distinguish between one-step ahead (ITDIR_1_STEP) and multi-step forecasts (ITDIR_IT and IT_DIR). This, however, is inevitable if we want to assess the impacts of iterated and direct forecasts, given that this distinction only applies to multi-step predictions.

Second, although important for readers who aim at understanding and perhaps even replicating results, the designs of the analyses are sometimes insufficiently documented. Whenever some of the characteristics we consider in our meta-analysis were missing from a paper, we sent an e-mail to (one of) the author(s). Although the authors were generally very helpful and we obtained responses very quickly, we were not able to fill all the gaps. Whenever an observation could not be related to the characteristics used as meta-independent

[21] We are grateful to an anonymous referee for pointing this possible bias out to us.

variables, we used this observation for the descriptive analysis but had to exclude it from the meta-regression analysis. Overall, the baseline meta-regressions for output and inflation are based on 18,782 and 20,263 observations, respectively.

4. Results

4.1. Descriptive analysis

This subsection presents descriptive statistics of relative RMSEs obtained from factor forecasts associated with the total sample (after removing outliers) and, separately for output and inflation, the different countries/country groups and benchmark models. We test whether means and medians of relative RMSEs differ significantly from 1 using a *t*-test and a Wilcoxon sign rank test, respectively. Empirical distributions for the entire sample and for output and inflation separately are shown in Figure 1. Table 2 provides descriptive statistics.

The means for the entire sample, inflation and output are roughly 0.97, 0.98 and 0.96, respectively; the medians are slightly higher. Although these numbers (means and medians) are only slightly below 1, the tests indicate that, on average, factors models perform significantly better than the respective benchmark models in predicting output and inflation. Lower relative RMSEs associated with inflation than with output go, however, along with more mass in the tails of the empirical distribution of inflation compared to output, as is apparent from the lower panel of Figure 1.

When looking separately at relative RMSEs of different benchmark models, it turns out that factor models generally outperform small-scale models, with one exception: factor models do worse than ARIMA models when predicting inflation. Factor models show the greatest improvements over random walks. Interestingly, forecast pooling significantly outperforms the pooling of data (i.e. factor models).

Likewise, alternative methods which are suited to exploit large datasets (OTH LARGE) do better than factor models when comparing the arithmetic means, however, no significant difference between these two models is found when comparison is made between the medians.

Descriptive statistics for individual countries/country groups clearly indicate that factor models outperform other models on average for the US and the euro area. This holds for both output and inflation. Interestingly, relative RMSEs associated with factor forecasts of euro-area inflation are much lower than relative RMSE associated with euro-area output. They are also lower than relative RMSEs associated with factor forecasts of US inflation. By contrast, the relative forecast performance of factor models for US output is better than for euro-area output. Our findings thus support the assertion put forward by Banerjee et al. (2005) that factor models are better at predicting nominal variables in the euro area compared to the US and real variables in the US compared to the euro area. Results for the UK are not the same for output and inflation; the mean of the relative RMSEs is slightly below 1 (the median does not differ significantly from 1) for output but exceeds 1 for inflation.

The descriptive analysis masks the fact that the various meta-independent variables may interfere with one another. To disentangle the effects, a regression approach is adopted in the next subsection.

4.2. Meta-regression

4.2.1. Baseline meta-regression

We estimate the meta-regression equation

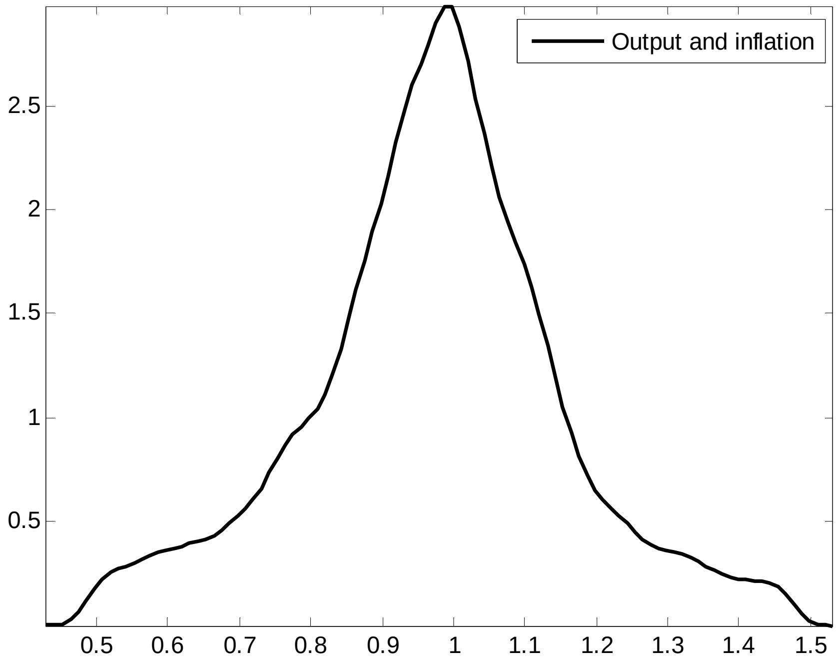
$$\psi_j = \mu + \phi' \mathrm{M}_j + \eta_j , \tag{5}$$

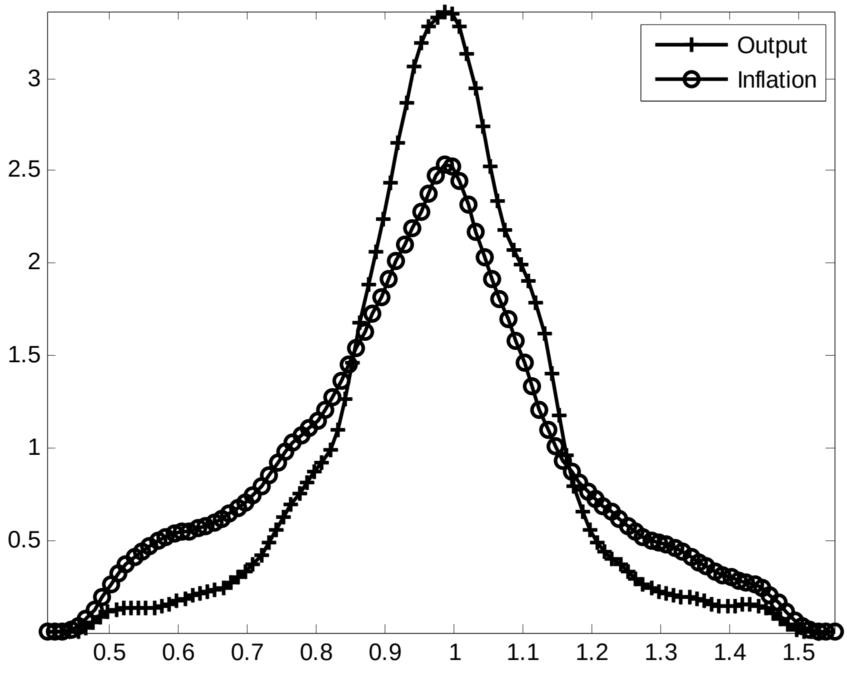
where M_j is the vector of explanatory (meta-independent) variables associated with observation j, ϕ is the corresponding vector of coefficients and μ refers to the overall constant. M_j comprises the continuous variables N, T and HORIZON (as deviations from their arithmetic means) as well as a set of dummy variables into which the discrete variables were transformed. Consider, for example, the variable FACTOR. The dummy variables for the cases FHLR and KM take values of 1 if f_t was estimated with the FHLR and the KM technique, respectively, and 0 otherwise. To avoid perfect collinearity, the SW case is omitted. Negative/positive signs of the coefficients of the included dummies indicate lower/higher relative RMSEs, i.e. a better/worse relative factor forecast performance compared to the omitted cases. The impacts of the omitted cases are summarized in the common intercept, which, hence, can be interpreted as the average relative RMSE conditional on the characteristics given by the omitted cases and on the means of the continuous variables.[22] We estimate equation (5) separately for output and inflation, which is suggested by an F-test.[23] In our baseline regression, we use OLS and report White-corrected standard errors to account for heteroskedasticity in the errors.[24]

[22] The constant and the variables' coefficients can then be used to compute the means of relative RMSEs conditional on any characteristics of interest to the reader.

[23] The F-test is constructed by augmenting equation (5) with interaction terms: $\psi_j = \mu + \phi' \mathrm{M}_j + \varphi'(D_j \widetilde{\mathrm{M}}_j) + \eta_j$, where D_j is the scalar dummy variable for output and $\widetilde{\mathrm{M}}_j$ corresponds to M_j, but does not contain the dummy for inflation (which is included in M_j). We estimate this equation and test the null that all elements of φ equal 0, i.e. equation (5) can be applied to the entire sample. We obtain $F(J, n-\tau) = 44.0969$ with degrees of freedom

Figure 1: Smoothed histograms of relative RMSEs





Note: after outlier adjustment.

$J = 25$ and $n - \tau = 38,993$, where τ is the number of freely estimated parameters, and therefore strongly reject the null.

[24] We also estimated equation (5) with $\log(\psi_i)$ instead of ψ_i on the left hand side to allow for negative values of the dependent variable. Results remained unaffected. Out of concerns about multi-collinearity, we also removed variables contained in M_j one by one. The coefficients and the significance level of the remaining variables remain basically the same.

Table 2: Descriptive statistics

	Mean		Median		Std.	Min.	Max.	# obs
			Output and inflation					
Total	0.973	***	0.980	***	0.180	0.486	1.477	43,389
ARIMA	0.987	***	0.990	***	0.191	0.486	1.477	9,649
RW	0.903	***	0.860	***	0.191	0.490	1.471	682
VAR	0.964	***	0.958	***	0.203	0.487	1.475	3,049
INDIC	0.925	***	0.936	***	0.194	0.486	1.475	15,486
FOREC POOL	1.040	***	1.034	***	0.155	0.486	1.475	6,510
OTH LARGE	1.004	***	0.993		0.102	0.648	1.473	8,013
US	0.969	***	0.972	***	0.133	0.486	1.475	20,736
UK	1.031	***	1.023	***	0.152	0.487	1.474	2,070
EA	0.933	***	0.946	***	0.224	0.486	1.475	12,006
OTHERS	1.024	***	1.022	***	0.199	0.487	1.477	8,577
			Output					
Total	0.984	***	0.985	***	0.152	0.486	1.477	20,540
ARIMA	0.957	***	0.965	***	0.184	0.487	1.477	4,627
RW	0.878	***	0.843	***	0.164	0.490	1.265	297
VAR	0.941	***	0.917	***	0.188	0.488	1.475	1,474
INDIC	0.939	***	0.934	***	0.182	0.487	1.475	2,264
FOREC POOL	1.025	***	1.026	***	0.143	0.486	1.475	4,157
OTH LARGE	1.003	**	0.994		0.096	0.655	1.473	7,721
US	0.978	***	0.980	***	0.130	0.486	1.475	12,338
UK	0.987	**	1.000	*	0.142	0.487	1.453	746
EA	0.987	***	0.996	***	0.171	0.488	1.475	3,103
OTHERS	0.997		0.996	*	0.190	0.487	1.477	4,353
			Inflation					
Total	0.963	***	0.971	***	0.202	0.486	1.476	22,849
ARIMA	1.015	***	1.010	***	0.194	0.486	1.476	5,022
RW	0.922	***	0.864	***	0.207	0.530	1.471	385
VAR	0.985		0.993	**	0.214	0.487	1.475	1,575
INDIC	0.922	***	0.937	***	0.196	0.486	1.475	13,222
FOREC POOL	1.066	***	1.053	***	0.171	0.486	1.474	2,353
OTH LARGE	1.026	**	0.983		0.197	0.648	1.472	292
US	0.956	***	0.959	***	0.138	0.486	1.474	8,398
UK	1.055	***	1.041	***	0.152	0.637	1.474	1,324
EA	0.915	***	0.914	***	0.237	0.486	1.475	8,903
OTHERS	1.051	***	1.060	***	0.204	0.487	1.476	4,224

Notes: ***/**/* indicates that values are significantly different from 1 at the 1/5/10 percent level. This is tested using a t-test for the means and a Wilcoxon-sign rank test for the medians.

The results of our baseline regression equations for output and inflation are given in the first four columns of Tables 3 and 4. While specification (1) refers to a one-time estimation of equation (5), specification (2) provides estimation results where insignificant variables are sequentially removed from the set of meta-independent variables[25]. Variables remaining in the equation in specification (2) have coefficients that are very similar to the corresponding coefficients in specification (1). To address concerns about the limitations of meta-analyses,

[25] The variable with the lowest *t*-statistic is removed from the set of meta-independent variables, and relative RMSEs are, again, regressed on the reduced set of meta-independent variables. This is repeated as long as only variables which are significant at the 5% level are left in the equation.

we also report the means of the dummy variables associated with our baseline specification (1) in the last columns of Tables 3 and 4. Small numbers indicate that coefficient estimates were obtained based on only relatively few results associated with these cases and that results may be affected once new papers will be included in (an updated version of) our meta-analysis.

As regards the meta-independent variables capturing the forecast environment, it turns out that, in line with the descriptive statistics, factor models tend to be better (worse) at predicting US than euro-area real (nominal) variables, although differences between US and euro-area inflation are insignificant. Factor forecasts of British variables are outperformed by factor forecasts of US variables. The coefficients of the variables capturing the different benchmark models suggest, in line with the descriptive statistics, that factor models generally perform relatively better in comparison to small-scale models than in comparison to pooled forecasts or to alternative methods suited to handle large datasets.[26]

Factor models seem to perform better at longer horizons than at shorter horizons for inflation, but the opposite holds for output (although the effects are weak economically with coefficients close to zero). The latter result is surprising given findings by Giannone et al. (2002) for the US and Altissimo et al. (2001) for the euro area. According to these studies, the common component of output is much more persistent than its idiosyncratic component. Coefficients of HORIZON remain the same if ITDIR is excluded from the equations and if relative RMSEs are only regressed on HORIZON and a constant.

The coefficients of N and T are small and tend to be negative, i.e. the size of the dataset has a favorable impact on the relative forecast performance of factor models, except for the coefficient of T associated with output which is positive, but insignificant. Regression results with respect to BAL are inconclusive for output and inflation: factor forecasts of output/inflation based on balanced/unbalanced panels outperform forecasts based on unbalanced/balanced panels. Results regarding the variable PRESEL are unexpected. The removal of variables with correlated errors and/or errors with large variances (PRESEL_1) as well as the removal of variables based on economic considerations (PRESEL_2) worsen the performance of factor models, although not or barely significantly in the case of output. The positive and significant coefficients for inflation are difficult to interpret. A possible explanation for the positive sign of PRESEL_2 is that not only papers which try to solve the oversampling problem discussed above, but also papers which aim at assessing the predictive content of certain groups of variables, yet whose main focus is not to improve the overall

[26] The coefficient of BENCH_OTH LARGE and associated to inflation is very large. All results are taken from one single study (De Mol et al., 2006).

forecast performance of the factor model fall in this category. However, the positive sign of PRESEL_1 is puzzling.[27]

Another result is that output forecasts based on a recursive forecasting scheme perform better than forecasts based on a rolling scheme. This suggests gains from using long time series and also that factor models may be relatively well suited to cope with structural breaks as emphasized by Stock and Watson (1998). ROLREC_ROL does not enter the inflation equation significantly. It should be emphasized, however, that only 5% of the results were obtained based on a rolling forecasting scheme; see the last columns of Tables 3 and 4.

Factor models tend to perform relatively well when forecasters rely on monthly rather than quarterly data, although FREQ_Q is insignificant for inflation in specification (1). For output, relative RMSEs are 4 percentage points lower when forecasts are made on a monthly rather than on a quarterly basis. This points to exploiting extra information contained in monthly data.

The FHLR and KM factor estimation methods outperform the SW method for output. On average, FACTOR_SW yields relative RMSEs which exceed relative RMSEs produced based on FACTOR_FHLR and FACTOR_KM techniques by 4 and 18 percentage points, respectively, in specification (1). For inflation, by contrast, there are no significant differences between FACTOR_KM and FACTOR_SW, and we find advantages of FACTOR_SW over FACTOR_FHLR. We carefully conclude that it may be worthwhile to account for the dynamic relationships between the variables and that efficiency gains which may result in more precise factor estimation can be obtained from weighting time series according to their signal-to-noise ratios, but gains from using more complex factor estimation methods are not guaranteed. The literature is also inconclusive: on the one hand, Forni et al. (2003) find that their estimation approach outperforms the SW factor estimation approach for output and inflation, and Schumacher (2006) finds modest improvements of the FHLR and the KM factor

[27] Interestingly, when among the (only four) papers, which pre-select the dataset before predicting inflation – and we refer to PRESEL_1 here – only the Boivin and Ng (2006) paper or only the Banerjee et al. (2006b) paper is left in the dataset, the coefficient of PRESEL_1 turns significantly negative. Likewise, the coefficient of PRESEL_1 for output turns significantly negative when, among the papers, which pre-select the dataset before predicting output, only Van Nieuwenhuyze (2006), Schneider and Spitzer (2004), Den Reijer (2005), Boivin and Ng (2006) or Banerjee et al. (2006b) is left in the dataset (the coefficient is significant at the 10% level for the latter study).

Table 3: Meta-regression results for output

	Baseline				Outliers				Sampling bias				Dependency				Quality		
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		
	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	mean
CONSTANT	0.93	95.14	0.93	169.37	0.97	83.06	1.08	40.14	0.99	76.74	0.98	98.63	0.93	12.75	0.93	8.40	0.88	66.68	-
COUNTRY_UK	0.02	2.07	-	-	-0.01	-1.17	-0.07	-3.30	0.02	1.95	0.02	2.89	0.02	1.12	0.02	1.21	0.02	1.39	0.037
COUNTRY_EA	0.05	7.84	0.05	10.66	0.06	7.41	0.06	3.13	0.05	6.07	0.05	7.90	0.05	2.63	0.05	2.51	0.06	5.93	0.153
COUNTRY_OTHER	0.01	1.54	-	-	0.02	2.71	0.04	1.57	0.02	2.09	0.01	1.14	0.01	0.46	0.01	0.66	0.01	0.61	0.225
BENCH_RW	-0.05	-4.39	-0.05	-4.82	-0.06	-5.22	-0.07	-2.49	-0.02	-1.82	-0.01	-0.53	-0.05	-1.37	-0.05	-1.49	-0.05	-3.56	0.016
BENCH_VAR	-0.02	-3.58	-0.02	-3.41	-0.03	-3.65	-0.07	-5.85	-0.03	-3.26	-0.02	-3.23	-0.02	-2.28	-0.02	-1.87	-0.02	-2.06	0.071
BENCH_INDIC	-0.01	-1.72	-	-	-0.02	-3.14	-0.05	-3.73	-0.02	-2.49	-0.00	-0.38	-0.01	-0.90	-0.01	-1.16	-0.01	-0.81	0.088
BENCH_FOREC POOL	0.07	15.41	0.07	16.31	0.08	15.30	0.11	9.23	0.05	10.12	0.05	11.30	0.07	2.19	0.07	1.59	0.07	12.75	0.195
BENCH_OTH LARGE	0.06	15.81	0.06	16.55	0.07	16.48	0.05	3.88	0.05	10.28	0.04	10.74	0.06	2.58	0.06	1.51	0.06	12.58	0.411
HORIZON (demeaned)	0.01	17.08	0.01	17.48	0.02	20.69	0.03	15.32	0.01	15.14	0.01	16.98	0.01	2.69	0.01	1.70	0.01	11.77	-
log(N) (demeaned)	-0.01	-3.26	-0.01	-5.12	-0.01	-4.78	-0.02	-4.78	-0.01	-2.74	-0.01	-4.94	-0.01	-0.82	-0.01	-0.52	-0.01	-2.32	-
log(T) (demeaned)	0.00	0.42	-	-	-0.04	-5.71	-0.14	-9.02	-0.03	-3.76	-0.02	-3.03	0.00	0.07	0.00	0.05	0.03	3.90	-
BAL_NO	0.02	2.30	0.01	2.22	0.01	1.66	0.03	2.01	0.01	0.84	0.02	3.62	0.02	2.34	0.02	1.48	0.03	3.51	0.171
PRESEL 1	0.01	1.97	-	-	0.01	0.79	0.00	0.26	0.02	1.61	0.02	2.35	0.01	0.75	0.01	0.69	0.02	1.82	0.087
PRESEL 2	0.02	1.27	-	-	-0.02	-1.23	-0.05	-1.31	-0.02	-1.13	0.02	1.81	0.02	0.56	0.02	0.48	0.03	2.05	0.007
ROLREC_ROL	0.05	5.63	0.04	4.99	0.04	4.56	0.03	1.40	0.04	3.92	0.07	7.89	0.05	3.00	0.05	4.44	0.06	5.35	0.047
FREQ_Q	0.04	3.11	0.04	8.38	-0.04	-2.74	-0.19	-6.48	-0.02	-1.29	0.01	0.95	0.04	0.51	0.04	0.37	0.08	5.39	0.381
FACTOR_FHLR	-0.04	-5.44	-0.04	-6.29	-0.04	-5.49	-0.05	-2.26	-0.04	-3.85	-0.03	-4.31	-0.04	-2.39	-0.04	-2.68	-0.04	-4.58	0.059
FACTOR_KM	-0.18	-10.89	-0.18	-11.55	-0.18	-11.20	-0.22	-2.89	-0.18	-12.09	-0.17	-8.62	-0.18	-4.23	-0.18	-3.70	-0.17	-7.89	0.002
RESTR_YES	0.04	4.88	0.04	4.88	0.07	6.78	0.09	3.66	0.04	2.95	0.04	4.27	0.04	1.71	0.04	2.06	0.05	4.45	0.034
ITDIR_IT	0.03	6.29	0.02	5.19	0.02	4.38	0.02	1.12	0.03	4.17	0.03	5.88	0.03	1.81	0.03	1.37	0.03	4.48	0.521
ITDIR_1 STEP	0.02	4.72	0.02	3.66	0.02	4.53	0.03	2.01	0.02	2.57	0.02	3.72	0.02	2.83	0.02	2.54	0.03	4.54	0.163
PUBL_YES	0.04	7.06	0.04	6.99	0.05	7.60	0.09	6.35	0.04	5.38	0.05	8.12	0.04	2.47	0.04	2.16	0.04	5.84	0.237
AUTHOR_YES	-0.03	-6.03	-0.04	-9.98	-0.03	-5.45	-0.05	-3.34	-0.03	-4.08	-0.04	-6.73	-0.03	-2.76	-0.03	-1.73	-0.04	-5.64	0.222
FOCUS_DFM YES	-0.04	-5.91	-0.04	-6.36	-0.05	-5.45	-0.07	-3.18	-0.06	-6.07	-0.08	-10.89	-0.04	-1.51	-0.04	-1.46	-0.04	-3.88	0.413
# obs	18,782		18,782		19,333		20,935		18,782		18,782		18,782		18,782		18,782		
R^2adj	0.114		0.114		0.107		0.110		0.935		0.935		0.114		0.114		0.882		
R^2adj*	-				-		-		0.087		0.094		-		-		0.101		

Notes: (1) *Ad hoc* outlier adjustment (OLS with White corrected stderr.), unweighted, (2) as (1), but after successive removal of insignificant variables, (3) *ad hoc* outlier adjustment where the upper and lower percentiles of the data where removed, (4) robust estimation, Tukey or biweight function, (5) *ad hoc* outlier adjustment, WLS (equal weights for each obs. from each study and for each study), (6) *ad hoc* outlier adjustment, WLS (equal weights for each obs. from each dataset and for each dataset), (7) *ad hoc* outlier adjustment, robust clustering where each study represents a cluster, (8) *ad hoc* outlier adjustment, robust clustering where each dataset represents a cluster, (9) *ad hoc* outlier adjustment, WLS (weights according to standard deviation of residuals within each study). mean refers to means of the dummy variables, associated with observations included in specification (1). R^2adj* refers to the adjusted R^2 where the WLS coefficient estimates are applied to the unweighted data.

Table 4: Meta-regression results for inflation

	Baseline				Outliers				Sampling bias				Dependency				Quality		
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		
	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	mean
CONSTANT	0.90	99.62	0.89	126.10	0.92	87.57	0.94	48.83	0.90	81.44	0.90	95.51	0.90	22.97	0.90	16.06	0.91	79.33	-
COUNTRY_UK	0.08	10.42	0.08	12.13	0.08	9.64	0.10	5.19	0.08	8.10	0.07	9.18	0.08	3.76	0.08	3.16	0.08	9.36	0.062
COUNTRY_EA	-0.00	-0.34	-	-	0.00	0.22	-0.03	-1.63	-0.00	-0.13	-0.00	-0.43	-0.00	-0.06	-0.00	-0.05	-0.02	-2.15	0.359
COUNTRY_OTHER	0.00	0.19	-	-	0.00	0.08	-0.00	-0.02	0.01	0.80	0.01	0.64	0.00	0.08	0.00	0.11	-0.00	-0.28	0.201
BENCH_RW	-0.03	-2.29	-0.03	-2.47	-0.04	-3.21	-0.09	-3.72	-0.05	-3.16	-0.03	-1.88	-0.03	-0.97	-0.03	-0.82	-0.02	-1.59	0.019
BENCH_VAR	-0.02	-3.37	-0.03	-4.09	-0.04	-5.08	-0.09	-6.85	-0.02	-2.62	-0.02	-2.71	-0.02	-1.87	-0.02	-1.69	-0.03	-2.98	0.061
BENCH_INDIC	-0.07	-12.75	-0.07	-13.57	-0.07	-12.33	-0.09	-8.23	-0.06	-9.98	-0.05	-9.97	-0.07	-2.18	-0.07	-2.18	-0.08	-12.21	0.610
BENCH_FOREC POOL	0.06	10.59	0.06	10.63	0.05	7.89	0.06	4.29	0.06	7.38	0.05	8.18	0.06	2.94	0.06	2.06	0.07	9.67	0.095
BENCH_OTH LARGE	0.16	9.34	0.17	11.51	0.24	11.81	0.30	8.35	0.14	6.85	0.12	6.75	0.16	1.91	0.16	1.38	0.20	9.35	0.014
HORIZON (demeaned)	-0.00	-3.70	-0.00	-3.93	0.00	2.65	0.00	2.65	-0.00	-3.56	-0.00	-3.13	-0.00	-1.04	-0.00	-0.80	-0.00	-2.78	-
log(N) (demeaned)	-0.02	-7.94	-0.02	-8.16	-0.02	-7.86	-0.02	-4.42	-0.02	-6.40	-0.02	-7.53	-0.02	-2.64	-0.02	-2.57	-0.02	-6.83	-
log(T) (demeaned)	-0.01	-1.43	-	-	-0.04	-5.95	-0.09	-7.98	-0.01	-1.98	-0.00	-0.41	-0.01	-0.40	-0.01	-0.31	-0.01	-0.68	-
BAL_NO	-0.02	-2.98	-0.02	-3.24	-0.02	-2.93	-0.03	-2.86	-0.01	-1.85	-0.02	-3.12	-0.02	-1.10	-0.02	-1.04	-0.03	-3.45	0.173
PRESEL 1	0.02	2.37	0.02	2.68	0.02	2.20	0.01	0.53	0.02	1.89	0.00	0.22	0.02	1.05	0.02	1.01	0.02	1.70	0.089
PRESEL 2	0.04	11.89	0.04	11.72	0.04	10.60	0.06	7.85	0.03	8.20	0.05	11.84	0.04	4.75	0.04	4.82	0.04	8.98	0.324
ROLREC_ROL	0.01	0.87	-	-	-0.02	-1.37	-0.08	-3.79	0.01	0.84	0.01	1.18	0.01	0.37	0.01	0.29	0.01	0.69	0.048
FREQ_Q	0.02	1.62	0.03	5.73	-0.04	-3.22	-0.14	-7.80	0.01	0.61	0.02	2.01	0.02	0.75	0.02	0.55	0.01	1.22	0.444
FACTOR_FHLR	0.03	3.78	0.04	5.51	0.04	3.48	0.01	0.40	0.03	2.37	0.03	3.24	0.03	2.51	0.03	1.12	0.03	2.80	0.136
FACTOR_KM	-0.02	-1.71	-	-	-0.07	-4.81	-0.17	-4.69	-0.02	-1.32	-0.01	-0.39	-0.02	-0.41	-0.02	-0.31	-0.03	-1.66	0.010
RESTR_YES	-0.13	-11.40	-0.14	-13.51	-0.16	-12.08	-0.19	-6.92	-0.10	-7.30	-0.12	-9.88	-0.13	-2.23	-0.13	-2.13	-0.18	-11.97	0.034
ITDIR_IT	0.03	3.38	0.03	3.71	-0.01	-1.57	-0.07	-5.44	0.02	2.14	0.03	3.24	0.03	0.83	0.03	0.87	0.04	4.00	0.077
ITDIR_1 STEP	-0.04	-6.73	-0.04	-6.95	-0.04	-5.94	-0.04	-2.93	-0.05	-7.53	-0.03	-5.48	-0.04	-3.53	-0.04	-2.27	-0.03	-4.27	0.179
PUBL_YES	0.08	10.66	0.08	17.20	0.09	11.02	0.12	8.88	0.08	8.52	0.07	8.68	0.08	2.92	0.08	2.00	0.07	7.99	0.571
AUTHOR_YES	0.03	5.32	0.03	5.48	0.03	5.24	0.05	3.87	0.03	4.84	0.03	5.17	0.03	0.90	0.03	0.88	0.04	5.71	0.362
FOCUS_DFM YES	0.07	10.88	0.07	12.40	0.09	13.42	0.18	15.81	0.06	7.26	0.06	9.30	0.07	2.68	0.07	2.90	0.08	10.28	0.481
# obs	20,263		20,263		21,369		23,547		20,263		20,263		20,263		20,263		20,263		
R^2adj	0.172		0.173		0.168		0.105		0.851		0.844		0.172		0.172		0.757		
R^2adj*	-				-		-		0.168		0.168		-		-		0.167		

Notes: (1) *Ad hoc* outlier adjustment (OLS with White corrected stderr.), unweighted, (2) as (1), but after successive removal of insignificant variables, (3) *ad hoc* outlier adjustment where the upper and lower percentiles of the data where removed, (4) robust estimation, Tukey or biweight function, (5) *ad hoc* outlier adjustment, WLS (equal weights for each obs. from each study and for each study), (6) *ad hoc* outlier adjustment, WLS (equal weights for each obs. from each dataset and for each dataset), (7) *ad hoc* outlier adjustment, robust clustering where each study represents a cluster, (8) *ad hoc* outlier adjustment, robust clustering where each dataset represents a cluster, (9) *ad hoc* outlier adjustment, WLS (weights according to standard deviation of residuals within each study). mean refers to means of the dummy variables, associated with observations included in specification (1). R^2adj* refers to the adjusted R^2 where the WLS coefficient estimates are applied to the unweighted data.

estimation techniques over the SW technique for output, on the other hand, the differences between FACTOR_SW and FACTOR_FHLR are found to be only minor and unsystematic by Boivin and Ng (2005), D'Agostino and Giannone (2006), Stock and Watson (2005a) and Cheung and Demers (2007).[28] Our results (at least with respect to output) are remarkable given that, in practice, most forecasts (i.e. roughly 94% of output and 85% of inflation, as shown in the last columns of Tables 3 and 4) are based on the SW factor estimation method, which is much easier to implement.

Our findings also suggest that imposing the restrictions implied by the factor structure in the forecasting equation may improve the relative factor forecast performance for inflation, but tends to worsen the performance for output. This partly supports the results obtained by D'Agostino and Giannone (2006) who find that restrictions are not harmful to factor forecasts and partly those of Boivin and Ng (2005) who opt for the more robust unrestricted estimation. We also find gains from making direct rather than indirect multi-step forecasts. This differs from Boivin and Ng (2005). According to the authors, there are also only negligible differences between the iterated and direct forecasts for inflation variables. By contrast, they find some gains of applying an iterated rather than a direct multi-step forecast scheme to output variables.

Regarding our variables capturing the publication strategy, we find a publication bias: factor models appear to perform relatively better in unpublished papers than in published papers. Results with respect to the other two variables (AUTHOR and FOCUS_DFM) are unclear. While we find an author bias and a negative significant coefficient of FOCUS_DFM_YES for output, the corresponding signs are positive for inflation.

4.2.2. Robustness checks: outliers, sampling bias, dependency and qualitative differences

The appropriate weighting of observations is a key issue. We begin this subsection by presenting alternative ways of dealing with outliers. We then address problems which may arise due to sampling bias and dependency of observations in a second step and quality differences in a third step.

[28] All these studies allow estimation techniques to be disentangled from whether the forecasting equation is estimated restrictedly or unrestrictedly. Schumacher (2006) compares the three factor estimation techniques while estimating forecasting equations unrestrictedly. Forni et al. (2003) and Stock and Watson (2005a) compare the FHLR and SW factor estimation techniques while performing only a restricted and only an unrestricted estimation of the forecasting equation, respectively. Finally, Boivin and Ng (2005) and D'Agostino and Giannone (2006) consider both estimation techniques and perform restricted and unrestricted estimations for both.

Outliers

Our baseline meta-regression (as well as the descriptive analysis) relies on a dataset from which outliers were removed in a particular, rather *ad hoc* manner, as explained in section 3. As a robustness check, we adopt another *ad hoc* outlier removal technique which involves dropping the upper and lower 5 percentiles of the data, following Knell and Stix (2005b). An alternative way to deal with outliers is to apply robust regression methods to the entire dataset (including those observations which were detected as outliers by our *ad hoc* methods and removed from the sample) (cf. Rousseeuw and Leroy, 1987). We adopt a so-called M-estimator which falls into the class of robust regression methods. Instead of minimizing the sum of squared residuals, M-estimators minimize the sum of a function of the residuals. A number of functions which downweight observations with residuals being distant from zero have been proposed in the literature. We use the commonly employed Tukey or biweight weighting function proposed by Beaton and Tukey (1974) where weights gradually decrease with the distance of the residuals from zero and are set to zero from some point on.[29] The M-estimation was implemented with iterated reweighted least squares.

Sampling bias and dependency

There are two problems with using many observations from individual studies. First, studies which report a large number of results would be given a large weight relative to studies which report only a few results, also known as sampling bias (cf. Stanley, 2001; Weichselbaumer and Winter-Ebmer, 2005). Second, observations from the same study (conducted by the same researcher(s)) may not be independent, but may cluster within individual studies. These problems could be present in our sample where the number of results taken for our analysis ranges from 2 (Kabundi, 2004) to more than 9,000 (Lin and Tsay, 2005) (before removal of outliers). Dependency and sampling bias may also exist because researchers rely on similar or identical datasets from which they estimate the factors. For example, the quarterly dataset constructed by SW for the US (or a slightly modified version) is used in a total of 11 studies.

We address these problems as follows. First, to eliminate potential sampling bias and enhance efficiency, we construct a weighting matrix which attributes equal weight to each observation from a particular study $1/n_{g^s}$ where n_{g^s} is the number of observations taken from study $g^s = 1,...,52$. Accordingly, we attribute equal weights to results produced based on equal or similar datasets, namely $1/n_{g^d}$ where n_{g^d} is the number of observations produced with the dataset $g^d = 1,...,37$. Equation (5) is then estimated with weighted least squares (WLS) (see also Weichselbaumer and Winter-Ebmer, 2005; Knell and Stix, 2005). Second, we apply

[29] The weights equal $\left[1-(\eta_j / \kappa_T)^2\right]^2$ for $|\eta_j| \le \kappa_T$ and 0 for $|\eta_j| > \kappa_T$, where κ_T, the tuning constant associated with the Tukey or biweight function, is set at 4.685.

robust clustering to equation (5), which provides us with robust covariance estimates and thereby adjusts for within-cluster correlation, where each study and each dataset, respectively, represents a cluster.[30]

Quality

Many meta-analyses downweight low-quality results and give high-quality results larger weights (cf. Knell and Stix, 2005; Weichselbaumer and Winter-Ebmer, 2005). Most existing meta-analyses focus on regression coefficient estimates. The underlying studies generally also report some measure of estimation uncertainty such as standard errors, and attribute weights to the observations which are inversely related to this uncertainty. Measuring the quality of observations is, however, difficult in our context (as well as in other contexts: see, for example, Weichselbaumer and Winter-Ebmer, 2005). Some of the papers on which our analysis is based do not provide standard errors at all. Others provide standard errors or use diverse formal tests to find out whether some models significantly outperform others.[31] The different statistics are, however, barely comparable. Finding a common metric of estimation precision is further complicated by the fact that we often need to convert results reported in the papers to relative RMSEs, and it is unclear how such a conversion of the statistics translates to the uncertainty surrounding them.

Nevertheless, we try to address the issue of qualitative differences between observations. We follow Weichselbaumer and Winter-Ebmer (2005) and Longhi et al. (2005) and weight the observations with the inverse of the standard deviation of the residuals obtained from an OLS regression of equation (5) associated with individual studies and, thus, fit WLS to equation (5). This presumes that a relationship between large heterogeneity among relative RMSEs obtained from an individual study and less precise forecasts. Obviously, the drawback of such a weighting scheme is that all observations taken from one study are weighted equally, although some results of that particular study are probably more accurate than others. Also, large heterogeneity across results from an individual study might simply reflect the broad scope of that study and careful consideration of lots of models. Results should therefore be interpreted with caution.

Overall, results are relatively robust across different outlier adjustment methods and weighting schemes; see specifications (3) to (9) in Tables 3 and 4.[32] Two remarks, however,

[30] Under robust clustering, the estimated variance of the parameters is given by $\left(\sum_g G_g' G_g\right)^{-1}\left(\sum_g G_g' \eta_g \eta_g' G_g\right)\left(\sum_g G_g' G_g\right)^{-1}$, where $G = [1 \quad M']$ and subscript $g \in \{g^s, g^d\}$ refers to the group (the study or the dataset).

[31] They use, for instance, the Diebold and Mariano (1995) tests for non-nested models and the Clark and McCracken (2001) or Giacomini and White (2006) tests for nested models.

[32] Interestingly, WLS permits us to noticeably reduce the dispersion in the dependent variable, which is reflected in a large adjusted R^2 of the WLS regressions; see Tables 3 and 4.

are in order. First, a few coefficients switch signs with different outlier adjustment methods (specifications (3) and (4)) (although variables generally become also insignificant). The coefficient associated with output T turns significantly negative, which is the expected sign; the coefficients associated with inflation of FREQ_Q, ITDIR_IT, ROLREC_ROL turn negative and of HORIZON turns positive with both specifications (3) and (4). Second, dependency seems to be an important issue (cf. specifications (7) and (8) in Tables 3 and 4). This is reflected in generally much lower *t*-statistics for the coefficients. Some variables turn from significant to insignificant, such as BENCH_RW and ITDIR_IT for both output and inflation. Also, N, FREQ_Q and FOCUS_DFM YES no longer enter the equation for output significantly, and BENCH_VAR, BENCH_OTH LARGE, HORIZON, BAL_NO, PRESEL_1 and AUTHOR_YES become irrelevant in the equation for inflation. Dependency seems to be present between observations taken from individual studies as well as datasets.

5. Conclusion

In this study, we have taken a meta-analytic approach to empirically assess the relative forecast performance of large dynamic factor models for real economic activity and inflation. This approach has allowed us to systematically summarize findings from existing factor forecast applications. At the same time, we have been able to assess the relevance of a large number of determinants of the factor forecast performance and therefore hope that our analysis will help practitioners to improve their forecasts.

Our results broadly suggest that, on average, the forecast performance of large-scale dynamic factor models is slightly superior to that of other models. Among the variables determining the forecast environment (which cannot be influenced by the forecasters), the variable itself (whether output or inflation is predicted and from which country/country group the variable is taken) seems to influence the relative forecast performance of factor models: Factor models deliver better predictions of US variables than UK variables, of US output than euro-area output and of euro-area inflation than US inflation. According to our analysis, factor model forecasts are worse than pooled forecasts, but generally outperform small-scale models, and we conclude that it is worthwhile exploiting information from data-rich environments.

Among the variables governing the forecast design (which can be influenced by the forecaster), the size of the dataset positively affects the relative factor forecast performance: the cross-section and time dimensions tend to improve factor forecasts, and it seems worthwhile using monthly (and not only quarterly) information and doing recursive rather than rolling forecasts. Surprisingly, there is no evidence that pre-selecting the variables has positively influenced the factor forecasts in the past, which may, however, by due to rather *ad hoc* procedures adopted by most studies. In addition, it seems important to carefully specify the forecasting model. The dynamic, more complex, factor estimation approaches of FHLR and KM outperform the static and simpler SW approach for output (but differences between the factor estimation techniques are less important for inflation). Moreover, imposing restrictions implied by the dynamic factor structure in the forecasting equation has helped to improve inflation forecasts in the past.

III. Business Cycle Transmission from the US to Germany – a Structural Factor Approach[*]

1. Introduction

International business cycle linkages have recently returned to the focus of public interest. This renewed interest has its roots in the worldwide economic downturn in 2001. It has often been claimed that the downturn was caused (at least partly) by economic disturbances in the US (cf. Artis et al., 2007; Monfort et al., 2004). In the second half of the 1990s, the US economy experienced a prolonged phase of extraordinarily large productivity gains due to technological advances. Those productivity gains were expected to be long-lasting. Consequently, global demand was boosted, triggering a rapid and exaggerated boom in international stock markets. The subsequent bursting of the stock market bubble contributed notably to the global economic downturn. Its remarkable strength, speed and synchronicity across industrial countries were apparently unexplainable in terms of trade linkages alone. At the same time, financial markets and confidence measures around the globe were particularly affected during the downturn. This led researchers to examine international business cycle linkages more closely, with a particular focus on the international propagation of US shocks and the relevance of the various international transmission channels.

The present study investigates the transmission of macroeconomic shocks from the US to Germany using the large-scale structural dynamic factor model developed by Forni and Reichlin (1998). It addresses three main questions which are relevant for forecasters and policymakers.

First, to what extent do US shocks affect the German economy? Our modeling framework enables us to assess the impact of US shocks not only on German economic activity but also on many other variables such as prices and employment.

Second, what are the role and the relevance of different international transmission channels? Besides the more traditional trade channel, "new" channels, i.e. financial markets and the

[*] This chapter is based on Eickmeier (2007), „Business cycle transmission from the US to Germany – a structural factor approach“, *European Economic Review*, 51(3), 521-551. This paper was written while I was visiting the Université Libre de Bruxelles and the Deutsche Bundesbank. It was presented at the ECB workshop “The external dimension of the euro area: trade, capital flows and international macroeconomic linkages”, the APF conference in Frankfurt, the CFS summer school in Eltville and at seminars at the Deutsche Bundesbank, the German Council of Economic Experts and the University of Cologne. I am very grateful to Jörg Breitung, Jacopo Cimadomo, Olivier De Bandt, Juergen B. Donges, Jörg Döpke, Ralf Fendel, Heinz Herrmann, Michele Lenza, Giovanni Lombardo, Christian Schumacher, Harald Uhlig, Jens Ulbrich, Jens Weidmann, and two anonymous referees for very helpful comments and suggestions.

confidence channel, are considered. From a theoretical point of view, it is unclear whether shocks are spread positively or negatively through all these channels (see the detailed discussion in Kose et al., 2003d). There is some theoretical support for a positive international transmission. In terms of trade, for example, higher import demand in one country, for example, will boost exports in other countries (cf. Canova and Dellas, 1993). Also, productivity advances may spread internationally through vertical integration (cf. Kose and Yi, 2001; Elliott and Fatás, 1996). In terms of financial integration, it enables agents to diversify their risk by investing in different markets, and financial prices become more highly synchronized through arbitrage (Kose et al., 2003c, 2003d; Doyle and Faust, 2002). Other theoretical arguments, however, point in favor of a negative international propagation. If trade is accompanied by a larger degree of inter-industrial specialization, linkages should be loosened in the presence of industry-specific shocks (cf. Frankel and Rose, 1998). Greater inter-industrial specialization and negative spillovers can also result from mobile capital being reallocated to economies where it is used most productively (Canova and Marrinan, 1998; Kalemli-Ozcan et al., 2003; Imbs, 2004; Heathcote and Perri, 2004). Given the conflicting theories, the question of how globalization, i.e. real and financial integration as well as confidence linkages, affects the international propagation of macroeconomic shocks has to be answered empirically.

Whereas the shock propagation through trade and financial markets has been investigated extensively in the literature, not much attention has yet been directed to the confidence channel. The confidence channel denotes the international transmission of economic and financial shocks through sentiment and/or expectations that might, for example, be caused by "informational cascades" or "fads" (cf. IMF, 2001a; Anderton et al., 2004). It is reflected in a shock transmission which goes beyond or is smaller than the propagation through fundamental channels such as trade and financial markets and has been examined in some detail for emerging markets and for periods of financial stress (cf. Pericoli and Sbracia, 2003; Dornbusch et al., 2000). IMF (2001a) and Anderton et al. (2004) have recently emphasized the confidence channel for industrial countries and "normal" cyclical periods. Our analysis seeks to shed light on the role of the different transmission channels.

Third, we consider business cycle linkages between the US and Germany during two specific periods and ask how the boom experienced by the US between 1995 and 2000 and the subsequent US recession in 2001 spread to Germany.

This study is related to the empirical literature on cyclical interdependencies between countries. Business cycle linkages between the US and Germany have been examined previously by Canova and Marrinan (1998), Artis et al. (2006), Artis et al. (2007) and the German Council of Economic Experts (GCEE, 2001). Those analyses find that shocks to the US business cycle are transmitted significantly and positively to Germany. All these studies employ small-scale VAR models (with constant or time-varying parameters). In addition, the

GCEE (2001) uses the Deutsche Bundesbank's large-dimensional macroeconometric multi-country model. Numerous studies also investigate business cycle linkages between the US and Europe, cf. IMF (2001b), GCEE (2001), Artis et al. (2006), Dalsgaard et al. (2001), Artis et al. (2007), Adjemian et al. (2004), De Walque and Wouters (2004), Monfort et al. (2004), Dees et al. (2007). It goes without saying that it is of particular interest to investigate the transmission to Germany, the largest European economy in terms of GDP and population.

Empirical studies focusing on international transmission channels include Otto et al. (2001), Kose et al. (2003d), Imbs (2004) and Baxter and Kouparitsas (2005). Their work is based on cross-country regression models where some measure of co-movement is regressed on variables covering international transmission channels like trade and financial integration. There is some consensus that a positive relationship exists between two countries' business cycle correlations and trade intensity/financial linkages. IMF (2001a) and Anderton et al. (2004) find empirical evidence suggesting that the confidence channel may play a role for the shock transmission between the US and Europe.[33] We will discuss the relevance of the confidence channel by investigating how US shocks affect German confidence measures. In the present framework, it will, however, be difficult to disentangle fluctuations of German confidence which only reflect the international spreading of US shocks through fundamental channels from those which are not justified by movements in trade and financial markets.

As already outlined above, the models often employed in analyses focusing on aggregate business cycle linkages range from vector autoregressive (VAR) models[34], fully structural macroeconometric multi-country models[35] or, more recently, fully- fledged international DSGE models.[36] Among the studies which are mentioned above and which examine business cycle linkages between the US and European countries, only Monfort et al. (2004) use a factor model. Other factor-based analyses of international business cycle linkages with a somewhat different country-coverage include Stock and Watson (2005b), Kose et al. (2003a, 2003b, 2003c), Lumsdaine and Prasad (2003), Bayoumi and Helbling (2003), Norrbin and Schlagenhauf (1996), Eickmeier and Breitung (2006) and Eickmeier (2005). However, most of these studies use exact factor models and apply them to a small dataset containing only international output measures.[37] Most of the above-mentioned factor studies, in addition, do

[33] According to the IMF (2001a), international confidence measures are more strongly correlated than international real economic activity measures. In addition, residuals from a regression of business confidence on economic activity measures or leading indicators are found to be significantly correlated across the US and the euro area. Anderton et al. (2004) show that US confidence Granger causes euro-area confidence.
[34] Cf. Canova and Marrinan (1998), GCEE (2001), Artis et al. (2006), Dees et al. (2007).
[35] Cf. Dalsgaard et al. (2001), IMF (2001b), GCEE (2001).
[36] Cf. Adjemian et al. (2004), De Walque and Wouters (2004).
[37] There are three exceptions. Bayoumi and Helbling (2003) include 35 series of output, investment, consumption and exports of G7 countries and estimate a large-scale dynamic factor model. Eickmeier (2005) investigates economic comovements in the euro area based on a panel of more than 150 variables from key euro-area economies by means of a dynamic factor model. And Eickmeier and Breitung (2005) employ a factor model to study business cycle synchronization between the euro area and central and east European countries. They use more than 200 variables.

not perform structural analysis, i.e. they do not interpret economic co-movements, but confine themselves to estimating a common cycle.[38] Finally, note that our study is also related to other applications of large-scale structural dynamic factor models which have become popular in recent years (see, for example, Giannone et al., 2002, 2005; Sala, 2003 and Cimadomo, 2004 for monetary policy applications).

The main contribution of this chapter is that it applies a large-dimensional structural dynamic factor model to the topic of international business cycle transmission. To our knowledge, it is the first to do so. Studying international business cycles in such a framework has various advantages over VAR models[39] or structural models. Much information can be exploited in dynamic factor models which should allow us to estimate the common driving forces and their propagation more precisely. Moreover, and particularly favorable in our context, it allows us to assess the responses of many variables to macroeconomic shocks. It can therefore be employed to simultaneously assess the relevance of a large number of transmission channels, including the "new" channels, without needing to know the exact workings of these channels. VAR models and cross-country regression models have to cope with collinearity and scarce-degrees-of-freedom problems. In addition, cross-country regression models need to deal with endogeneity problems. Fully structural models can contain a large number of variables. These models at present, however, do not include all and detailed channels. For example, stock markets, confidence, foreign direct investment (FDI) and bank lending are generally missing. The reason may be a lack of consensus on how to model these "new" channels in a wholly structural framework. These models are found to be not able to fully account for the international output correlations, suggesting that the missing channels are of some importance (IMF, 2001a; GCEE, 2001). These advantages are accompanied by two drawbacks. The interpretation of the outcome is difficult for such a largely reduced form model. A further disadvantage of our model is that we can assess how trade variables, financial market variables etc. react to the shocks, but we do not know how movements in each of these variables *ceteris paribus* affect economic activity in Germany.

This chapter is organized as follows. Section 2 describes the data set. Section 3 presents the model and briefly describes estimation techniques and the identification of shocks. Section 4 characterizes US shocks. Section 5 illustrates the impact of US shocks on German key variables, including variables covering the transmission channels. Section 6 analyzes the contribution of US shocks to economic activity in Germany since the mid-1990s. Section 7 concludes.

[38] An exception is, again, Eickmeier and Breitung (2006) who identify structural euro-area shocks and investigate their propagation to new EU member states. Also, some of the studies correlate common international factors with potentially explanatory series, such as US shocks, oil prices, trade variables (Stock and Watson, 2005a; Kose et al., 2003a, 2003b; Monfort et al., 2003; Eickmeier, 2005).

[39] An exception is the global VAR model recently developed by Pesaran et al. (2004) and extended in Dees et al. (2007). Its estimation involves the estimation of small-scale VAR models only, which are then stacked to yield a global large-scale VAR model and, therefore, does not face the same drawbacks as the usual VAR models.

2. Data

We rely on a large data set with $N = 296$ variables observed over $T = 112$ quarters from 1975 to 2002. The data set contains domestic real and nominal variables for the US and German economies. In addition, it includes measures of the international integration of both countries, namely trade and financial variables and confidence indicators. Notice that we consider trade in goods and services separately. In addition, we are able to disentangle long-term interest rates and stock prices and different categories of capital flows, namely FDI, securities as well as credits. It would certainly be interesting (and possible in the present framework) to consider trade and financial market measures at further disaggregated levels. For instance, the relevance of direct versus third-market effects, or the role of trade of intermediate goods versus final goods, could be investigated. This is, however, beyond the scope of this chapter, and we leave it for future research. Global influences are finally captured by world commodity prices. Outliers are removed. Integrated series are made stationary through differencing and/or deterministic de-trending. All series are standardized to have means of 0 and unit variances. For details on the data we refer to Appendix A and Table A1.

3. The model, estimation and identification of US shocks

The series are collected in the $N \times 1$ vector $\mathrm{y}_t = [y_{1t} \quad ... \quad y_{Nt}]'$. It is assumed that y_t follows an approximate dynamic factor model (cf. Stock and Watson, 1998, 2002a; Bai and Ng, 2002) and can be represented as:

$$\mathrm{y}_t = \mathrm{x}_t + \xi_t = \Lambda' \mathrm{f}_t + \xi_t, \tag{1}$$

where $\mathrm{f}_t = [f_{1t} \quad \cdots \quad f_{rt}]'$ is a $r \times 1$ vector of common (static) factors and $\Lambda = [\Lambda_1 \quad \cdots \quad \Lambda_N]$ is an $r \times N$ matrix of factor loadings. x_t and ξ_t are $N \times 1$ vectors of common and idiosyncratic components. Typically, $r \ll N$.

The common factors are driven by common shocks which are global shocks and/or idiosyncratic shocks which are propagated to other variables, including international variables. Those factors and shocks are the same for all variables, but reactions to changes in the common factors may differ, the latter being reflected in possibly heterogeneous loadings. The vector ξ_t, by contrast, contains influences which are specific to individual variables or small groups of variables. The idiosyncratic component ξ_t thus is driven by idiosyncratic shocks which barely, or do not, spread to other variables. ξ_t may also reflect measurement error. The idiosyncratic components are allowed to be weakly cross-correlated and weakly serially correlated in the sense of Bai and Ng (2002).

We suppose that f_t has a VAR representation:

$$A(L)f_t = Qv_t, \qquad (2)$$

where $A(L) = I - A_1 L - ... - A_p L^p$ and v_t is a $r \times 1$ vector of orthogonal innovations. It is further assumed that structural common shocks $\tilde{w}_t$ are linearly related to the orthogonal innovations v_t:

$$\tilde{w}_t = Hv_t. \qquad (3)$$

Provided that there are enough identifying restrictions on H, the structural shocks $\tilde{w}_t$ can be recovered from the factor innovations. It is important to notice that we are interested only in a subgroup of structural shocks, namely in US shocks that are spread internationally, which we denote by w_t. The ultimate goal is to assess the transmission of these shocks to the German economy.

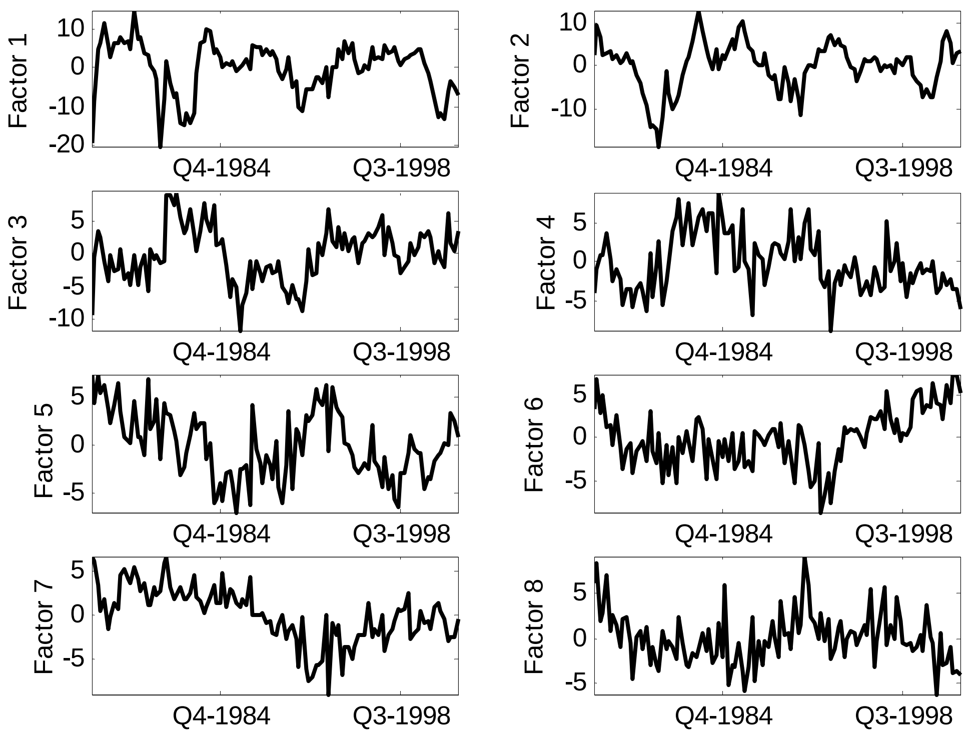
Table 5: Largest ten factor loadings and variance explained by each factor

Factor 1		Factor 2		Factor 3		Factor 4		Factor 5		Factor 6		Factor 7		Factor 8	
#	Loading	#	Loading	#	Loading	#	Loading	#	Loading	#	Loading	#	Loading	#	Loading
35	0.142	89	0.141	201	0.147	127	0.154	213	0.149	209	0.172	85	0.173	195	0.174
36	0.140	133	-0.136	219	0.145	125	0.152	241	-0.146	113	0.157	47	0.162	179	0.171
41	0.135	227	0.130	200	0.145	128	0.150	162	0.134	249	-0.155	62	0.150	177	0.169
39	0.135	96	0.130	202	0.138	126	0.146	186	-0.134	243	0.136	79	0.144	70	0.152
42	0.133	132	-0.129	240	0.137	161	0.135	176	0.133	237	0.135	84	0.140	175	0.152
118	0.132	104	0.128	292	0.129	246	0.127	175	0.132	108	0.131	219	-0.137	165	-0.147
40	0.131	99	0.126	159	-0.127	105	-0.125	166	0.130	244	0.130	162	0.137	178	0.142
55	0.131	225	0.125	206	0.127	245	0.123	179	0.119	254	-0.129	187	-0.133	134	-0.140
56	0.129	97	0.125	290	0.125	100	-0.121	177	0.116	255	-0.127	70	0.131	291	0.136
57	0.128	131	-0.124	205	0.123	247	0.117	78	0.116	205	0.122	169	0.131	176	0.127
Variance of the total panel explained by each factor/variance explained by the first eight factors															
0.29		0.22		0.12		0.09		0.08		0.07		0.07		0.05	

Notes: The variable description corresponding to # can be found in Table A1. The signs of the factors are normalized such that largest absolute loading is positive.

For this purpose, we first estimate f_t by applying static principal component analysis to y_t. The dimension of f_t, r, was estimated to be 8 on the basis of the Bai and Ng (2002) criteria. Notice that eight factors explain 50% of the total variance. Figure 2 shows the time series of the estimated factors. A glimpse of the loadings may help to interpret the factors in economic terms. Table 5 reports the largest ten loadings associated with each of the eight factors. The first factor which explains 29% of the factor space is a real US factor, being most closely related to US employment and industrial production. The second factor accounts for 22% of the factor space, and factor loadings are largest for US CPI inflation and interest rates. Factors three to eight contribute to a much lesser extent to economic comovements. The third factor is most closely related to capacity utilization in Germany, the fourth to interest rates in the US, the fifth to wages,

Figure 2: Times series of the first eight common factors



confidence, consumption and employment in Germany, the sixth to unit labor costs in Germany, total reserves in the US and interest rates in Germany, the seventh to the US labor market, and the eighth to the construction industry in Germany. Such an exercise should, however, not be conducted without caution. The factors are only identified up to a rotation and cannot be interpreted as such.[40] In addition, interpretation is made more difficult by the fact that f_t can be a linear combination between not just the dynamic factors, but also their lags. To address these concerns, in the following we identify the structural shocks behind common factors and try to disentangle US shocks that have an impact on the German economy from other common shocks.[41]

We estimate the innovations v_t by following Giannone et al. (2002) (and as described in Appendix B) and fit a VAR(1) model to the estimated vector of factors $\hat{f}_t$. The lag order of the VAR model was estimated with the Schwarz information criterion. The structural US shocks w_t cannot be recovered without imposing identification restrictions. Most studies disentangle spillovers from other common shocks by imposing that the former only have a delayed impact on other countries while the latter affect all countries at the same time (cf.

[40] With an exception: in the international business cycle literature, the first factor is often interpreted as a common cycle.
[41] An alternative is to rotate the factors themselves. See, for example, Eickmeier (2005).

Monfort et al., 2004). The problem with this approach is that spillovers within a period are attributed to common shocks. Since one of the goals of our study is to assess the propagation of US shocks to Germany, we refrain from restricting the impact *a priori*. Instead, we have chosen to adopt another approach. We focus on the main driving forces of the US economy. We then give them an economic interpretation using identification restrictions with respect to US variables and name them US shocks. The disadvantage of this approach is, of course, that it cannot be assured that those US shocks are shocks that have their origin in the US. In the next section, we will, however, provide some evidence that most of them should. But let us first explain in more detail how we recover these shocks.

Following Uhlig (2003), we extract those shocks that explain as much as possible of the forecast error variance of (the common component of) US GDP, where, as in Uhlig (2003), US GDP is taken as a proxy for US economic activity and the forecast horizon is zero to five years. Principal component analysis is applied to (some function of) the impulse responses of US GDP to v_t. It turns out that two shocks are sufficient to explain the overwhelming forecast error variance share of the common component of US GDP, 98%.[42] We therefore only focus on two shocks. The next step involves giving the two shocks a structural interpretation as is done in the SVAR literature. Following Peersman (2005), we use a relatively agnostic identification approach and impose short-run sign restrictions on impulse responses of key US variables to identify a US supply shock and a US demand shock. This prevents us from using zero restrictions which are at odds with some theoretical models (see the discussions in Peersman, 2005, and in Canova and de Nicólo, 2003). As in Peersman (2005), we impose the following restrictions. A positive supply shock has non-negative effects on output and non-positive effects on prices contemporaneously and during the first four quarters after the shock; the short-term interest rate does not increase on impact. A positive demand shock affects output and prices non-negatively instantaneously and during the first four quarters after the shock; the immediate effect on the short-term interest rate is non-negative. These conditions are consistent with the standard aggregate supply-aggregate demand framework and with more complex structural models such as the DSGE model outlined in Smets and Wouters (2003).[43] The vector of impulse-response functions of variable i to the shocks $\mathrm{w}_t = \begin{pmatrix} w_{1t} & w_{2t} \end{pmatrix}'$ at horizon h is given by $\Theta_{ih} = \partial y_{it+h} / \partial \mathrm{w}_t'$. The median impulse responses and the corresponding 90% bootstrapped confidence bands are shown in Figure 3. For details on the estimation of the model, the identification of the shocks and the construction of the confidence bands see Appendix B.

[42] This needs to be put in relation to the variance share of US GDP growth explained by the common component, which amounts to 70% (Table 6).

[43] In such a model, a productivity shock raises production, lowers marginal cost and thus prices and interest rates, the latter being determined by a Taylor-style monetary policy reaction function. See also Canova and de Nicoló (2003) who sketch theoretical models that are consistent with our restrictions.

As already emphasized above, Uhlig's (2003) identification scheme helps to disentangle the main US driving forces that spill over to Germany and other common shocks. Besides this advantage, it has two other advantages. First, it prevents us from determining the number of common structural or dynamic shocks. As already indicated above, the number of shocks is not identical with the dimension of f_t if f_t contains, not only the dynamic factors but also their lags. There exist a number of formal (Breitung and Kretschmer, 2005; Bai and Ng, 2007b) and informal criteria (Forni et al., 2000) to determine the number of structural shocks; however, they sometimes lead to inconclusive results. Second, focusing on a smaller shock space when identifying the shocks involves computational gains. They can be substantial, especially when employing an agnostic identification procedure based on a grid search as is done here.

4. Characterization of the US shocks

The identified supply shock accounts for 87% and the real demand shock captures 10% of the forecast error variance of the common component of US GDP over five years.[44] Although not imposed, the impact of the supply shock on the real US economy is long-lasting. The large explanatory power of this shock for fluctuations of US GDP is consistent with the real business cycle view.[45] The US demand shock has temporary real effects. Note that private investment increases quite strongly in response to the demand shock, whereas government expenditures do not go up significantly, and private consumption only increases shortly before entering into a prolonged decline. If one supposes that the price effect precedes the capacity effect, the demand shock could therefore be related to investment rather than to private consumption or fiscal policy. The demand shock has a larger impact on prices than the supply shock: the forecast error variance of the common component of CPI inflation explained by the demand shock is 16%, compared to 14% which is explained by the supply shock.

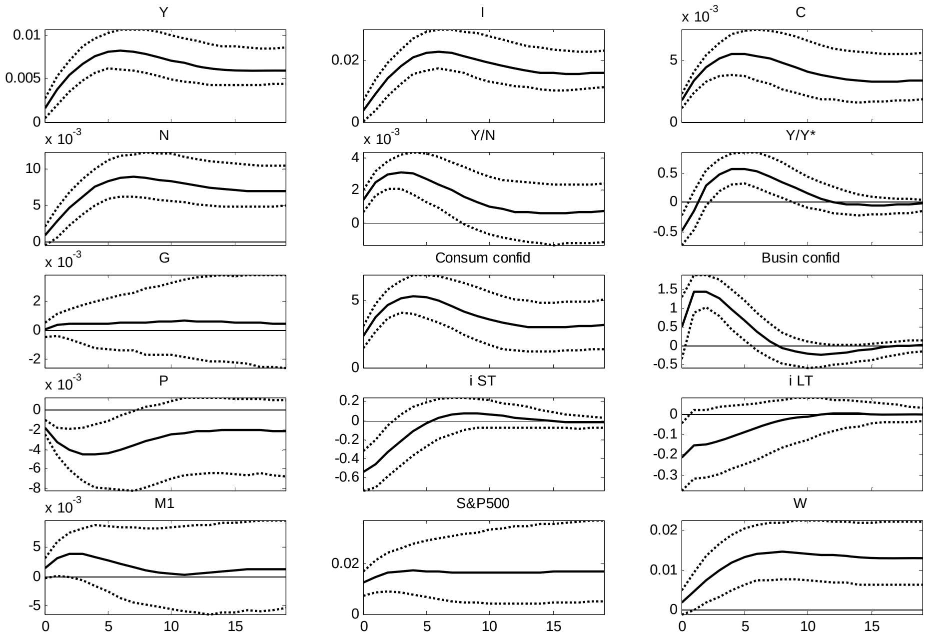
As already pointed out above, there is some evidence suggesting that the two shocks have their origin in the US. We performed the factor analysis and the identification for a data set containing US variables only. The impulse responses of US variables based on this reduced

[44] In the following, we always refer to a 0 to 5-year forecast horizon and the median response.

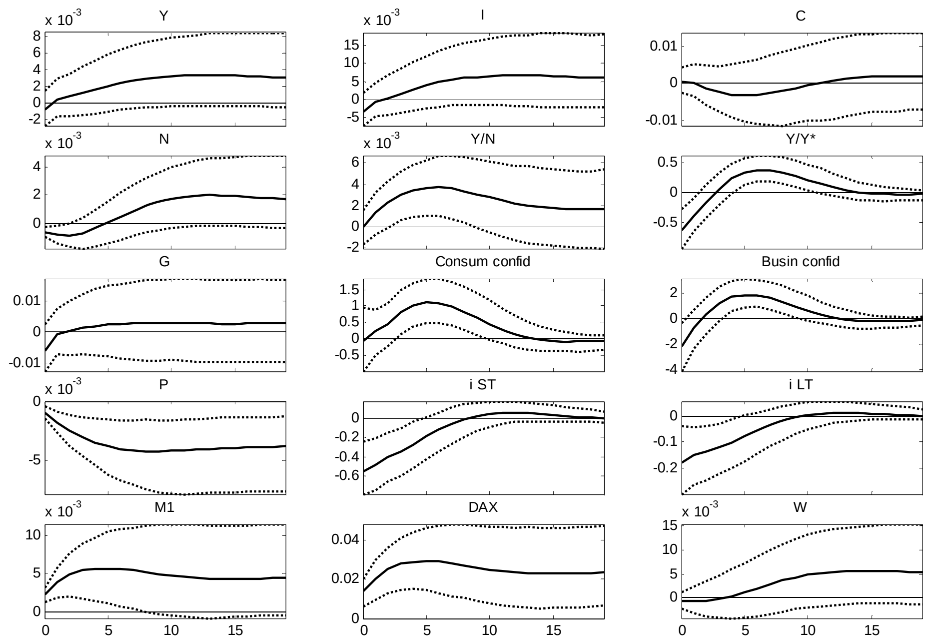
[45] This literature claims that productivity shocks account for the bulk of output variation. On the other hand, we are aware that the contribution of these shocks to business cycle fluctuations crucially depends on the identification scheme (see, for example, Galí, 1999; Peersman, 2005; Canova and de Nicoló, 2003, who all apply distinct identification techniques and find relatively low contributions of productivity shocks).

Figure 3: Impulse-response functions to US shocks

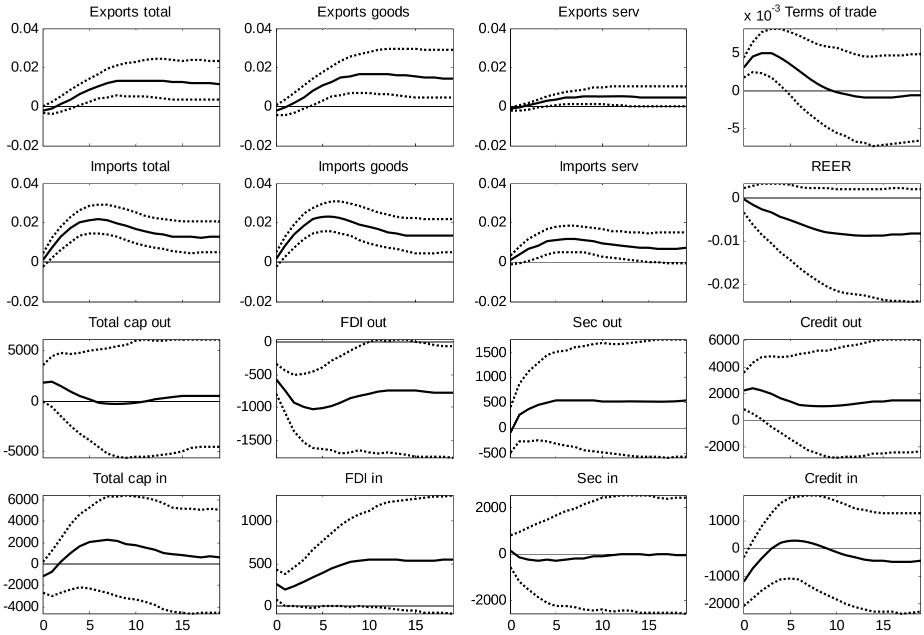
a) US supply shock → US variables



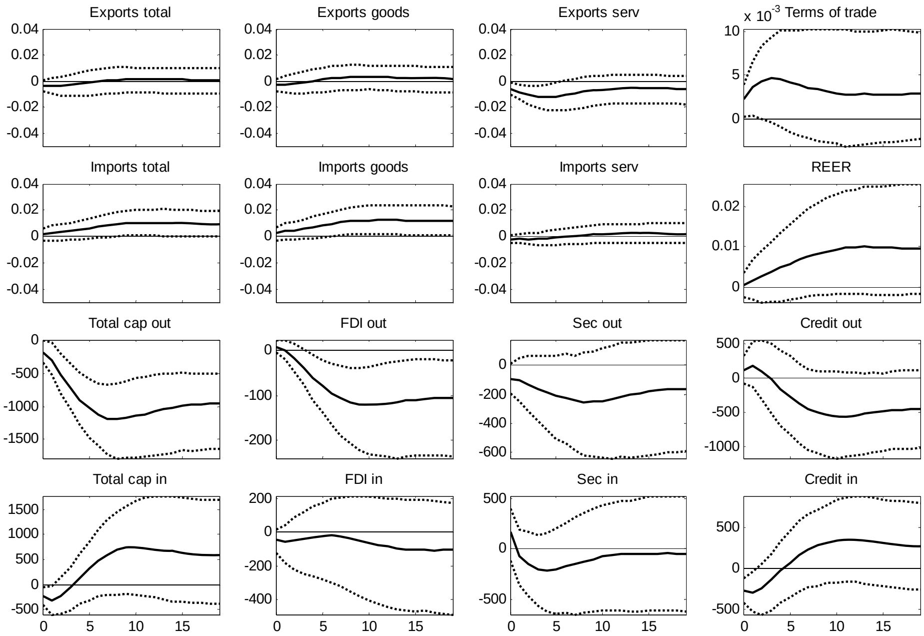
b) US supply shock → German variables



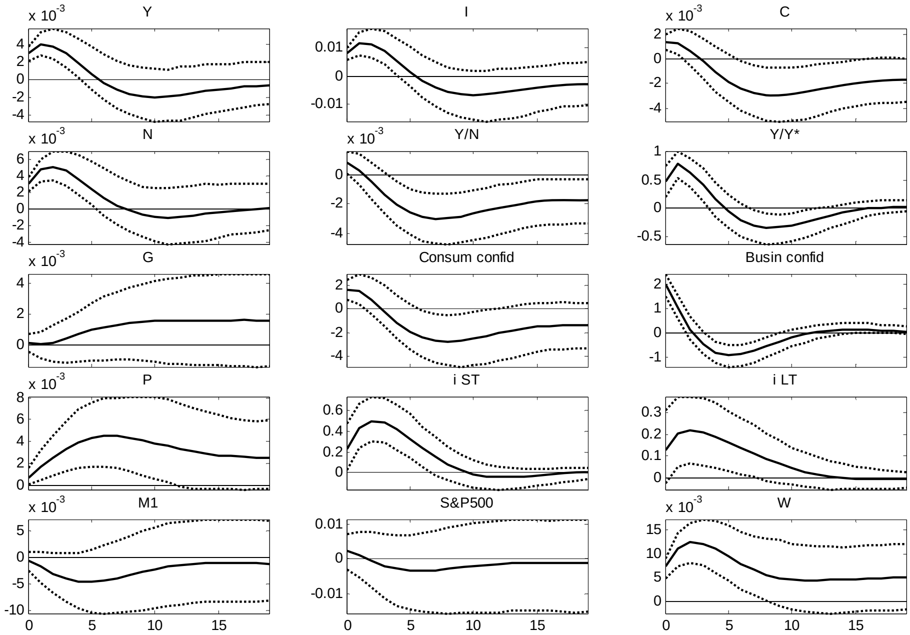
c) US supply shock → US variables



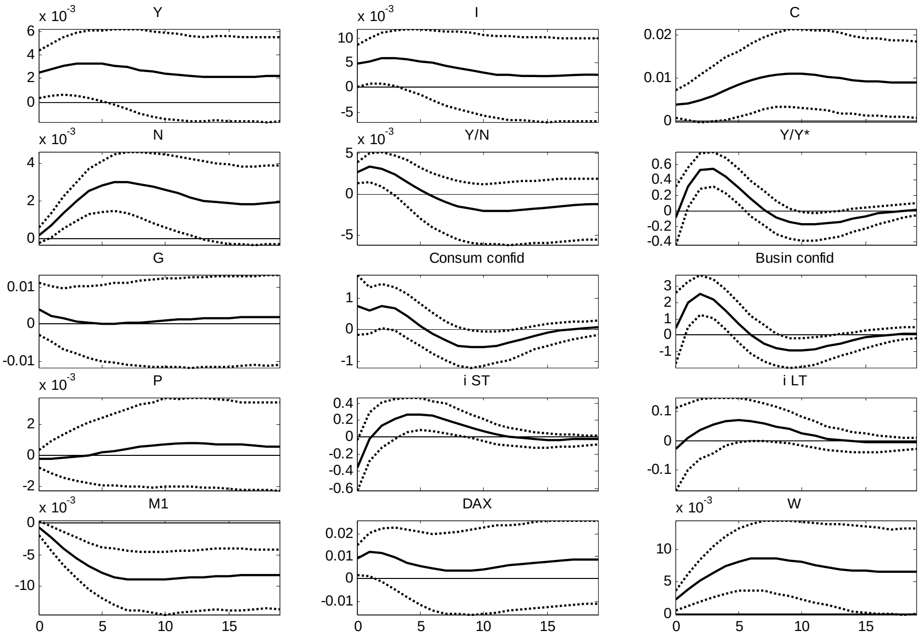
d) US supply shock → German variables



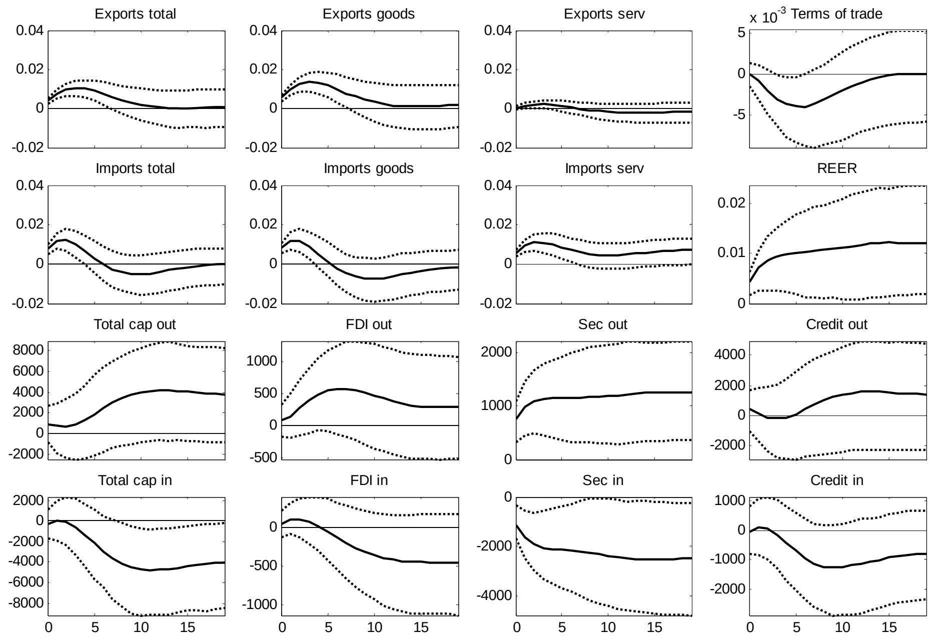
e) US demand shock → US variables



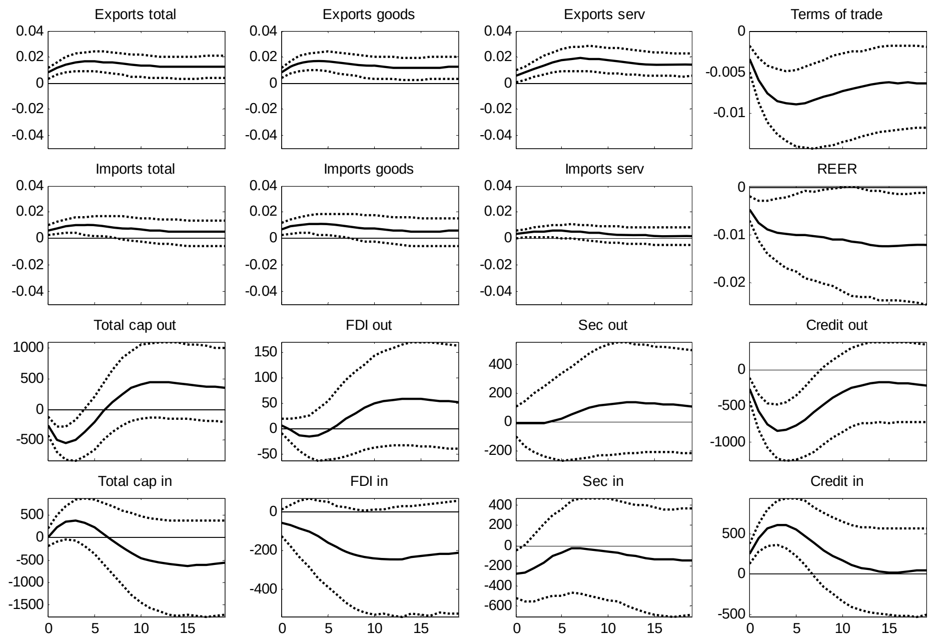
f) US demand shock → German variables



g) US demand shock → US variables



h) US demand shock → German variables



Notes: The size of the shocks is one standard deviation. The median and 90% confidence interval are reported here. Y: real GDP. I: real private investment. C: personal consumption expenditure. N: employment. G: real government expenditures. Y/N: labor productivity. Y/Y*: capacity utilization. P: CPI. iST: short-term interest rate. iLT: long-term interest rate. W: wages. REER: real effective exchange rate. Total cap/FDI/Sec/Credit out: net capital outflows. Total cap/FDI/Sec/Credit in: net capital inflows.

data set and on the bi-national data set look very similar. Let us further anticipate some of our results. We find that the contribution of the shocks to the variance of the forecast error of the common components of world oil prices and other world export prices is small – the US supply shock captures 5% of the forecast error variance of world oil prices and only 15% of other world export prices. The US demand shock explains 8% and 14% respectively. The common component's shares are only 34% and 36% for world oil prices and other world export prices respectively. This indicates that the extracted shocks are not world oil price shocks or other global commodity price shocks which may also drive the common component. It further turns out that the US shocks have larger contemporaneous median effects on US GDP than on German GDP, and the shocks display their maximum impact with some delay in Germany compared to the US. Overall, we therefore carefully conclude that most of the two shocks should stem from the US. This is consistent with Ahmed and Park (1993) and Ahmed et al. (1993), who find that country-specific shocks are more important in explaining output fluctuations in OECD countries than external shocks. While even a large and relatively closed economy like the US may be substantially influenced by external shocks during specific crises periods, this is not plausible for a longer period such as the one considered here.

5. Impact on the German economy and transmission channels

Let us now focus on impulse responses and variance decompositions of German economic variables to the two US shocks. According to Figure 3, the US shocks display effects in Germany that are largely symmetric to the effects in the US. The US supply shock leads to a gradual long-lasting rise in German real GDP of up to 0.3%. The median contribution of this shock to the variance of the forecast error of the common component of German GDP amounts to 9%.[46] The output increase, however, is statistically not significant. Note, in contrast, the significantly positive reaction of labor productivity. Prices respond strongly to the shock. They decline significantly on impact and decrease further to -0.4%, where they remain. Monetary policy seems to play a role in accommodating the US supply shock.

The US demand shock has an immediate impact on German GDP of +0.3%. The impact declines thereafter, becoming insignificant after roughly one year. The corresponding median variance contribution is identical to that of the supply shock (9%). This share is quite large given the relatively low explanatory power of this shock for US economic activity. Investment, consumption, employment, capacity utilization, and productivity also increase. Prices rise gradually, but not significantly. German monetary policy seems to counteract the US demand shock after some delay.

[46] The variance share of the common component of German GDP growth is 79%.

Next, we describe the impact of the US shocks on the variables covering the transmission channels. We focus on real trade variables, exchange rates and the terms of trade, stock prices, long-run interest rates and capital flows, as well as German confidence measures. We address exchange rates within the trade block, although they are determined by capital flows as well. Let us also point out that it is difficult to isolate the confidence channel. We focus on survey-based confidence measures yet bear in mind that financial prices and short-term capital flows also reflect movements in confidence.

A look at the variance shares of the common components of the measures covering German transmission channels yields a first impression of their relevance for business cycle comovements between the US and Germany. Shares are relatively large for German exports and imports (50% and 61%; see Table 6).

Table 6: Variance decompositions

	Variance shares of the common components (CC) of		Forecast error variance of the CC of US variables explained by the		Forecast error variance of the CC of GER variables explained by the	
	US variables	GER variables	US supply shock	US demand shock	US supply shock	US demand shock
Gross domestic product	0.70	0.79	0.87 (0.72 0.94)	0.10 (0.05 0.25)	0.09 (0.01 0.37)	0.09 (0.01 0.31)
Private investment	0.79	0.77	0.76 (0.58 0.87)	0.12 (0.05 0.28)	0.08 (0.01 0.34)	0.05 (0.01 0.24)
Personal cons. expend.	0.59	0.81	0.61 (0.35 0.82)	0.17 (0.04 0.41)	0.04 (0.01 0.22)	0.19 (0.02 0.48)
Employment	0.84	0.73	0.77 (0.57 0.88)	0.10 (0.05 0.26)	0.11 (0.02 0.36)	0.22 (0.03 0.56)
Productivity	0.36	0.37	0.17 (0.06 0.42)	0.24 (0.06 0.49)	0.11 (0.02 0.34)	0.08 (0.03 0.26)
Capacity utilization	0.85	0.87	0.28 (0.14 0.47)	0.31 (0.17 0.44)	0.28 (0.15 0.42)	0.25 (0.13 0.42)
Government expend.	0.17	0.50	0.03 (0.00 0.23)	0.06 (0.00 0.30)	0.03 (0.00 0.20)	0.02 (0.00 0.14)
Consumer confidence	0.36	0.64	0.54 (0.30 0.73)	0.13 (0.04 0.36)	0.19 (0.07 0.41)	0.13 (0.04 0.31)
Business confidence	0.81	0.82	0.37 (0.24 0.50)	0.42 (0.29 0.55)	0.20 (0.09 0.35)	0.20 (0.09 0.35)
Consumer prices	0.89	0.76	0.14 (0.02 0.44)	0.16 (0.02 0.46)	0.33 (0.07 0.61)	0.03 (0.00 0.17)
Short-term int.	0.91	0.86	0.24 (0.08 0.48)	0.36 (0.15 0.60)	0.31 (0.10 0.54)	0.18 (0.07 0.34)
Long-term int.	0.88	0.71	0.12 (0.02 0.38)	0.21 (0.04 0.47)	0.22 (0.04 0.49)	0.09 (0.02 0.24)
M1	0.61	0.37	0.04 (0.00 0.24)	0.05 (0.01 0.25)	0.16 (0.01 0.48)	0.40 (0.12 0.71)
Stock prices	0.21	0.15	0.27 (0.05 0.55)	0.03 (0.00 0.15)	0.33 (0.06 0.62)	0.05 (0.01 0.24)
Nominal wages	0.71	0.44	0.49 (0.18 0.73)	0.17 (0.05 0.45)	0.09 (0.01 0.37)	0.24 (0.03 0.58)
Exports total	0.38	0.50	0.32 (0.06 0.57)	0.10 (0.04 0.35)	0.04 (0.00 0.17)	0.34 (0.09 0.61)
Goods	0.36	0.50	0.33 (0.08 0.58)	0.11 (0.04 0.39)	0.04 (0.00 0.20)	0.33 (0.08 0.61)
Services	0.08	0.13	0.19 (0.02 0.47)	0.05 (0.01 0.23)	0.10 (0.01 0.37)	0.32 (0.09 0.58)
Imports total	0.59	0.61	0.48 (0.24 0.67)	0.10 (0.04 0.24)	0.14 (0.01 0.42)	0.11 (0.02 0.38)
Goods	0.54	0.57	0.47 (0.23 0.67)	0.10 (0.04 0.27)	0.17 (0.01 0.44)	0.11 (0.02 0.38)
Services	0.28	0.17	0.29 (0.04 0.60)	0.20 (0.04 0.52)	0.05 (0.01 0.22)	0.07 (0.01 0.31)
Terms of trade	0.57	0.58	0.08 (0.02 0.25)	0.06 (0.01 0.27)	0.07 (0.01 0.34)	0.29 (0.08 0.55)
Real effective exch. rate	0.66	0.72	0.08 (0.00 0.39)	0.19 (0.01 0.49)	0.09 (0.00 0.40)	0.17 (0.01 0.46)
Net capital outflows, total	0.04	0.03	0.03 (0.01 0.26)	0.11 (0.01 0.36)	0.47 (0.22 0.69)	0.09 (0.03 0.28)
FDI	0.03	0.02	0.20 (0.03 0.49)	0.06 (0.01 0.27)	0.23 (0.03 0.53)	0.06 (0.01 0.26)
Securities	0.07	0.05	0.06 (0.00 0.33)	0.28 (0.04 0.57)	0.07 (0.00 0.32)	0.04 (0.00 0.23)
Credits	0.06	0.06	0.06 (0.01 0.33)	0.05 (0.00 0.22)	0.12 (0.01 0.37)	0.15 (0.05 0.38)
Net capital inflows, total	0.03	0.05	0.05 (0.01 0.26)	0.16 (0.01 0.46)	0.09 (0.01 0.33)	0.07 (0.02 0.28)
FDI	0.02	0.07	0.13 (0.01 0.37)	0.07 (0.01 0.28)	0.04 (0.00 0.28)	0.12 (0.01 0.46)
Securities	0.02	0.14	0.04 (0.00 0.24)	0.24 (0.01 0.56)	0.04 (0.00 0.18)	0.03 (0.00 0.16)
Credits	0.04	0.02	0.04 (0.01 0.16)	0.05 (0.01 0.23)	0.09 (0.01 0.29)	0.11 (0.04 0.35)

Notes: The variance shares of the common components refer to I(0) series, i.e., GDP refers to the first difference of GDP etc. The forecast error variance refers to forecast horizons 0 to 5 years and the levels of the series.

Percentages are notably smaller for trade in services compared to goods. The common components of the real effective exchange rate and the terms of trade also exhibit high variance shares (72% and 58%). Similarly high values are found for German confidence measures. As regards financial markets, the variance share of German long-term interest rates explained by the common component amounts to 71%, but other variance shares (of capital flows and stock prices) are relatively low (2% to 15%). Let us now turn to impulse responses of these variables and variance decompositions.

Real trade and relative prices. The US supply shock reduces net exports in Germany. This reduction is based on an immediate significant and temporary decline in German exports of services. In addition, German imports become significantly positive after six quarters. The decline in German net exports is accompanied by a real appreciation of the currency – possibly because German prices decline less than US prices during the first year after the shock – and an improvement in the terms of trade due to strongly declining import prices (not reported here). The US demand shock, by contrast, triggers a significantly positive and permanent response of the trade balance, caused by persistently rising exports. Imports also exhibit a positive response; however, this is transitory and smaller than the response of exports. In contrast to the supply shock, the terms of trade worsen and the German currency strongly and permanently depreciates in real terms. The variance decompositions also show that trade is affected by the US shocks to a non-negligible extent. The supply shock explains 4% of the forecast error variance of exports and 14% of imports, and the demand shock captures 34% of exports and 11% of imports. The former accounts for 7% and 9% of the terms of trade and the real effective exchange rate; the corresponding values with respect to the latter are 29% and 17%.

Financial markets. The US supply shock leads to a lasting increase in German stock prices, paralleling the reaction of US stock prices. German long-term interest rates decline. Net outflows of FDI decrease. The negative instantaneous reaction of net inflows of credits are, however, difficult to interpret. The US real demand shock does not have a significant effect on US stock prices. Consequently, German stock prices are barely affected; only a small immediate increase of the DAX is observed. Long-term interest rates rise after a delay. Net outflows of credits decrease, and net inflows of credits rise significantly. Variance decompositions underscore the fact that the influence of the supply shock on financial prices is larger than that of the demand shock, explaining 33% compared with 5% of the forecast error variance of the DAX and 22% compared with 9% of the forecast error variance of long-term interest rates.[47]

Confidence. The US supply shock triggers delayed positive responses of German business and consumer confidence. The demand shock affects both measures positively in the short run. They decline thereafter.

[47] The large contributions of the US supply shock to fluctuations of the common component of net total capital and FDI outflows are striking. As pointed out above, the variance shares explained by the common components of these variables are, however, very small.

In summary, trade seems to be most important for the international business cycle transmission. The US supply shock may have increased the US supply of inputs, lowered German import prices and raised German real imports. The US demand shock may have triggered an increase in US import demand and hence German exports. Moreover, trade seems to be influenced by relative prices. Real exchange rates and the terms of trade alter competitiveness, consumer spending power and, ultimately, trade in Germany; they also dampen the transmission of the supply shock and enhance the propagation of the demand shock. The trade channel seems to be more important for spreading the demand shock than for transmitting the supply shock. Monetary policy reacts as expected to the strong price movements after the US supply shock, accommodating the latter. It may display real effects in the medium run, consistent with the transmission lags of monetary policy. It is difficult to draw unambiguous conclusions on the role of financial markets and confidence. Stock price and confidence movements could have enhanced the transmission of the US shocks, especially the US supply shock. Capital flows seem to support shock transmission. However, the picture is somewhat blurred, and the variance shares explained by the common components are small. At present, it is not possible to draw clear conclusions on which categories of capital flows are particularly relevant.

Our results regarding the aggregate impact of US shocks on Germany are roughly consistent with results obtained on the basis of VAR models (cf. GCEE, 2001; Artis et al., 2006; Artis et al., 2007; Canova and Marrinan, 1998). Comparison is somewhat exacerbated by the fact that these studies focus on aggregate shocks to economic activity and do not distinguish between different structural shocks. The effects found in the present study are somewhat larger than what is usually found by means of structural macroeconometric multi-country models (cf. Dalsgaard et al., 2001; IMF, 2001b; GCEE, 2001) – which was to be expected, given that not all channels are included in those models. Our findings with respect to the transmission channels are roughly in line with those of Artis et al. (2007) and Canova et al. (2007a). The former study finds that exchange rates and monetary policy are most important for the international shock transmission and that capital flows may also be supportive, while the latter indicates a significant role for trade, monetary policy and consumer spending power.

6. Latest US boom and recession - to what extent was Germany affected?

In this section, we investigate to what extent Germany was able to benefit from the boom in the US between 1995 and 2000, and whether the current German slump was caused by the US recession in 2001. We assess the historical decomposition of German economic activity with respect to the shocks derived from our model. For this purpose we compare true series with hypothetical series that would have evolved had they been driven solely by the US shocks. Hypothetical series are generated by setting shocks other than the US supply and demand shocks for all points in time to zero and making use of the estimated matrices of factor loadings and VAR coefficients, as well as the matrices associated with the shock identification.[48]

German GDP growth was slightly lower than it would have been had it been driven solely by US shocks between 1995 and 2000. The mean (annualized) growth rate of German GDP was 1.7%, and the corresponding rate for the hypothetical series is 1.9%. Interestingly, it was mainly the US demand shock which spread positively to Germany: German GDP would have grown by 2.3% if it had only been driven by the US demand shock, whereas its growth rate would have amounted to 1.7% if it had only been affected by the US supply shock. Other influences must have held down economic activity in Germany and overcompensated positive US influences. The negative movement of employment and private demand in the second half of the 1990s when compared with the movement if only driven by US shocks is striking and supports the widespread view that domestic influences, like overcapacity in the construction sector after the unification and small productivity gains coupled with relatively high nominal wage increases in the eastern part of Germany, suppressed these variables.

The mean growth rate of German GDP in 2001 was 0.5%, whereas the rate of change of the hypothetical series was lower (0.1%). This finding is consistent with Artis et al. (2005) who also attribute most of the 2001 downturn in Germany to a shock to US output growth. Our analysis suggests further that it was the negative US demand shock that contributed substantially to the economic slowdown in Germany: German GDP would not have changed at all if it had been driven by the US demand shock only, whereas its growth rate would have amounted to 2.2% if it had been determined exclusively by the US supply shock.

As regards the transmission channels, hypothetical trade series highlight the large US influence. They move very much in parallel with the true series. Hypothetical financial market series and the corresponding true series, by contrast, do evolve similarly. Interestingly, confidence was not much influenced by US shocks in the mid-1990s, but the lines move much more in parallel since the end of the 1990s. This may simply reflect the stronger business cycle transmission via trade or financial markets by the end of the sample period. However, another interpretation is that the confidence channel has become relevant only in the last few years. This would also indicate that it is too early to expect this channel to show up in existing empirical studies. Our results are finally consistent with IMF (2001a) and Anderton et al. (2004) who find an increase in the correlation between US and euro-area confidence measures in recent years.

[48] $f_{1975:1}$ is chosen as a starting value for f_t. Notice that not a single, but 6 rotations of the first two principal component shock point estimates satisfy our identification restrictions. We refer here to the median point estimates of the two structural US shocks and the corresponding rotation matrix. Notice also that we re-added the means (but not deterministic trends) to the hypothetical series. Differences between growth rates discussed in this section and official growth rates published by the statistical offices may arise due to the removal of deterministic trends from our series. Detailed results (with respect to different variables and the two shocks) are not reported in the paper, but are available upon request.

It should also be noted that our model probably overestimates the US contribution to the German expansion and underestimates the contribution of the US recession to the German slump. The reason is that our linear model cannot account for transmission asymmetry. Studies employing non-linear empirical models find that negative real shocks are transmitted to a larger extent internationally than positive shocks (Artis et al., 2007; GCEE, 2001; Canova et al., 2007a; Osborn et al., 2005). These asymmetries can be explained in terms of nominal rigidities (which are stronger downwards than upwards), menu costs, difficulties for firms facing stronger demand to expand their capacities, and informational asymmetry between lenders and borrowers (Ball and Mankiw, 1994; Peersman and Smets, 2002).

7. Conclusion

In this chapter, we have investigated the transmission of US macroeconomic shocks to the German economy between 1975 and 2002 by means of a large-scale structural dynamic factor model. This framework allows us to simultaneously assess the responses of a large set of real and nominal German variables and investigate the role of many transmission channels, including "new" channels such as stock markets, foreign direct investment, international bank lending and the confidence channel. To that extent, it has advantages over other models used in this context, which are not able to investigate as many transmission channels simultaneously.

We have identified two US shocks: one supply shock and one real demand shock. We find that these shocks affect the US economy and the German economy largely symmetrically. That is, the supply shock raises output and lowers prices and interest rates, while the demand shock increases all three variables in both countries. Reactions of German output to the US supply shock and of German prices to the US demand shock, however, are not statistically significant.

As concerns the transmission channels: trade, influenced by relative price movements, seems to play the dominant role in the transmission, especially for the propagation of US demand shocks. Besides trade, monetary policy reacts to relatively strong German price movements following US supply shocks and seems to influence the medium-run impact of US shocks. No clear conclusion can be drawn on the role of financial markets and confidence. Interestingly, German confidence has been driven notably by US shocks only since the end of the 1990s. This might indicate that the confidence channel has become relevant only in the last few years.

Historical decompositions, finally, show that negative domestic factors more than compensated for positive US influences during the US boom between 1995 and 2000 in

Germany. By contrast, the US recession in 2001 seemed to be the main culprit for the German slump.

Appendix A

The data set incorporates 296 variables. The data are selected such that the US and German economies, as well as the international integration of both countries, are represented in a balanced way. Less data are generally available for Germany than for the US. We thus mainly confine ourselves to including series which are available for both countries. The data set contains variables covering the real side and the nominal side of the US economy and the German economy. In addition, variables approximating the international economic integration of the two countries and those capturing global factors are included, as outlined in the main text. The data are taken from various national and international sources. They are seasonally adjusted and quarterly. This frequency is chosen in order to include national accounts series, which are generally not available on a monthly basis. Originally monthly series were converted into quarterly series. The X11 seasonal adjustment method was applied to originally not seasonally adjusted series. Logarithms are taken for all non-negative series that were not already in rates or percentage units.

The study is carried out for the period from the first quarter of 1975 to the fourth quarter of 2002. One reason for selecting this period is that important capital controls were abolished in Germany in 1974. In the US, the last capital controls were abandoned in 1973. Moreover, this period corresponds to the post-Bretton Woods flexible exchange rate regime period. Another advantage of this starting date is that potentially extraordinary influences of the first oil price shock in 1973-74 are eliminated.

When constructing the data set, one problem to be addressed is the break in the series resulting from German unification. Most German series were extended by applying west German growth rates to the German levels retrospectively from 1991 onwards. Visual inspection of the series does not suggest a break.

Non-stationarity is removed by differencing and/or deterministic detrending. We performed the analysis with several data sets which differ in the treatment of the series. Results are robust across different specifications. A detailed description of the set finally used may be found in Table A1.

Table A1: Data

#	Variables	Treatment	Source
	US variables		
1	Gross domestic product, real	2	BEA
2	Personal consumption expenditures, real	2	BEA
3	Durables, total	2	BEA
4	Durables, motor vehicles and parts	2	BEA
5	Durables, furniture and household equipment	2	BEA
6	Durables, other	2	BEA
7	Nondurables, total	2	BEA
8	Food	2	BEA
9	Clothing shoes	2	BEA
10	Gasoline fuel oil and other energy goods	2	BEA
11	Other	2	BEA
12	Sevices, total	2	BEA
13	Housing	2	BEA
14	Housing operation	2	BEA
15	Transportation	2	BEA
16	Medical care	2	BEA
17	Other	2	BEA
18	Gross private domestic investment, real	2	BEA
19	Private fixed investment	2	BEA
20	Nonresidential	2	BEA
21	Nonresidential, structures	2	BEA
22	Nonresidential builings incl. farms	2	BEA
23	Utilities	2	BEA
24	Mining exploration, shafts, and wells	2	BEA
25	Other structures	2	BEA
26	Nonresidential, equipment and software	2	BEA
27	Information processing and related equipment	2	BEA
28	Industrial equipment	2	BEA
29	Transportation equipment	2	BEA
30	Residential	2	BEA
31	Structures	2	BEA
32	Equipment	2	BEA
33	Change in priv. inventories, real	0	BEA
34	Government cons. expend. and gross investment, real	2	BEA
35	Employment, total private	2	BLS
36	Goods-producing	2	BLS
37	Natural resources and mining	2	BLS
38	Construction	2	BLS
39	Manufacturing	2	BLS
40	Service-producing	2	BLS
41	Private service producing	2	BLS
42	Trade, transportation and utilities	2	BLS
43	Wholesale trade	2	BLS
44	Retail trade	2	BLS
45	Transportation and utilities	2	BLS
46	Information	2	BLS
47	Financial activities	2	BLS
48	Professional and business services	2	BLS
49	Education and health services	2	BLS
50	Leisure and hospitality	2	BLS
51	Government	2	BLS
52	Unemployed, civilian labor force	1	BLS
53	Unemployment rate	0	BLS
54	Average weekly hours, manufacturing	2	BLS
55	Hours worked, nonfarm business sector	2	BLS
56	Industrial production, total	2	FRB
57	Manufacturing	2	FRB

58	Consumer goods, total	2	FRB
59	Durables	2	FRB
60	Nondurables	2	FRB
61	Business equipment	2	FRB
62	Defense and space equipment	2	FRB
63	Construction supplies	2	FRB
64	Other business supplies	2	FRB
65	Materials	2	FRB
66	Capacity utilization, total	0	FRB
67	Manufacturing	0	FRB
68	Durables	0	FRB
69	Nondurables	0	FRB
70	Other	0	FRB
71	Mining	3	FRB
72	Electric and gas utilities	0	FRB
73	Computers, communications equipm., semiconductors	0	FRB
74	Primary and semifinished processes	3	FRB
75	Finished processing	3	FRB
76	Productivity (output per hour, all persons), nonfarm bus.	2	BLS
77	Unit labor costs, all persons, nonfarm business sector	2	BLS
78	Unit nonlabor payments, all persons, nonfarm bus.	2	BLS
79	Wages and salaries, total, nominal	2	BEA
80	Private industries	2	BEA
81	Commodities-producing industries	2	BEA
82	Manufacturing	2	BEA
83	Distributive industries	2	BEA
84	Service industries	2	BEA
85	Government	2	BEA
86	Disposable personal income, real	2	BEA
87	Personal savings, nominal	2	BEA
88	Saving rate	0	BEA
89	CPI, total	2	BLS
90	Food and beverages	2	BLS
91	Housing	2	BLS
92	Apparel	2	BLS
93	Transportation	2	BLS
94	Medical care	2	BLS
95	Other goods and services	2	BLS
96	Commodities	2	BLS
97	PPI, total (finished goods)	2	BLS
98	Capital equipment	2	BLS
99	Intermediate materials and supplies	2	BLS
100	Materials and components for construction	2	BLS
101	Processed fuels and lubricants	2	BLS
102	Crude materials	2	BLS
103	Implicit price deflator, GDP	2	BEA
104	Personal consumption expenditures	2	BEA
105	Private fixed investment	2	BEA
106	Government cons. and investment expend.	2	BEA
107	Private inventories	2	BEA
108	M1	2	FRB
109	M2	2	FRB
110	M3	2	FRB
111	Consumer installment loan: total outstanding	2	FRB
112	Real balances (M1/GDP deflator)	2	FRB/BEA
113	Total reserves	2	FRB
114	Non borrowed reserves	2	FRB
115	Consumer confidence, total	3	Conference Board
116	Consumer confidence, current situation	3	Conference Board
117	Consumer confidence, expectations	0	Conference Board
118	Business confid. (ISM Purchasing Managers Index, Manuf. Survey)	0	ISM
119	Business confid. expect. (FRB Philad., 6 months forec., diff. index)	0	FRB Philad.
120	Federal Funds Rate	0	FRB

121	Commercial paper 3 months (non financial)	0	FRB
122	Commercial paper 3 months (financial)	0	FRB
123	CDs secondary market 3 months	0	FRB
124	Treasury bills, constant maturity, 1 year	0	FRB
125	Treasury bills, constant maturity, 10 year	0	FRB
126	Conventional mortgage rates 30 years	0	FRB
127	Moody's AAA seasoned	0	FRB
128	Moody's BBB seasoned	0	FRB
129	Interest rate spread, 3m CDs - Federal Funds Rate	0	FRB
130	Interest rate spread, 1y treasury yield - Federal Funds Rate	0	FRB
131	Interest rate spread, 10y treasury yield - Federal Funds Rate	0	FRB
132	Interest rate spread, Moody's AAA - Federal Funds Rate	0	FRB
133	Interest rate spread, Moody's BBB - Federal Funds Rate	0	FRB
134	S&P 500 Composite, Index	2	Datastream
135	S&P Price Earnings Ratio	0	Datastream
136	Exports total, real	2	BEA
137	Exports goods	2	BEA
138	Exports services	2	BEA
139	Imports total, real	2	BEA
140	Imports goods	2	BEA
141	Imports services	2	BEA
142	Price exports total	2	BEA
143	Price exports goods	2	BEA
144	Price exports services	2	BEA
145	Price imports total	2	BEA
146	Price imports goods	2	BEA
147	Price imports services	2	BEA
148	Terms of Trade (price of exports/price of imports)	2	BEA
149	Net capital outflows, total, nominal	3	BEA
150	Direct investment	3	BEA
151	Securities	3	BEA
152	Credits banks	3	BEA
153	Credits non-banks	3	BEA
154	Net capital inflows, total, nominal	3	BEA
155	Direct investment	3	BEA
156	Securities	3	BEA
157	Credits banks	3	BEA
158	Credits non-banks	3	BEA
159	Exchange rate (US Broad Index), real	2	FRB
160	Current account balance	3	BEA
GER variables			
161	Gross domestic product, real	2	StaBu
162	Consumption of private households, real	2	StaBu
163	Food beverages and tabacco	2	StaBu
164	Clothing and footwear	2	StaBu
165	Housing water and energy	2	StaBu
166	Furnishings and households equipment	2	StaBu
167	Transport and communications	2	StaBu
168	Recreation and culture	2	StaBu
169	Hotel and restaurant services	2	StaBu
170	Other purposes	2	StaBu
171	Private gross fixed capital formation, real	2	StaBu
172	Machinery and equipment	2	StaBu
173	Metal products and machinery	2	StaBu
174	Transport equipment	2	StaBu
175	Construction	2	StaBu
176	Housing	2	StaBu
177	Other construction	2	StaBu
178	Building construction	2	StaBu
179	Civil and underground engineering	2	StaBu
180	Other products	2	StaBu

181	Change in private and public inventories, real	0	StaBu
182	Government consumption expenditures, real	2	StaBu
183	Government fixed investment, real	2	StaBu
184	Employment, total	2	StaBu
185	Agriculture, forestry and fishing	2	StaBu
186	Industry, incl. energy, excl. construction	2	StaBu
187	Construction	2	StaBu
188	Trade and transport	2	StaBu
189	Financial renting and business activities	2	StaBu
190	Other service activities	2	StaBu
191	Unemployment (registered unemployed)	2	OECD
192	Unemployment rate	0	Bundesbank
193	Industrial production, total	2	Bundesbank
194	Manufacturing	2	Bundesbank
195	Construction	2	Bundesbank
196	Chemicals	2	Bundesbank
197	Investment goods	2	Bundesbank
198	Consumer goods	2	Bundesbank
199	Intermediate goods	2	Bundesbank
200	Capacity utilization, manufacturing	0	IFO Inst. Munich
201	Manufacturing trade excl. foodstuffs	0	IFO Inst. Munich
202	Capital goods production	0	IFO Inst. Munich
203	Consumer goods	3	IFO Inst. Munich
204	Durable goods	3	IFO Inst. Munich
205	Nondurable goods	0	IFO Inst. Munich
206	Intermediate goods excl. chemicals	0	IFO Inst. Munich
207	Raw materials production excl. chemicals	0	IFO Inst. Munich
208	Productivity (output per hour), industry	2	Bundesbank
209	Unit labor costs, total	2	StaBu
210	Wages and salaries, total	2	StaBu
211	Agriculture, forestry and fishing	2	StaBu
212	Industry, incl. energy, excl. construction	2	StaBu
213	Construction	2	StaBu
214	Trade and transport	2	StaBu
215	Financial renting and business activities	2	StaBu
216	Other service activities	2	StaBu
217	Disposable income	2	StaBu
218	Personal savings	2	StaBu
219	Savings ratio	0	StaBu
220	CPI, total	2	OECD
221	Non food, non energy	2	OECD
222	Energy	2	OECD
223	Food and alcohol free drinks	2	OECD
224	Housing rental services	2	OECD
225	PPI, total	2	OECD
226	Investment goods	2	OECD
227	Manufacturing industry	2	OECD
228	Deflator, GDP	2	StaBu
229	Consumption (private and public)	2	StaBu
230	Gross investment (private and public)	2	StaBu
231	Gross fixed investment (private and public)	2	StaBu
232	M1	2	Bundesbank
233	M2	2	Bundesbank
234	M3	2	Bundesbank
235	Loans, lending to domestic enterprises and households	2	Bundesbank
236	Lending to domestic enterprises	2	Bundesbank
237	Lending to domestic self-employed households	2	Bundesbank
238	Lending to domestic employees and other individuals	2	Bundesbank
239	Real balances (M1/GDP deflator)	2	Bundesbank/StaBu
240	Consumer confidence	0	European Comm.

241	Business confidence	0	IFO Inst. Munich
242	Business confidence, expectations (next 6 months)	0	IFO Inst. Munich
243	Overnight rate	0	Bundesbank
244	3m money market rate	0	Bundesbank
245	Long-term government bonds	0	Bundesbank
246	Mortgage rates 3 years	0	Bundesbank
247	Mortgage rates 10 years	0	Bundesbank
248	Interest rate spread (3m - overnight rate)	0	Bundesbank
249	Interest rate spread (10y - overnight rate)	0	Bundesbank
250	DAX	2	Bundesbank
251	Exports total, real	2	StaBu
252	Exports goods	2	StaBu
253	Exports services	2	StaBu
254	Imports total, real	2	StaBu
255	Imports goods	2	StaBu
256	Imports services	2	StaBu
257	Price exports total	2	StaBu
258	Price exports goods	2	StaBu
259	Price exports services	2	StaBu
260	Price imports total	2	StaBu
261	Price imports goods	2	StaBu
262	Price imports services	2	StaBu
263	Terms of Trade (price of exports/price of imports)	2	Bundesbank
264	Exports total to the US, nominal	2	Bundesbank
265	Exports total to the EU, nominal	2	Bundesbank
266	Exports total to other countries, nominal	2	Bundesbank
267	Imports total from the US, nominal	2	Bundesbank
268	Imports total from the EU, nominal	2	Bundesbank
269	Imports total from other countries, nominal	2	Bundesbank
270	Net capital outflows, total, nominal	3	Bundesbank
271	Direct investment	3	Bundesbank
272	Equity capital	3	Bundesbank
273	Reinvested earnings	3	Bundesbank
274	Intercompany debt	3	Bundesbank
275	Securities	3	Bundesbank
276	Stocks	3	Bundesbank
277	Bonds	3	Bundesbank
278	Credits banks	3	Bundesbank
279	Credits non-banks	3	Bundesbank
280	Net capital inflows, total, nominal	3	Bundesbank
281	Direct investment	3	Bundesbank
282	Equity capital	3	Bundesbank
283	Reinvested earnings	3	Bundesbank
284	Intercompany debt	3	Bundesbank
285	Securities	3	Bundesbank
286	Stocks	3	Bundesbank
287	Bonds	3	Bundesbank
288	Credits banks	3	Bundesbank
289	Credits non-banks	3	Bundesbank
290	Real effective exchange rate DM (Euro)	2	Bundesbank
291	Current account balance	3	Bundesbank
Global variables			
292	Exchange rate US dollar/DM (Euro), nominal	2	Bundesbank
293	World crude petroleum price	2	Datastream
294	World price all exports excl. fuel	2	Datastream
295	World gold price	2	Datastream
296	World food price	2	Datastream

Notes: 0: no transformation, 1: logarithm, 2: first difference of logarithm, 3: first difference; BEA: Bureau of Economic Analysis, BLS: Bureau of Labor Statistics, FRB: Federal Reserve Board, ISM: Institute for Supply Management, StaBu: Federal Statistical Office, Germany.

Appendix B

This appendix describes the estimation of the factor model, the identification of the structural US shocks and the construction of the confidence bands. We first estimate f_t by applying static principal component analysis to y_t.

$$\hat{\mathrm{f}}_t = \hat{\mathrm{V}}\mathrm{y}_t, \tag{A1}$$

where $\hat{\mathrm{V}}=[\hat{\mathrm{V}}_1 \quad \cdots \quad \hat{\mathrm{V}}_{\mathrm{N}}]$ is the $r \times N$ matrix of eigenvectors corresponding to the largest r eigenvalues of the sample correlation matrix. $\hat{\mathrm{V}}$ is an estimate of the matrix of factor loadings Λ. The estimated vector of factors $\hat{\mathrm{f}}_t$ has a VAR(1) representation:

$$\hat{\mathrm{f}}_t = \mathrm{A}_1\hat{\mathrm{f}}_{t-1} + \mathrm{u}_t. \tag{A2}$$

OLS is applied to each equation yielding the estimated coefficient matrix $\hat{\mathrm{A}}_1$ and the reduced form VAR residuals $\hat{\mathrm{u}}_t$.

The residuals $\hat{\mathrm{u}}_t$ are orthogonalized, as in Uhlig (2003), by means of the Cholesky decomposition, but any other orthogonalization would work as well. Hence $\mathrm{cov}(\hat{\mathrm{u}}_t) = \hat{\mathrm{Q}}\hat{\mathrm{Q}}'$, with $\hat{\mathrm{Q}}$ being the $r \times r$ lower triangular Cholesky matrix. Then

$$\hat{\mathrm{v}}_t = \hat{\mathrm{Q}}^{-1}\hat{\mathrm{u}}_t \tag{A3}$$

and $\mathrm{cov}(\hat{\mathrm{v}}_t) = \mathrm{I}_r$. The vector of impulse response functions of variable y_{it} in period k to v_t is $\varphi_{ik} = \Lambda_i'\mathrm{A}_1^{\mathrm{k}}\mathrm{Q}$, where Λ_i is the i th column of Λ, and the corresponding variance of the k-step-ahead forecast error is $\varphi_{i0}\varphi'_{i0} + \quad \ldots \quad + \varphi_{ik}\varphi'_{ik}$.

We will now attempt to identify the main driving forces of the US economy. We label the vector of common driving forces $\omega_t = [\omega_{1t} \quad \cdots \quad \omega_{rt}]'$, with ω_{jt}, $j = 1,\ldots,r$, being scalars, and we suppose that v_t is linearly related to them through the $r \times r$ matrix G: $\mathrm{v}_t = \mathrm{G}\omega_t$. The aim is to choose G so that the first shock explains as much of the forecast error variance of (the common component of) US GDP over a certain horizon k as possible, the second shock (which is orthogonal to the first) explains as much of the remaining forecast error variance as possible, etc. US GDP is a proxy of economic activity. We choose k to be 19, which yields the variance of the five–years-ahead forecast error.[49] Because v_t is the vector of orthogonal shocks, we can write the forecast error variance accounted for by r shocks as the sum of the forecast error variance accounted for by each shock. Let us now focus on a single or the first shock. The forecast error variance accounted for by this shock is $\varphi_{i0}\mathrm{g}_1\mathrm{g}'_1\varphi'_{i0} + \quad \ldots \quad + \varphi_{ik}\mathrm{g}_1\mathrm{g}'_1\varphi'_{ik}$, where i stands for US GDP and g_1 is the first column of G.

[49] Uhlig (2003) and Altig et al. (2002) also choose US output and $k = 19$.

Uhlig (2003) shows that g_1 should be chosen such that $g'_1 S_{ik} g_1$ is maximized, where $S_{ik} = (k+1-0)\varphi'_{i0}\varphi_{i0} + \ \ldots \ + (k+1-k)\varphi'_{ik}\varphi_{ik}$, subject to $g'_1 g_1 = 1$, which is a normalization condition.[50] The Lagrangian may be set up as follows:

$$L = g'_1 S_{ik} g_1 - \gamma(g'_1 g_1 - 1), \tag{A4}$$

where γ is the Lagrange multiplier. The solution to this problem g_1 is the first eigenvector of S_{ik} and γ is the corresponding eigenvalue. The shock associated with g_1 may therefore be called the first principal component shock. Hence, G is the matrix of eigenvectors of S_{ik}, $(g_1 \quad g_2 \quad \ldots \quad g_r)$, where g_j is the eigenvector corresponding to the j^{th} principal component shock and $j = 1,\ldots,r$.

As explained in the main text, two shocks explain the bulk, i.e. 98%, of the variance of the forecast error of the common component of US GDP, given by the sum of the first two eigenvalues of S_{ik}. Hence, we concentrate on the first two estimated principal component shocks, represented by $\hat{\omega}_t = [\hat{\omega}_{1t} \quad \hat{\omega}_{2t}]'$, and neglect the remaining ones.

Up to now, the principal component shocks are identified up to a rotation. The vector of principal component shocks ω_t is linearly related to the vector of structural shocks $w_t = [w_{1t} \quad w_{2t}]'$ through the 2×2 orthonormal rotation matrix R. Hence,

$$\hat{w}_t = R\hat{\omega}_t, \tag{A5}$$

with $\mathrm{cov}(\hat{w}_t) = I_2$ and $R'R = I_2$. The objective is now to fix a rotation which yields plausible results in terms of impulse-response functions and variance decompositions. Suppose that the first two columns of QG correspond to the $r \times 2$-dimensional impulse matrix associated with the first and the second principal component shocks, called Υ. Then the 2×1 vector of the impulse response functions of variable i to the shocks w_t at horizon k can be computed[51] as $\Theta_{ik} = \Lambda_i' A_1^k \Upsilon' R'$ and the corresponding forecast error variance as $\Theta_{i0}\Theta'_{i0} + \ \ldots \ + \Theta_{ik}\Theta'_{ik}$.

R has to be chosen such that the identifying restrictions specified in the main text are satisfied. Any two-dimensional rotation matrix R can be parameterized as follows

$$R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{A6}$$

To systematically explore the factor space, the rotation angle θ is varied on a grid from 0 to π. Further rotations would only result in repetitions, possibly with a flipped sign. The number of grids is set at 24, and θ is fixed such that the imposed restrictions are satisfied.

[50] For a detailed derivation see Uhlig (2003).
[51] Impulse responses are also multiplied by the variables' standard deviations.

As outlined in the main text, we identified an aggregate supply and an aggregate real demand shock. We also tried to identify a monetary policy shock, where, as in Peersman (2005), we restricted output and prices not to fall instantaneously and during the first four quarters after a positive monetary policy shock and short-term interest rates not to rise contemporaneously. By rotating the first two principal component shocks, we tried to identify such a monetary policy shock together with a supply shock and a monetary policy shock together with a demand shock. In addition, we rotated the first three principal component shocks and tried to simultaneously identify a supply shock, a demand shock and a monetary policy shock. But no rotation which satisfied the respective restrictions at the same time was found, even after increasing the number of grids. This suggests no monetary policy shock among the most important drivers of US economic real activity and is consistent with the view that monetary policy shocks contribute little to business cycles movements (cf. Uhlig, 2003).

Since $N \gg T$, the uncertainty involved with the factor estimation can be neglected (cf. Bernanke and Boivin, 2003). In order to account for the uncertainty involved in the estimation of the VAR model on the factors, we construct confidence bands by means of the bootstrap-after-bootstrap techniques based on Kilian (1998). These techniques allow us to remove a possible bias in the VAR coefficients which can arise due to the small sample size of the VAR model (for details on the bootstrap see Kilian, 1998). Most draws deliver not just one, but a set of shocks which all satisfy the restrictions. In this case, we follow Peersman (2005) and draw and save one of them. Some draws, however, do not deliver any shocks satisfying the restrictions. We draw until we have saved 500 shocks (536 draws were needed). For more details on the identification, the reader is referred to Peersman (2005).

IV. Comovements and Heterogeneity in the Euro Area Analyzed in a Non-Stationary Dynamic Factor Model[*]

1. Introduction

Although the member states of the European Monetary Union (EMU) are closely linked through trade and financial markets, economic comovements are still far from perfect, and there is still persistent heterogeneity across individual countries' output and price developments (Figure 4). Economic comovements (at business cycle and low frequencies) and heterogeneity were investigated and discussed intensively in the run-up to the EMU. This was reflected in the Maastricht criteria which stress common long-run tendencies (converged inflation and interest rates, a solid fiscal situation and stable exchange rates) and in other optimum currency area criteria[52], including a high degree of business cycle synchronization. These criteria are now widely accepted as being important prerequisites for a successful monetary union and are currently being re-applied to the central and eastern European EMU accession candidates. Comovements and heterogeneity have recently returned to the focus of interest among the public, academics and policymakers in light of persistent output growth differentials between the large euro-area economies since the mid-1990s and in light of an increase in inflation differentials observed since 2000 (cf. EBC, 2003). This renewed interest is reflected in a growing literature and numerous conferences on these issues organized by major European policy institutions.[53]

Heterogeneity in the euro area is not necessarily harmful and does not automatically call for policy intervention. Output and inflation differentials may partly reflect the catching-up process, in the course of which countries with lower initial incomes experience higher output

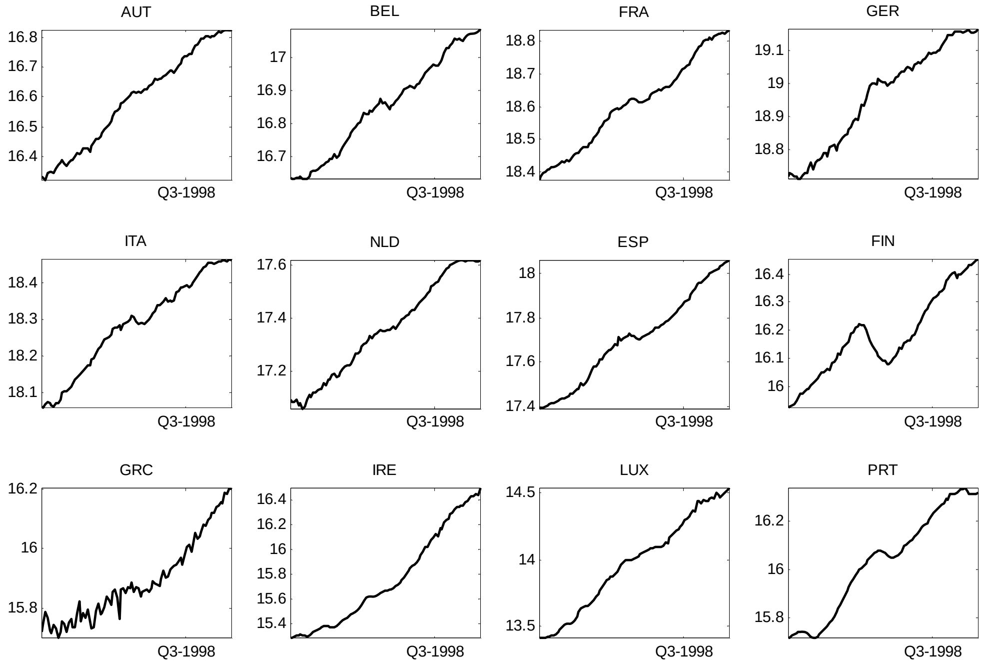
[*] This chapter is based on Eickmeier (2006), „Comovements and heterogeneity in the euro area analyzed in a non-stationary dynamic factor model", Bundesbank Discussion Paper 31/2006. This paper has been presented at seminars at the Deutsche Bundesbank, the Czech National Bank, the Joint ECB/DG ECFIN Workshop on "Dynamic Adjustment within EMU" in Frankfurt and the "Fourth Conference on Growth and Business Cycles in Theory and Practice", CGBCR/University of Manchester. It has benefited from comments by Jushan Bai, Jörg Breitung, Jörg Döpke, Uli Fritsche, Heinz Herrmann, Johannes Hoffmann, Carlos Lenz, Massimiliano Marcellino, Katerina Smídková, Jens Ulbrich, Jürgen Wolters, three anonymous referees and seminar and workshop participants.

[52] Cf. Mundell (1961), MacKinnon (1963).

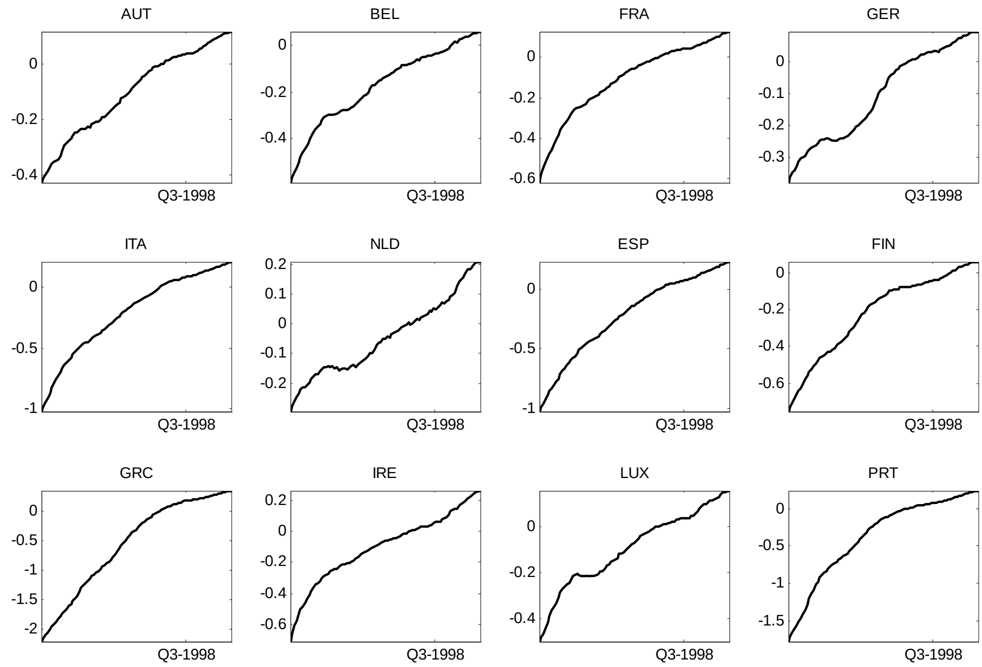
[53] Conferences and workshops hosted and/or organized by the European Central Bank include "What effects is EMU having on the euro area and its member countries?" (June 2005) and "Monetary policy implications of heterogeneity in a currency area" (December 2004); see http://www.ecb.int/events/conferences/past/html/index.en.html. Both the ECB and the European Commission organized a joint workshop on "Dynamic Adjustment within EMU" in March 2007. The European Commission, for example, hosted a conference on "Business Cycles and Growth in Europe" (October 2004). The themes of the first workshop of the Euro Area Business Cycle Network (EABCN) included "international business cycles".

Figure 4: Raw output and price data (1981Q2-2003Q4)

a) Output



b) Prices



Notes: The series shown here are raw data, i.e. they are seasonally adjusted, but not yet standardized nor outlier adjusted. Output is GDP in logs of 10 000 Euros, prices are logs of harmonized consumer price indices (1990Q1=100).

growth and inflation and dispersion is inevitable. In addition, adjustments in individual countries to shocks naturally trigger temporary inflation dispersion. If, however, such adjustments are slow due to nominal rigidities and imperfect factor mobility, this may lead to long-lasting undesirable output and inflation differentials. In addition, heterogeneity may also reflect inappropriate national economic policies or other unwarranted domestic developments, such as wage increases out of line with productivity and employment considerations, or excessive profit margin and demand developments caused, for example, by overconfident investors in asset markets. In these cases, if not counteracted by economic policies, heterogeneity may persist and result in large welfare losses for individual countries.

This chapter seeks to establish stylized facts about comovements and heterogeneity of output and price developments in EMU member states and their determinants by means of a large-scale dynamic factor model.

Factor models are particularly well suited to analyze economic comovements and heterogeneity among countries. They assume that variables such as output and prices are driven by a few common factors or shocks and by idiosyncratic shocks. Common shocks are shocks that hit all variables in the dataset simultaneously or with some lags, while idiosyncratic shocks are shocks that affect individual variables (or groups of variables) and do not affect all other variables at any point in time. In the international macro context, the former shocks may be either global shocks or shocks that occur in a country and are transmitted to others or correlated disturbances such as the implementation of similar national policies. Idiosyncratic shocks, by contrast, may include unexpected national economic policies and other country-specific economic developments such as those mentioned above. In the factor framework, idiosyncratic shocks lead to diverging economic developments, whereas common shocks are generally responsible for economic comovements. However, if the latter are spread asymmetrically to individual variables or countries — possibly due to differences in economic structures, such as nominal rigidities or factor mobility, economic policies or expectation formation processes — they might contribute to dispersion as well. Factor models therefore allow us to decompose heterogeneity in its components.

Large dynamic factor models also have various advantages over simple correlation analyses, VAR models or structural models which are more frequently used in the context of international economic linkages (cf. Bergman, 2004; Giannone and Reichlin, 2006; Croux et al., 2001; Ciccarelli and Rebucci, 2006; Clements et al., 2001). Much information can be extracted by means of dynamic factor models; this should allow us to estimate the common driving forces and their propagation more precisely. VAR modelers, by contrast, rapidly face scarce degrees of freedom problems. In addition, the potentially heterogeneous responses of many variables, i.e. all variables in the panel, to the common shocks can be assessed. It is also advantageous that we can remain relatively agnostic about the structure of the economy and do not need to rely on excessively tight restrictions, as is sometimes the case in structural models. The only restrictions we impose serve to give the common shocks an economic interpretation.

In this chapter, we combine the recently developed PANIC (Panel Analysis of Nonstationarity in Idiosyncratic and Common components) approach of Bai and Ng (2004, henceforth BN) and the structural factor setup based on Forni and Reichlin (1998, henceforth FR) and Forni, Giannone, Lippi and Reichlin (2005a, henceforth FGLR) and apply them to a newly constructed partly non-stationary dataset containing a total of 173 quarterly macroeconomic times series from 1981 to 2003, most of which capture economic developments in euro-area countries and some external influences. The former method allows us to estimate the common and idiosyncratic components of individual countries' output and prices, while the latter enables us to perform structural analysis, i.e. to identify common structural shocks and assess their propagation to these variables. We will also decompose heterogeneity: we will first assess to what extent heterogeneity is due to idiosyncratic shocks and adjustments to these shocks and to what extent it is due to the asymmetric spread of common shocks. We will then decompose the latter determinant further and investigate whether some common shocks trigger more heterogeneity than others.

This study is not the first that studies economic linkages in the euro area based on a large dynamic factor model. Marcellino et al. (2000) estimate factors from a large euro-area dataset and give them an economic meaning by correlating them with national factors. Altissimo et al. (2001) estimate a coincident indicator out of a large set of euro-area variables. Forni and Reichlin (2001) fit a large dynamic factor model to a panel of European regions' output, extract a common factor and assess its relevance. Beck et al. (2006) explain regional inflation in the euro area. Altissimo et al. (2004) fit a factor model to a large euro-area dataset which contains highly disaggregate inflation differentials and other macroeconomic variables. Sala (2003) studies the transmission of a common monetary policy shock to individual euro-area countries. Eickmeier and Breitung (2006) focus on the transmission of euro-area shocks to central and east European economies, but also provide results for current EMU members.

We go beyond the literature in mainly two respects. First, the studies just mentioned all fit stationary factor models to stationary datasets. The BN approach allows us to examine comovements and heterogeneity without imposing restrictions on the persistence of the variables and their components (which are allowed to be non-stationary) and hence also on comovements and heterogeneity. This is a particularly favorable feature: as we have pointed out above, it is not short-run but persistent heterogeneity that may indicate structural rigidities or inappropriate policies and that may be relevant for policy makers. Second, most of the analyses extract common factors from large datasets and examine how much of the variation they explain. Of the studies mentioned above, only Sala (2003) and Eickmeier and

Breitung (2006) are concerned about the economic interpretation of the factors.[54] They fit a structural dynamic factor model to the euro-area dataset and identify the structural shocks driving the common factors. While Sala (2003) focuses on a monetary policy shock, Eickmeier and Breitung (2006) identify aggregate euro-area supply, demand and monetary policy shocks. We will identify a richer set of shocks, namely five euro-area shocks.

The chapter is organized as follows. Section 2 describes the data. Section 3 presents the methodology. It first explains the non-stationary factor model of BN and then the structural dynamic factor setup suggested by FR and FGLR. Section 3 also identifies the common structural shocks. Section 4 presents the results. It first provides historical decompositions of individual countries' output and price developments into their components and computes their dispersion. It then examines the transmission of the specific common shocks to individual countries' output and prices. Section 5 concludes.

2. Data

We rely on a large dataset which is composed of 20 to 22 macroeconomic time series for each of the core euro-area countries (Austria, Belgium, France, Germany, Italy, the Netherlands and Spain), representing the real, the nominal and the external sides in a possibly balanced way. It further contains GDP and consumer prices from the other small and mostly peripheral euro-area countries (Finland, Ireland, Greece, Luxembourg and Portugal).[55] There is evidence that economic developments in the latter countries somewhat differ from developments in the core EMU countries (cf. Carvalho and Harvey, 2005). As shown by Boivin and Ng (2006), including too many variables with important idiosyncratic components in the dataset may distort factor estimates. We nevertheless aim at producing results for key variables of all and not only the core EMU countries. We therefore decided to include fewer, only two variables, of the smaller countries in the dataset. The dataset also comprises a few global variables which possibly have an impact on economic activity in the euro area, such as world energy prices, key US variables and the nominal US dollar/euro exchange rate. We further add five aggregate euro-area variables, which will help us to identify common structural euro-area macro shocks as will be apparent below: GDP, the harmonized consumer price index, the nominal short-term interest rate, real wages and investment. The aggregate euro-area series are taken from the dataset underlying the ECB's area-wide model (AWM), most remaining series are taken from OECD statistics. The dataset contains a total of $N = 173$ variables. The variables are listed in more detail in Table 7.

[54] Marcellino et al. (2000) do not identify the shocks behind the factors, but directly give the factors an economic meaning and therefore also do some structural analysis.
[55] Slovenia is not included in the analysis.

Table 7: Data

Country/region	Variable	Treatment	Source
Core EMU member countries (AUT, BEL, FRA, GER, ITA, NLD, ESP)	GDP, real	3	OECD
	Government expenditure	3	OECD
	Private final consumption expenditure	3	OECD
	Private total fixed capital formation, vol.	3	OECD
	Industrial production	3	OECD
	Capacity utilization rate manufacturing	0	OECD
	Total employment	3	OECD
	Unit labor costs (business sector)	3	OECD
	Productivity	3	OECD
	CPI, harmonized	3	OECD
	PPI	3	OECD
	GDP deflator	3	OECD
	Short-term interest rate nominal	0	OECD
	Long-term int. rate (gvt. bonds) nom.	0	OECD
	M1	3	Bundesbank
	M3	3	Bundesbank
	Main stock price index	3	OECD
	Industrial confidence	3	OECD
	Imports (goods and services), vol.	3	OECD
	Exports (goods and services), vol.	3	OECD
	Real effective exchange rate	3	IMF
	Current account balance	2	OECD
Remaining EMU countries (FIN, GRC, IRE, LUX, PRT)	GDP, real	3	OECD
	CPI, harmonized	3	OECD
World	Energy prices	1	HWWA
	Non-energy commodity prices	1	HWWA
	World trade	3	OECD
	Euro/US Dollar nominal	3	OECD
	US GDP, volume	3	OECD
	US CPI	3	OECD
	US nominal short-term interest rate	0	OECD
	UK GDP, volume	3	OECD
	JPN GDP, volume	3	OECD
Aggregate euro-area variables	GDP, real	3	ECB
	CPI, harmonized	3	ECB
	Short-term interest rate nominal	0	ECB
	Real wages	3	ECB
	Gross investment, real	3	ECB

Notes: This table describes which and how variables enter y_t. The dataset does not contain government expenditure, but government consumption expenditure for Germany and Spain. It does not contain private total fixed investment, but total fixed investment for Spain. Not PPI, but WPI for Austria is included. Productivity and industrial confidence are missing for Belgium, Italy and Spain and for Austria and Spain, respectively. Regarding the treatment of the data, 0/1/2/3 indicates that variables enter the dataset in levels/log levels/differences/log differences. If necessary, series were seasonally adjusted. Most German series are taken from the Bundesbank database. For further details on the data, see the text.

Data are quarterly, and our observation period ranges from 1981Q2 to 2003Q4; hence our time dimension T equals 91. One of the reasons for the choice of this period is data availability. Moreover, our reporting period is roughly the same as the period considered by Cavalho and Harvey (2005), termed stabilization and restructuring period by the authors, and it is long enough to comprise at least two entire business cycles according to the CEPR definition.[56]

Where necessary, the raw series were seasonally adjusted using the X-11 method and/or converted from monthly to quarterly series. Logarithms were taken of all non-negative series that were not already in ratios or percentage form. In constructing the dataset, one problem that needed to be addressed was the break in some series caused by German unification in 1990. Those German series for which a break was apparent were extended by applying West German growth rates to the German levels retroactively from the end of 1991 on. After this transformation, visual inspection of these series did not suggest a break anymore. The raw data are shown in Figure 4.

As pointed out in the introduction, our analysis mainly focuses on output and price dynamics, and the method of BN enables us to handle non-stationary data. We start, however, by constructing a stationary dataset and by fitting (in the methodological section) the more familiar stationary factor model to it. As usual in stationary factor analyses, we difference the variables until they are I(0). This involves taking a stance on the degree of integration of the individual variables. Most variables, such as output variables, are treated alike in the literature, but some variables are not. It is, for example, economically plausible to treat prices and interest rates as I(1) and I(0) variables, respectively. Empirical tests, however, sometimes suggest that they may be I(2) and I(1), respectively.[57] In the present study we treat them as I(1) and I(0) variables. Table 7 shows how each variable enters the dataset.

Factor analysis requires further manipulation of the dataset. All (now stationary) series are demeaned. Whenever a series exhibits structural breaks in the mean, we account for these breaks. [58] Notice that the demeaning procedure eliminates differences in mean growth rates of I(1) variables such as output and prices over the entire sample and, hence, also eliminates some of the heterogeneity. This, however, does not prevent heterogeneity to exist and persist for sustained periods of time, and it is still interesting to look at that dispersion.

---

[56] http://www.cepr.org/data/Dating/.

[57] There is no consensus in existing factor applications. Some studies treat prices as I(2) variables (cf. Stock and Watson, 2002; Forni et al., 2005a; Giannone et al., 2002), others as I(1) variables (cf. Marcellino et al., 2000; Cristadoro et al., 2005; Giannone et al., 2005). Interest rates are sometimes treated as I(1) variables (cf. Stock and Watson 2002; Forni et al., 2005a; Giannone et al., 2002, 2005; Cristadoro et al., 2005), sometimes as I(0) variables (cf. Marcellino et al., 2000).

[58] Breakpoints were detected by applying the sequential multiple breakpoint test of Bai and Perron (1998, 2003) (and the Gauss routines provided by Pierre Perron on his webpage) to all series of our (stationary) dataset, and we subtract the (possibly shifted) means from the series. Break dates are available upon request. Evidence for break in the means of inflation and related variables in the euro area can be found in Corvoisier and Mojon (2005) and Benati and Kapetanios (2003).

The stationary series are also normalized to have unit variances. This is done to account for the difference in measurement units in the dataset, which can influence factor estimates, and to guarantee that the variables with a relatively large variance do not dominate the factor estimates. Finally, outliers were removed.[59]

3. Methodology

Our analysis combines two dynamic factor setups suitable to analyze large datasets, the non-stationary (non-structural) factor model developed by BN and the structural factor setup of FR and FGLR. The former is presented in subsection 3.1. As mentioned in the previous section, it is convenient to start by presenting the more familiar stationary factor model, then explain, how it relates to the non-stationary factor setup of BN (PANIC) and how the latter can be used to estimate the possibly non-stationary common and idiosyncratic components of euro-area countries' output and prices. Subsection 3.2. then presents the structural factor setup.

3.1. The non-stationary factor model

The stationary series (first differences of I(1) variables and the I(0) level variables) are collected in the $N \times 1$ vector $\mathrm{y}_t = [y_{1t} \quad y_{2t} \quad \ldots \quad y_{Nt}]'$. It is assumed that y_{it}, $i = 1,....,N$, follows an approximate dynamic factor model (e.g. Stock and Watson, 1998, 2002; Bai and Ng, 2002) and can be represented as

$$y_{it} = x_{it} + \xi_{it} = \Lambda_i(L)'\mathrm{f}_t^d + \xi_{it} = \Lambda_i'\mathrm{f}_t + \xi_{it}, \tag{1}$$

where x_{it} and ξ_{it} are the scalar common and idiosyncratic components of variable y_{it}. The idiosyncratic components are allowed to be weakly cross-correlated in the sense of Bai and Ng (2002). $\mathrm{f}_t^d = [f_{1t}^d \quad \cdots \quad f_{qt}^d]'$ is a $q \times 1$ vector of common dynamic euro-area factors and $\Lambda_i(L) = \Lambda_{i0} + \Lambda_{i1}L + ... + \Lambda_{ig}L^g$ denotes the lag polynomial of $q \times 1$ vector of factor loadings associated with lags 0 to g. The loadings can differ across variables. $\mathrm{f}_t = [f_{1t} \quad \cdots \quad f_{rt}]'$ is a vector of $r \geq q$ 'static factors' that comprises the dynamic factors f_t^d and all lags of the factors that enter with at least one non-zero weight in the factor representation. The $r \times 1$ vector Λ_i comprises the loadings for the vector of static factors. Typically, $r \ll N$.

[59] Outliers are defined as data lying outside 6 times the interquartile range (cf. Watson, 2003). They are removed by being set to the latter.

So far, the analysis is confined to stationary series contained in y_t, such as output growth and inflation. The focus of this analysis, however, is on the levels of output and prices, and the BN procedure comes into play here. Cumulating equation (1) yields

$$Y_{it} = X_{it} + \Xi_{it} = \Lambda_i(L)'\mathrm{F}_t^d + \Xi_{it} = \Lambda_i'\mathrm{F}_t + \Xi_{it}, \tag{2}$$

where capital letter variables denote the cumulated lower case letter variables, i.e. $Z_t = \sum_{s=2}^{t} z_s$ and or, equivalently, $z_t = \Delta Z_t$ for any variables Z and z. Y_{it} may be the levels of output or prices (or other variables that were originally I(1)). F_t^d and F_t denote the $q \times 1$ and $r \times 1$ vectors of common dynamic and static factors affecting Y_{it}, and X_{it} and Ξ_{it} are its common and idiosyncratic components.

The elements of F_t may be stationary, non-stationary or both. Let r_0 denote the number of stationary and $r_1(= r - r_0)$ the number of non-stationary factors or common trends. In addition, the idiosyncratic components Ξ_{it} may be I(0) or I(1).[60] The source of non-stationarity in Y_{it} can thus be pervasive, idiosyncratic or both.

The goal is now to estimate equation (2). Following BN, we first estimate the stationary static factors $\mathrm{f}_t = \Delta\mathrm{F}_t$ by applying static principal component analysis to y_t: $\hat{\mathrm{f}}_t = \hat{\mathrm{V}}\mathrm{y}_t$, where $\hat{\mathrm{V}} = \begin{bmatrix} \hat{\mathrm{V}}_1 & \cdots & \hat{\mathrm{V}}_N \end{bmatrix}$ is the $r \times N$ matrix of eigenvectors corresponding to the largest r eigenvalues of the correlation matrix of y_t. $\hat{\mathrm{V}}_\mathrm{i}$ is an estimate of the vector of factor loadings Λ_i. The estimated common and idiosyncratic components of y_{it} are $\hat{x}_{it} = \hat{\mathrm{V}}_i'\hat{\mathrm{f}}_t$ and $\hat{\xi}_{it} = y_{it} - \hat{x}_{it}$. An estimate for F_t is then obtained through cumulation: $\hat{\mathrm{F}}_t = \sum_{s=2}^{t} \hat{\mathrm{f}}_s$, and common and idiosyncratic components of Y_{it}, X_{it} and Ξ_{it}, are estimated accordingly: $\hat{X}_{it} = \sum_{s=2}^{t} \hat{x}_{is}$ and $\hat{\Xi}_{it} = \sum_{s=2}^{t} \hat{\xi}_{is}$.

The reader may have noticed that our presentation of the model and of the estimation steps differs somewhat from the presentation in BN. The authors depart from a possibly non-stationary dataset, suggest differencing the entire dataset, extracting principal components from the differenced data, interpreting them as differenced factors and cumulating these differenced factors to recover the factors in levels, which drive the variables in the original dataset. Our original dataset contains I(1) and I(0) variables, and there is no need to difference the latter. Instead of differencing the entire dataset, we therefore differenced only the I(1) variables and extract principal components from a dataset which is composed of the differenced I(1) variables and the levels of I(0) variables.[61]

[60] BN state that, if Ξ_{it} is I(0), $\xi_{it} = \Delta\Xi_{it}$, although over-differenced, is still stationary and weakly correlated, and hence, the conditions for the consistent estimation of the number of factors and the factors themselves are not violated.

[61] Below, we will test the degree of integration of the common factors, and it turns out that results do not change when we apply these tests to cumulated factors estimated from a dataset of differenced I(1) and differenced I(0) variables.

Notice further that extracting F_t directly from the original dataset which contains all (I(1) and I(0)) variables in levels by applying principal component analysis to that dataset, is no option here. It would have been an option if idiosyncratic components were all I(0) (cf. Bai, 2004). However, this generally cannot be assured *a priori* (and it will turn out below that this is indeed not the case). If idiosyncratic components instead contain non-stationary elements, a regression of the series on the factors is spurious, even if the factors have been observed, and estimates for the loadings and thus of the idiosyncratic components are not consistent.

The dimension of F_t (and f_t), r, is set to be 5. Five factors combine to explain 37% of the total variance (of y_t).[62] Table 8 shows that also the variance shares of key euro-area aggregate variables explained by f_t are large, at 79% for GDP growth and at 43% for consumer inflation. Our choice is supported by the Bai and Ng (2002) PC_{p2} criterion[63] and by existing studies which find that four to six static factors explain between 37% and 55% of the total variance in euro-area macroeconomic datasets (cf. Marcellino et al., 2000; Altissimo et al., 2001; Eickmeier and Breitung, 2006; Altissimo et al., 2004).

Table 8: Variance of individual countries' and euro-area macro variables explained by the common factors f_t

	Output growth	Inflation
AUT	0.32	0.24
BEL	0.47	0.42
FRA	0.59	0.37
GER	0.44	0.64
ITA	0.41	0.25
NLD	0.23	0.32
ESP	0.21	0.18
FIN	0.41	0.55
GRC	0.07	0.12
IRE	0.26	0.30
LUX	0.17	0.48
PRT	0.40	0.07
Euro area	0.79	0.43

Notes: Output is GDP, inflation is consumer inflation.

The number of common stationary factors or trends r_1 are determined by applying the criteria proposed by BN. These criteria, denoted as MQ_c^τ and MQ_f^τ in BN, test whether the smallest eigenvalue of an autoregressive coefficient matrix is unity; see BN for a detailed description. The test statistics MQ_c^τ and MQ_f^τ take values -25.588 and -26.287, respectively, when we test

[62] The cumulated total variance shares explained by the first 10 factors are at 13%, 22%, 28%, 33%, 37%, 40%, 43%, 46%, 49% and 51%.

[63] The PC_{p2} criterion takes the values 0.8942 (1 factor), 0.8407 (2 factors), 0.8187 (3 factors), 0.8095 (4 factors), 0.8085 (5 factors), 0.8114 (6 factors), 0.8166 (7 factors), 0.8257 (8 factors), 0.8359 (9 factors), 0.8495 (10 factors). Details on the criterion are provided in Bai and Ng (2002).

the null hypothesis $r_1 = r$. They exceed the critical values reported in BN[64], which implies that we cannot reject the null, and hence r_1 is estimated to be five, i.e. all static factors are non-stationary and none is stationary.[65]

The BN framework also allows us to examine whether the source of non-stationarity of individual countries' output and prices is purely pervasive or also idiosyncratic, i.e. whether persistent shocks to these variables are only common or can also be country- (or series-) specific. This is of interest for policymakers, forecasters and other economic agents who perceive shocks in the economy and have – at least to some extent – an idea about the nature of these shocks, i.e. their structural interpretation, but also whether they are idiosyncratic or pervasive. They will only be able to adequately react to the specific shocks and/or make precise predictions, if they correctly anticipate how persistent the effects are that these shocks have on key economic variables.

We construct two panels, one containing the idiosyncratic components of individual countries' GDPs and one of individual countries' consumer prices. We apply panel rather than univariate unit root tests because the latter are known to have low power (cf. Breitung and Pesaran, 2006). Of the many existing panel unit root tests, we decide to apply the panel Modified Sargan-Bhargava (henceforth PMSB) test recently suggested by Bai and Ng (2007a) and the tests developed by Harvey and Bates (2003, henceforth HB) and Breitung and Das (2005, henceforth BD) to our two panels. The PMSB test is a natural choice, since it has been designed for the PANIC framework and has been shown to perform relatively well in contexts such as ours where variables originally exhibited a trend. While the PMSB test assumes independent idiosyncratic errors, the HB and BD tests are robust with respect to cross-section dependence of the units. This is appropriate for our panels of idiosyncratic components, since those are allowed to be weakly cross-correlated in the sense of Bai and Ng (2002, 2004) in approximate factor models. Since the PMSB test was originally designed for large N, we simulated the critical values for our smaller panels. For details on the PMSB test, we refer to Bai and Ng (2007a). The HB and the BD tests are described in Appendix B. Again, we needed to simulate the critical values, something which is also outlined in that appendix.

The tests indicate that idiosyncratic components of output are all I(1) (see the row referring to 'total dataset' in Table 9). This suggests that output in individual euro-area countries is not only driven by permanent common shocks, but also by permanent idiosyncratic shocks. This is consistent with our impression from visual inspection of the upper panel of Figure 5. Results are different for prices: the HB and the BD tests reject the null that all idiosyncratic

---

[64] The critical values are reported in BN, Table I, p. 1136. To compute the MQ_c^τ criterion, we set $J = 4ceil[\min(N,T)/100]^{1/4}$. The VAR order for the computation of the MQ_f^τ criterion is determined with the Akaike criterion.

[65] Our results differ somewhat from the findings by Luginbuhl and Koopman (2004), who find three common trends between 1970 and 2001 and two between 1987 and 2001; this is possibly due to the fact that they have included the output of five core euro-area countries in their dataset, whereas we work with a much larger dataset.

components are I(1) at the 5% significance level, and the PMSB test rejects it at the 10% significance level. Augmented Dickey-Fuller tests applied to idiosyncratic components of prices individually only reject the null of a unit root for Spain.

Table 9: Panel unit root tests applied to idiosyncratic components of individual countries' output and prices

	PMSB	HB	BD	PMSB	HB	BD
		Output			Prices	
Total dataset	-0.411	-3.954	-3.046	-1.423 *	-5.604 **	-5.027 **
1985Q1-2003Q4	0.808	-3.440	-2.608	-1.139	-4.595	-3.989
1990Q1-2003Q4	-1.718 ***	-3.072	-2.922	-1.716 ***	-1.364	-2.239
1995Q1-2003Q4	-1.690 ***	-1.316	-2.509	-0.558	-0.696	-1.834
E7, r = 5	0.156	-3.261	-2.890	-1.024	-4.803 ***	-3.789 *
E7, r = 3	0.120	-2.552	-2.411	1.168	-3.511	-2.444
E4, r = 5	-0.039	-2.165	-1.932	-0.858	-3.523 *	-3.425 *
E4, r = 3	-0.052	-1.764	-1.964	0.605	-2.584	-2.104
Only real variables, r = 5	2.695	-3.045	-2.351	-	-	-
Only real variables, r = 3	2.957	-3.215	-2.730	-	-	-
Only nominal variables, r = 5	-	-	-	-1.354 *	-5.978 ***	-6.446 ***
Only nominal variables, r = 3	-	-	-	-1.447 *	-6.127 ***	-5.942 ***

Notes: 'Total dataset' refers to our benchmark, i.e. the period 1981Q2-2003Q4 and r = 5. Results for the different periods are obtained with r = 5. E7 includes the core euro-area countries as defined in Table 7, E4 includes the largest countries (France, Germany, Italy and Spain). '***'/'**'/'*' indicates rejection of the null hypothesis of all idiosyncratic components being I(1) at the 1%/5%/10% significance level. HB and BD refer to t_{gls} and t_{rob} in Appendix B.

Three remarks are in order. First, the finding that most idiosyncratic components of output and prices are non-stationary implies that some individual countries' output and prices may have diverged during the sample period. Whether this is worrying against the background of a common monetary policy or not crucially depends on what drives our results. To shed some light on the sources of the idiosyncratic components' non-stationarity, we investigate to what extent our results depend on the period and on the composition of our dataset.

Regarding the period, current euro-area member countries departed from very different economic situations at the beginning of the 1980s and went (and are still going) through phases of convergence with different adjustment processes. With more and more countries having completed the convergence process, we would expect, *ceteris paribus*, the non-stationarity of idiosyncratic components to vanish over time. We fit the factor model to three alternative periods, starting in 1985Q1, 1990Q1 and 1995Q1 and ending all in 2003Q4. There is some – although weak – evidence that different economic developments at the beginning of the sample period indeed may have played some role: the PMSB rejects the null of all idiosyncratic components being I(1) for output in the two latest periods (Table 9). This is, however, not supported by the other two tests which find unit roots in all idiosyncratic components of output and also prices in all periods. One reason for the finding of the PMSB test may be that adjustments were still ongoing in many countries in later parts of the sample period. Another reason may be that the ongoing convergence process was counteracted by a few relatively large country-specific shocks (unrelated to the convergence process) and subsequent protracted and painful adjustment processes that occurred in the 1990s (such as the German unification in 1991 and the Finish banking crisis at the beginning of the 1990s).

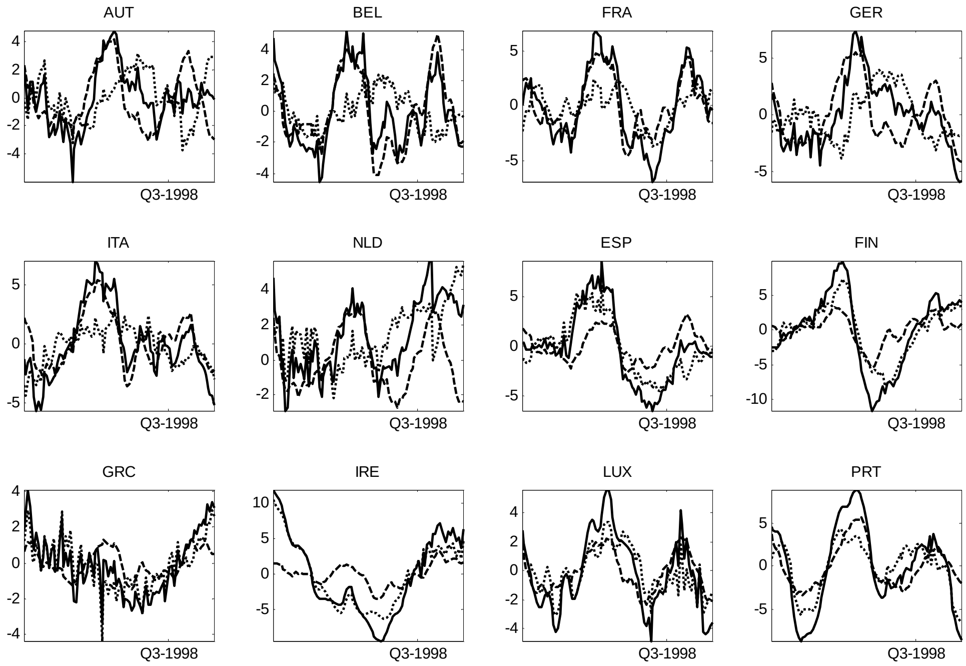
Regarding the composition of our dataset, it is useful to remember that our framework enables us to estimate factors which are common to all variables in the dataset, but not factors which only affect subgroups of variables. The latter will, in our framework, be reflected in the variables' idiosyncratic rather than common parts. This is certainly a drawback of our methodology given previous findings in the literature that certain countries in the euro area form 'clusters' (cf. Busetti et al., 2006) or 'convergence clubs' (cf, Canova, 2004). Moreover, there may be permanent factors driving output but not prices and *vice versa* which would be reflected in non-stationary idiosyncratic rather than common components in our framework. The Bayesian factor model proposed by Kose et al. (2003a) (and developed further by Del Negro and Otrok, 2005) can deal with large datasets and, at the same time, account for factors which are common to subgroups of variables only. However, this does not come without a cost. The Bayesian factor model imposes more structure with respect to the presence of certain factors and the variables they load or do not load and is more complicated to estimate. Moreover, it has only been applied to stationary datasets so far. To investigate whether the non-stationarity of most idiosyncratic components can be explained by this particular characteristic of our model, we estimate factors from a set of only real (nominal) variables[66] and test the degree of integration of the idiosyncratic components of output (prices). We also extract factors from sets of variables associated only to the seven core euro-area countries (E7) and to the four largest euro-area countries – France, Germany, Italy and Spain – (E4). To allow for the possibility that smaller datasets are driven by fewer factors, we consider, besides five factors, also three factors. It turns out that results which were obtained based on the total dataset are roughly reproduced (Table 9).

Overall our results are therefore not simply driven by the characteristics of our particular modeling framework. By contrast, we cannot exclude that large idiosyncratic shocks (unrelated to the convergence process) and slow adjustments to them are responsible for non-stationary idiosyncratic components. There is also some (however very weak) evidence that economic developments differed at the beginning of the sample, but became more similar

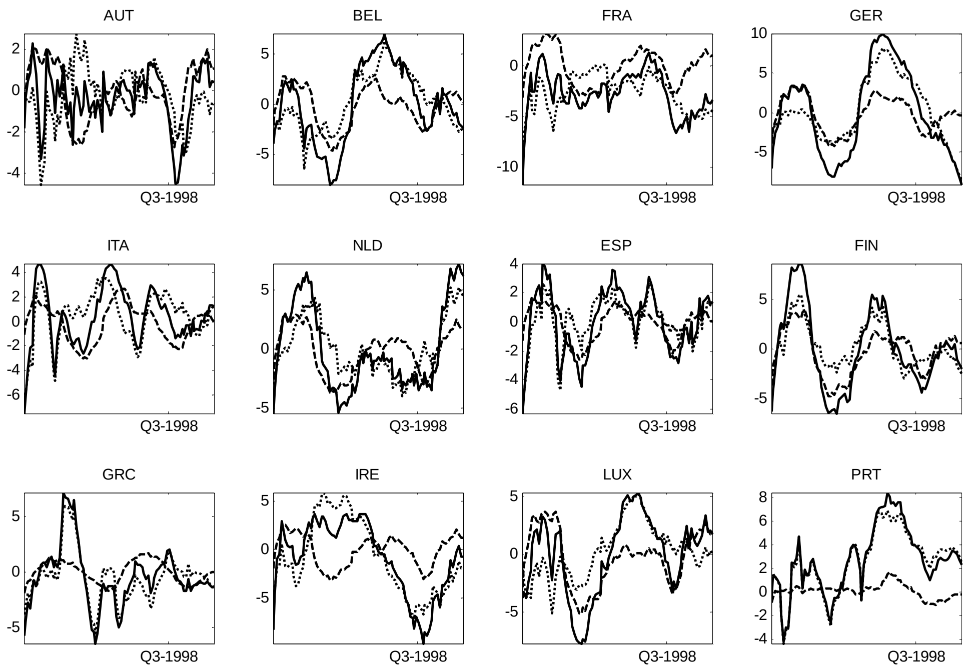
[66] Real effective exchange rates and current account balances are not included in either datasets.

Figure 5: Macroeconomic series and common and idiosyncratic components (1981Q2-2003Q4)

a) Output



b) Prices



Notes: Series (solid), common components (dashed), idiosyncratic components (dotted).

towards the end in terms of vanishing non-stationarity of idiosyncratic components. Interestingly, the results given by the PMSB test sometimes differ from the results given by the other two tests, suggesting that it may matter whether cross-correlation of idiosyncratic components is accounted for in the unit root tests or not.

Second, our modeling framework assumes constant factor loadings. One might ask whether our model is also still applicable if the European integration process has changed comovements within the euro area. The underlying hypothesis is that economic comovements are, like other optimum currency area criteria, endogenous, i.e. a monetary union should enhance trade and the integration of financial markets, which should tighten economic linkages between member states (cf. Frankel and Rose, 1998).[67] Based on bivariate VAR models and structural break tests, Eickmeier and Breitung (2006) find no evidence that EMU has altered linkages between changes in output and inflation of individual euro-area countries and the corresponding euro-area aggregates which supports our constant parameter approach. Similarly, according to Canova et al. (2007b) who use a time-varying panel VAR approach, the Maastricht Treaty and the inception of the European Central Bank do not seem to represent clear structural breaks.[68] Disregarding this evidence, our model is robust in the sense that the principal component estimator remains consistent with respect to mild time variation in the factor loadings as long as $T/N \to 0$, as shown by Stock and Watson (1998).

Third, the individual factors are not interpretable as such, since they are identified only up to a rotation. Recently, researchers have attempted to give them an economic meaning. Some focus on the factors themselves: see, for example, Marcellino et al. (2000) who investigate the relationship between the set of factors and individual variables or groups of variables or other factors, using multivariate correlation measures; Bai and Ng (2006) have developed formal tests to assess such relationships. Others give the factors an economic meaning by identifying the shocks, which drive the factors. This is the way we go here and describe in the next subsection.

3.2. The structural factor setup

By construction, the common factors are driven by q shocks that result from the VAR(p) representation of the factors:

$$A(L)f_t^d = Qv_t, \tag{3}$$

[67] Although theoretically not clear (cf. Kose et al., 2003d), trade and financial market linkages have been shown in empirical studies to enhance business cycle synchronization (cf. Imbs, 2004; Kose et al., 2003a; Baxter and Kouparitsas, 2005).

[68] Some changes in the transmission of German and external shocks are, nevertheless, found by the authors after both events and the creation of the ECB, respectively.

with $\mathrm{A}(L) = \mathrm{I} - \mathrm{A}_1 L - ... - \mathrm{A}_p L^p$. Matrix Q is chosen such that the innovations v_t are orthonormal. The shocks w_t are related to v_t through the structural equation

$$\mathrm{w}_t = \mathrm{R}\mathrm{v}_t, \tag{4}$$

where $\mathrm{R}'\mathrm{R} = \mathrm{I}_q$. Provided that there are enough identifying restrictions on R, the structural shocks w_t can be recovered from the factor innovations. The $N \times q$ matrix of impulse responses to the shocks $\mathrm{w}_t = \begin{bmatrix} \mathrm{w}_{1t} & \cdots & \mathrm{w}_{qt} \end{bmatrix}'$ at horizon h, $\partial \mathrm{y}_{t+h} / \partial \mathrm{w}_t' = \Theta_h$, is obtained from

$$\Theta(L) = \Theta_0 + \Theta_1 L + \Theta_2 L^2 + ... = \Lambda(L)' \mathrm{A}(L)^{-1} \mathrm{Q}\mathrm{R}'. \tag{5}$$

Cumulating the impulse response matrix given in (5) yields impulse responses of Y_{it}. One important objective of the analysis is to identify w_t and to assess impulse responses of individual variables to these shocks.

To estimate the VAR innovations v_t, we fit a VAR(2) model to $\hat{\mathrm{f}}_t = \Delta \hat{\mathrm{F}}_t$, thereby exploiting our previous finding that all factors in $\hat{\mathrm{F}}_t$ were estimated to be non-stationary. The lag order of the VAR model was estimated with the Akaike information criterion.

It is important to note that the VAR representation for $\hat{\mathrm{f}}_t$ is singular if the r-dimensional vector $\hat{\mathrm{f}}_t$ (and $\hat{\mathrm{F}}_t$) is driven by $q < r$ shocks. To estimate the q-dimensional vector v_t from the r-dimensional vector of residuals of the fitted VAR based on $\hat{\mathrm{f}}_t$, a principal component analysis is employed. This yields the linear combination of the q non-zero components in the residual vector of the VAR model. Let $\hat{\mathrm{v}}_t$ denote the resulting vector of orthogonal factor innovations. In our case, however, there is no need to employ any of the criteria proposed in the literature to estimate q formally (cf. Breitung and Kretschmer, 2005; Bai and Ng, 2007b; Amengual and Watson, 2006) or informally (cf. Forni et al., 2000). Instead, we set q equal to 5, which is consistent with our estimate of $\hat{r}_1 = \hat{r}$. Let us explain this reasoning with an example. Suppose that the vector of static factors F_t comprises four dynamic factors and one lagged dynamic factor, i.e. $\mathrm{F}_t = \begin{bmatrix} \mathrm{F}^d_{1t} & \mathrm{F}^d_{1t-1} & \mathrm{F}^d_{2t} & \mathrm{F}^d_{3t} & \mathrm{F}^d_{4t} \end{bmatrix}'$. Suppose further that all elements of F_t, i.e. all static factors, are I(1). Then, there exists a linear combination of F_t, which is stationary, namely $\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \end{bmatrix} \times \mathrm{F}_t$, which would have been detected by the BN tests as a stationary factor, i.e. $\hat{r}_1$ would have been 4. However, this was not the case, which suggests that all static factors are also dynamic factors. It thus follows from our estimate of $\hat{r}_1 = \hat{r}$ that $\hat{q} = \hat{r} = 5$.

The common structural shocks can now be recovered as in the structural VAR literature. The matrix R is chosen such that certain identifying restrictions that need to be specified are satisfied. This is achieved by applying the identification scheme initially proposed by Uhlig (2005) and Faust (1998) for monetary policy shocks and extended by Peersman (2005), Canova and de Nicoló (2003) and Peersman and Straub (2006, henceforth PS) to a larger set of shocks which consists in imposing sign restrictions on short-run impulse responses. This prevents us from using contemporaneous restrictions which are generally hard to justify theoretically or long-run restrictions commonly employed in the structural VAR (and structural dynamic factor) literature which are at odds with some theoretical models.

It is beyond the scope of this chapter to explicitly write down a theoretical model and to derive sign restrictions from this model. Instead, we borrow sign restrictions from PS who derive them from a New Keynesian dynamic stochastic general equilibrium model which includes standard frictions such as nominal rigidities, habit formation in consumption, investment adjustment costs and variable capital adjustment costs (see, for example, Christiano et al., 2005 or Smets and Wouters, 2003).[69]

Common euro-area shocks are here defined as shocks that explain a notable part of fluctuations in key aggregate euro-area variables. We identify two euro-area supply shocks, two euro-area real demand shocks and one monetary policy shock using the restrictions of PS.[70] Notice that PS identify seven shocks: three supply shocks (technology, labor supply and price markup), three demand shocks (preference, government spending, investment) and a monetary policy shock. However, our dataset is driven by only five shocks. We identify those shocks that were shown to individually explain the bulk of output and price fluctuations in the euro area in PS and Smets and Wouters (2003) and summarize some of the shocks identified in the former analysis, as explained below. Notice further that the shocks are all consistent with the model outlined in PS and hence no external shocks are separately identified, but will be partly reflected in the euro-area shocks.

We impose the following restrictions on impulse responses of aggregate euro-area real GDP, consumer prices, short-term nominal interest rates, real wages, and real investment, which are also summarized in Table 10. All common shocks are normalized to have a positive effect on euro-area output. The two euro-area supply shocks move output and prices in opposite directions. The two demand shocks and the monetary policy shock raise output and prices. While the demand shocks also lead to an increase in short-term interest rates, the latter decline after the monetary policy shock. We distinguish the two supply shocks as follows: the first supply shock, labeled ‘supply shock 1’, leads to an increase in real wages (which is consistent with the technology shock and the price mark-up shock in PS) and the second supply shock, labeled ‘supply shock 2’, lowers real wages (which is consistent with a labor supply shock in PS). The two real demand shocks are identified based on the following restrictions: ‘demand

[69] The authors show that the sign restrictions are consistent with a range of reasonable parameter values. The theoretical model provides more restrictions than necessary for identification. PS relax those restrictions which are not necessary to uniquely identify the shocks, and they also relax those restrictions that are controversial in the literature. We will rely on their relaxed restrictions.

[70] It is not unusual to identify euro-area monetary policy shocks even before the ECB superseded the national central banks as monetary authorities in 1999. Peersman and Smets (2002), for example, also identified common monetary policy shocks using synthetic euro-area data.

Table 10: Sign restrictions

	Output	Prices	Interest r.	Wages	Investment
Supply shock 1	≥	≤	?	≥	?
Supply shock 2	≥	≤	?	≤	?
Demand shock 1	≥	≥	≥	?	≤
Demand shock 2	≥	≥	≥	?	≥
Monetary policy shock	≥	≥	≤	?	?

Notes: Output is real GDP, prices are consumer prices, interest rates are short-term interest rates, wages are real, investment is real gross investment. Restrictions are imposed contemporaneously and on the first 4 quarters after the shocks. '?' indicates that no restrictions is imposed. For details on the underlying model, see Peersman and Straub (2006).

shock 1' raises investment (and is consistent with an investment shock in PS), whereas investment declines after 'demand shock 2' (being consistent with a preference shock and a government spending shock in PS).

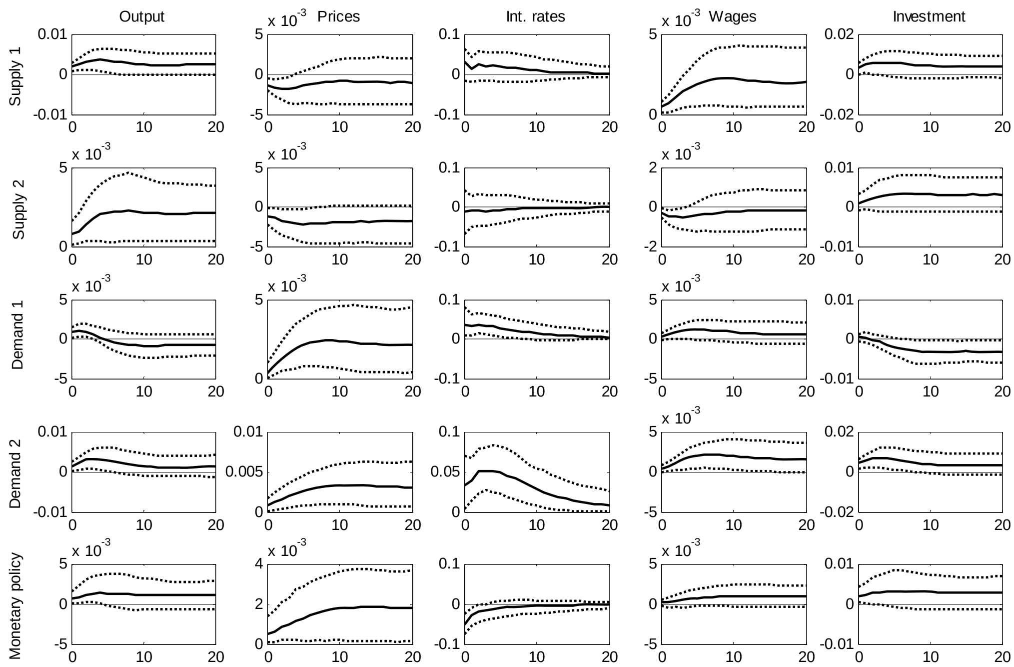
Restrictions are imposed on the contemporaneous impulse responses and on the first four lags. Results are robust with respect to the number of restricted lags. They remain unaffected when we only identify one of these shocks at a time and do not care about the other shocks, as done in PS.[71] We report the median impulse responses and 90% confidence bands which were constructed using bootstrap techniques. More details on the theoretical underpinnings can be found in PS. For details on the structural analysis and the bootstrap, see Appendix A.

In the following, we briefly characterize the main sources of economic fluctuations in the euro area before assessing their transmission to individual EMU economies. Impulse responses of euro-area output, prices, short-term interest rates, real wages and investment — the variables which were used to identify the common shocks — are shown in Figure 6. They appear roughly consistent with those found in the literature. The effects of the supply shocks on output tend to be quite persistent. Long-run neutrality with respect to output of the monetary policy shock cannot be rejected. The demand shocks trigger a temporary hump-shaped response of euro-area output. All shocks but supply shock 1 lead to long-lasting price responses.

Most of the variance of the forecast error of the common component of euro-area output is explained by supply shock 1 and demand shock 2: variance shares are at 41% and 29%, respectively, at short horizons (up to one year) and at 34% and 17%, respectively at medium horizons (up to five years); supply shock 2 also notably contributes to output fluctuations at medium horizons (Table 11). The contribution of the other shocks to output fluctuations are modest. Prices are mainly determined by the two demand shocks. The supply shocks play a greater role at short than at medium horizons, and the monetary policy shock explains 10% of the variance of the forecast errors of the common component of euro-area prices at medium horizons.

[71] We are grateful to an anonymous referee whose comment led us to conduct this exercise.

Figure 6: Impulse responses of key macro variables to structural shocks



Notes: Median (solid), 90% confidence bands (dotted).

4. Results

This section illustrates the evolution of common and idiosyncratic components and the transmission of specific common shocks to individual countries' output and prices. Heterogeneity is first decomposed into heterogeneity due to idiosyncratic shocks and heterogeneity due to the asymmetric spread of common shocks (subsection 4.1.). The latter heterogeneity is then further decomposed into heterogeneity due to the asymmetric spread of specific common shocks (subsection 4.2.).

4.1. Common and idiosyncratic individual countries' output and price components

The historical decomposition of individual countries' output and price developments (i.e. their deviations from deterministic trends) into common and idiosyncratic components is shown in Figure 5.[72] From the upper panel, it is apparent that GDPs of France, Belgium, but also Italy

[72] The BN procedure of differencing, de-meaning the differenced data and cumulating them yields series (and factors) which have endpoints of 0 and are therefore not reliable at the beginning and the end of the sample. This is no problem for the tests. However, when we are interested in the series and components themselves (which is the case in this subsection), we apply an OLS detrending to variables which were originally I(1) prior to estimating the factors, instead of using the BN procedure. This delivers series and factors which do not have the

Table 11: Forecast error variance decomposition of key euro-area macro variables

	Sup 1	Sup 2	Dem 1	Dem 2	Mon pol
Forecast horizon of 0 to 1 year					
Output	0.41	0.09	0.03	0.29	0.05
Prices	0.19	0.20	0.12	0.20	0.05
Interest rates	0.10	0.06	0.20	0.33	0.14
Wages	0.29	0.04	0.16	0.32	0.04
Investment	0.31	0.05	0.01	0.44	0.08
Forecast horizon of 0 to 5 years					
Output	0.34	0.19	0.03	0.17	0.06
Prices	0.08	0.13	0.16	0.33	0.10
Interest rates	0.11	0.06	0.17	0.39	0.09
Wages	0.40	0.02	0.08	0.33	0.08
Investment	0.24	0.10	0.08	0.26	0.09

Notes: The median is shown. Output is real GDP, prices are consumer prices, interest rates are short-term interest rates, wages are real, investment is real gross investment.

are closely related to the euro-area average, suggested by the fact that common components and series move in close parallel. Important idiosyncratic output movements are found for Greece and Ireland, which may be explained by catching-up processes. Finish and German output also diverged temporarily from the euro-area average. In Finland, the banking crisis at the beginning of the 1990s represented a strong negative idiosyncratic shock; in addition, Finland had tied trade links with the USSR and was affected more than other euro-area economies by the collapse of the Soviet economic system at the end of the 1980s. Germany first experienced a post-unification boom after 1991, which was largely idiosyncratic. Interestingly, Germany's weak economic performance in the second half of the 1990s is due primarily to idiosyncratic influences: unlike the common component, the idiosyncratic component of German output almost continuously tends to fall. As concerns historical decompositions for price developments, Spanish, French, Belgian and Finish price developments are to some extent driven by the common factors, whereas prices in the Portugal, Greece and Ireland are dominated by idiosyncratic factors representing the other extreme.

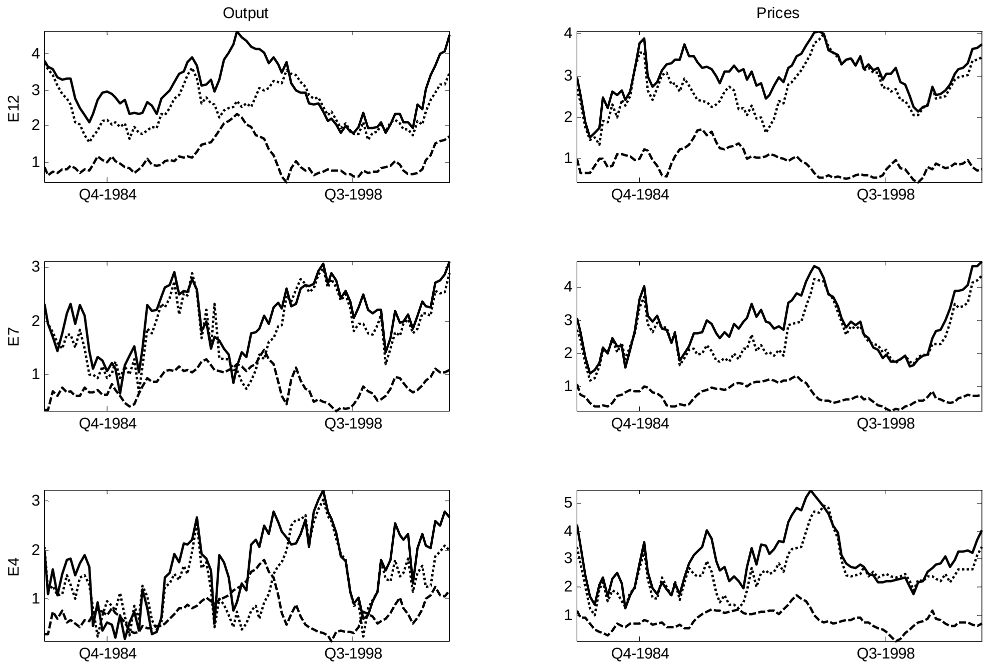
These findings based on visual inspection are confirmed by Table 8 which reports variance shares of output growth and inflation explained by their components. Common factors (f_t) are important for output growth in France, Belgium and Germany with variance shares at 59% , 47% and 44%. By contrast, they are relatively unimportant in Greece and Luxembourg with shares at 7% and 17%, respectively. Common factors play a relatively important role for

described undesired property. Notice further that the break dates which were detected (as described in section 2) for the means of output growth and inflation were also taken as break dates for the linear trends of the levels series.

inflation in Germany, Finland, Luxembourg and Belgium (between 42% and 64%), whereas they are rather unimportant for in inflation in the Portugal (7%) and Greece (12%).

The two upper panels of Figure 7 show standard deviations of the series of output and price developments (again, as deviations from deterministic trends) and of common and idiosyncratic components across all countries (E12).

Figure 7: Dispersion of output and price developments and their components across EMU countries (1981Q2-2003Q4)



Notes: The panels shows (unweighted) standard deviations of individual countries' GDP and consumer price developments and their components. Series (solid), common component (dashed), idiosyncratic component (dotted). E12/7/4 refers to the groups of 12 euro-area countries, the 7 core euro-area countries and the 4 largest euro-area countries.

Dispersion of both output and price series look quite persistent and somewhat unstable. Most of output and price dispersion seems to be due to idiosyncratic shocks. In the entire sample, the standard deviations of the idiosyncratic components of output and prices exceed the standard deviations of the common components. Idiosyncratic shocks were responsible for the relatively large price dispersion in the second half of the 1980s/beginning of the 1990s, when Germany and also the Netherlands experienced very low (and even negative inflation rates) – see the lower panel of Figure 4; the subsequent decline in dispersion marks the end of this phase. Idiosyncratic shocks also have contributed to the rise in output dispersion at the end of the 1980s/beginning of the 1990s, when Germany experienced a unification boom and then entered the recession later, in 1993, than its European neighbors. The subsequent decline in output and also in price dispersion may reflect adjustments in individual countries in the run-up to EMU. Interestingly, dispersion rises after Stage Three of EMU. Both common and idiosyncratic components seem to contribute to the increase in output dispersion, whereas the increase in price dispersion is mainly idiosyncratic. The four lower panels of Figure 7 show that output dispersion in the euro area until the end of the 1980s is mainly due to dispersion across the smaller euro-area countries, while this does not hold anymore for the 1990s when dispersion across the four large euro economies is as large as dispersion across all euro-area countries. By contrast, no clear difference is apparent between price dispersion associated with the different country groups.

Our findings are roughly in line with Giannone and Reichlin (2006) and Buisán and Restoy (2005). Based on bivariate VAR models and counterfactual correlations, the former paper also finds that idiosyncratic shocks explain the bulk of output (growth) dispersion. This result is confirmed by Buisán and Restoy (2005). Based on a large macroeconometric model (NIGEM) they estimate that individual countries' output and prices respond only moderately heterogeneously to common shocks. Our results regarding price dispersion, however, differ from those of Altissimo et al. (2004) who attribute most of inflation differentials to the common component. There exist several differences between their and our approaches. The authors focus on a different period (1990 to 2004) and on year-on-year inflation rates, whereas we have included prices (as deviations from their trends) in our analysis; in their analysis, inflation dispersion is measured as the unweighted average of differences between individual inflation rates and weighted euro-area inflation; their underlying dataset is different, including, among others, 60 time series of disaggregated inflation dispersion. Our results match those of Altissimo et al. (2004) insofar, as they also attribute an important role to idiosyncratic shocks for the increase in inflation differentials in 2000.

4.2. Individual countries' output and price impulse responses to specific common shocks

In the following, we leave apart idiosyncratic components and focus only on the common components. Figure 8 illustrates impulse responses of output and prices of individual euro-area countries to euro-area macro shocks and Table 12 the corresponding forecast error variance decompositions. Overall, impulse responses of output and prices look quite similar across countries and similar to the euro-area aggregates. But there are a few exceptions. Regarding output responses, the first euro-area supply shock leads to similar impulse response functions in the EMU countries, which are, however, not significant in Spain, Finland, Greece and Ireland. Its explanatory power for output in Portugal and Germany is largest. In these two countries, output does, however, not react significantly to supply shock 2. At least marginally significant responses to that shock, by contrast, are found for Belgium, France, Italy and Finland. Table 12 suggests that the variance shares explained by supply shock 2 are relatively large for these countries. Positive responses of output to the euro-area demand shocks are immediate and short-lived in most countries. However, output declines in the Netherlands and Luxembourg after the first demand shock. The demand shocks do not trigger significant responses in Finland; in addition, the first demand shock leads to insignificant responses in Austria and Portugal. Variance decompositions suggest that the first demand shock is most important for output fluctuations in Luxembourg, the Netherlands and Greece and that the second demand shock is most important in Austria and Germany. The euro-area monetary policy shock triggers temporary positive output reactions in most countries, which are, however, insignificant in Italy, the Netherlands, Finland, Luxembourg and Portugal. The forecast error variance of the common component of output accounted for by this shock is quite low in most countries except for Ireland, France, Spain and Greece where it explains 22% to 43% of fluctuations in the common component of output at medium horizons.

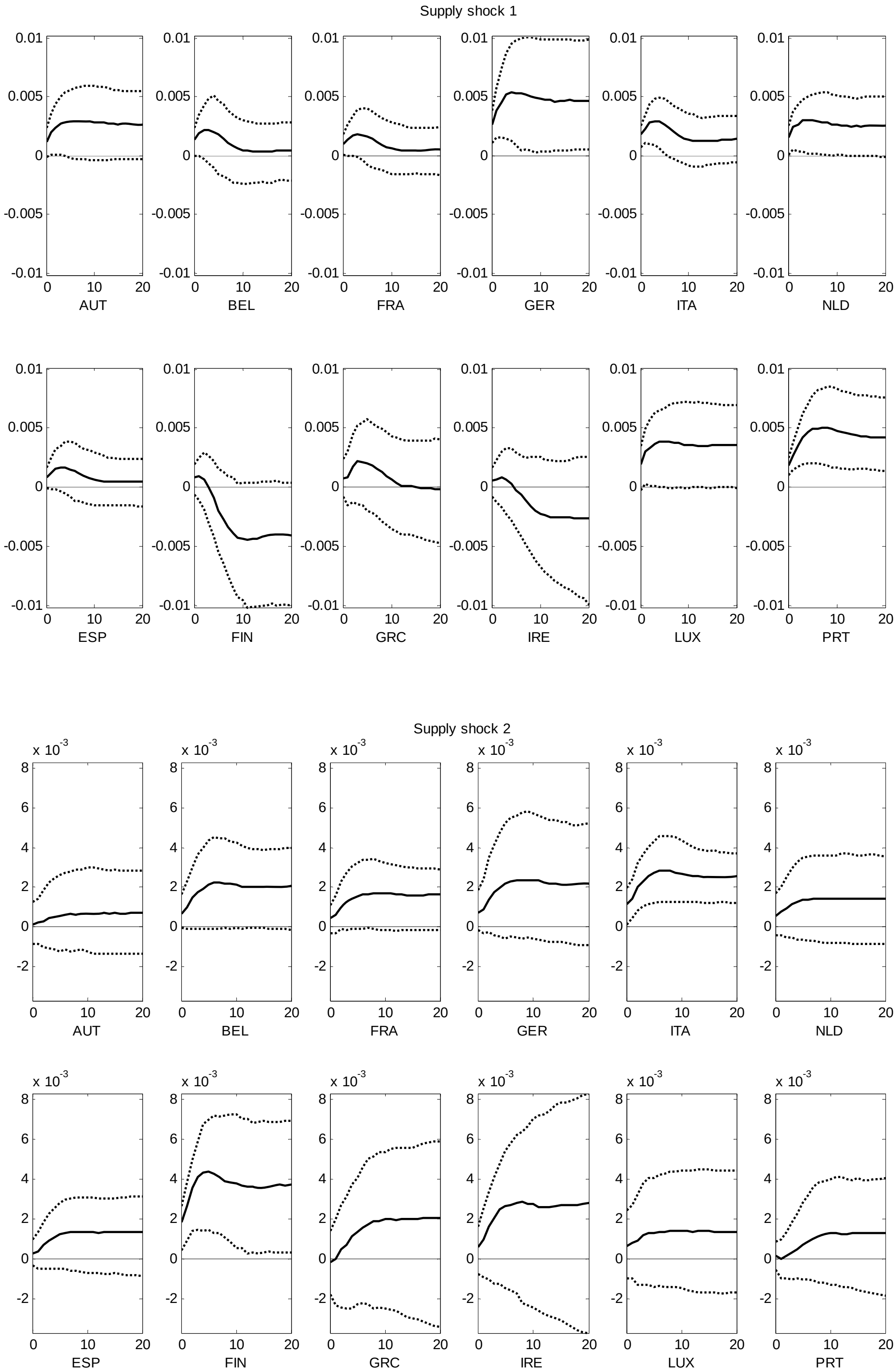
Price responses to the common shocks are also quite similar. However, Greek and Portuguese prices react somewhat atypically to the monetary policy and the second demand shock: they decline, although not significantly, while all other price responses are positive (or insignificant).

To assess the heterogeneity of the propagation of individual common shocks more formally, we compute cross-country standard deviations of impulse responses which are shown in Figure 9. The graphs do not clearly suggest that some shocks lead to impulse responses which differ more across countries than other shocks. Supply shock 2, demand shock 1 and the monetary policy shock seem to trigger somewhat less disperse output responses than supply shock 1 and demand shock 2; however, confidence bands associated to the latter shocks are also wider. Figure 9 also suggests that dispersion of impulse response functions of output to common shocks are mainly due to the smaller, more peripheral countries: dispersion across all 12 member states exceeds dispersion across the seven core or the largest four countries. Dispersion of price impulse responses does not differ notably across country groups and across shocks. Only price dispersion in response to the monetary policy shock tends to be larger across the group of all EMU countries compared to smaller country groups.

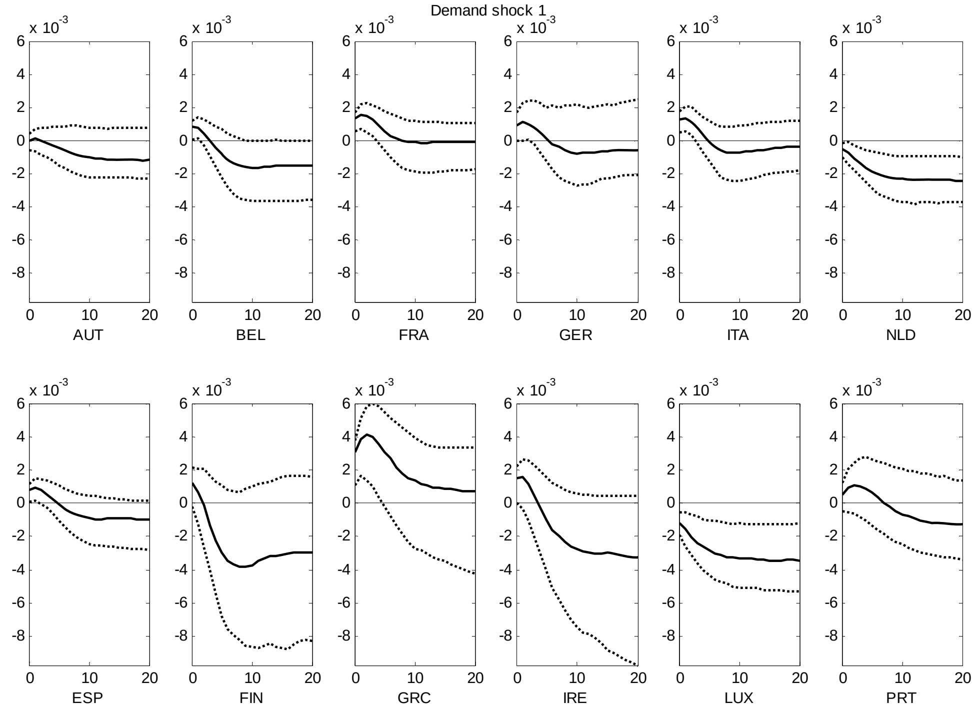
Our results are roughly in line with those of other analyses that investigate the propagation of common shocks to individual euro-area countries. Most of this literature also finds only moderate differences in the transmission of common monetary policy shocks (Ciccarelli and Rebucci, 2006; Clements et al., 2001; Sala, 2003; Eickmeier and Breitung, 2006), of aggregate euro-area supply and demand shocks (Eickmeier and Breitung, 2006). However, papers sometimes lack consensus on the ordering of individual countries in terms of deviations from the euro-area average response to common shocks. Our results are consistent

Figure 8: Impulse responses of individual countries' macro variables to euro-area shocks

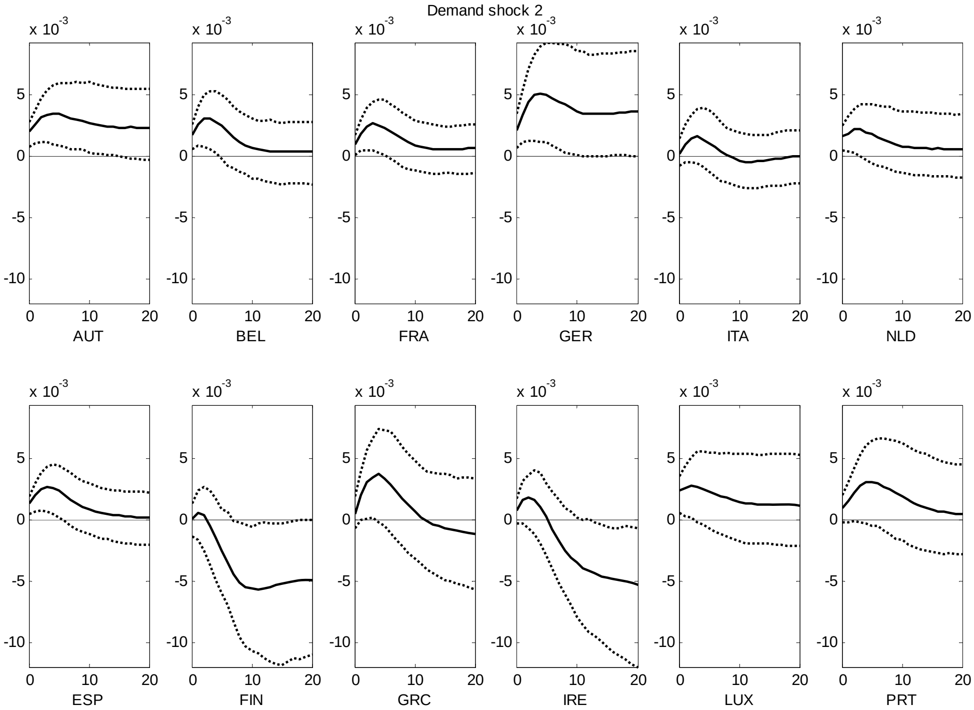
a) Output

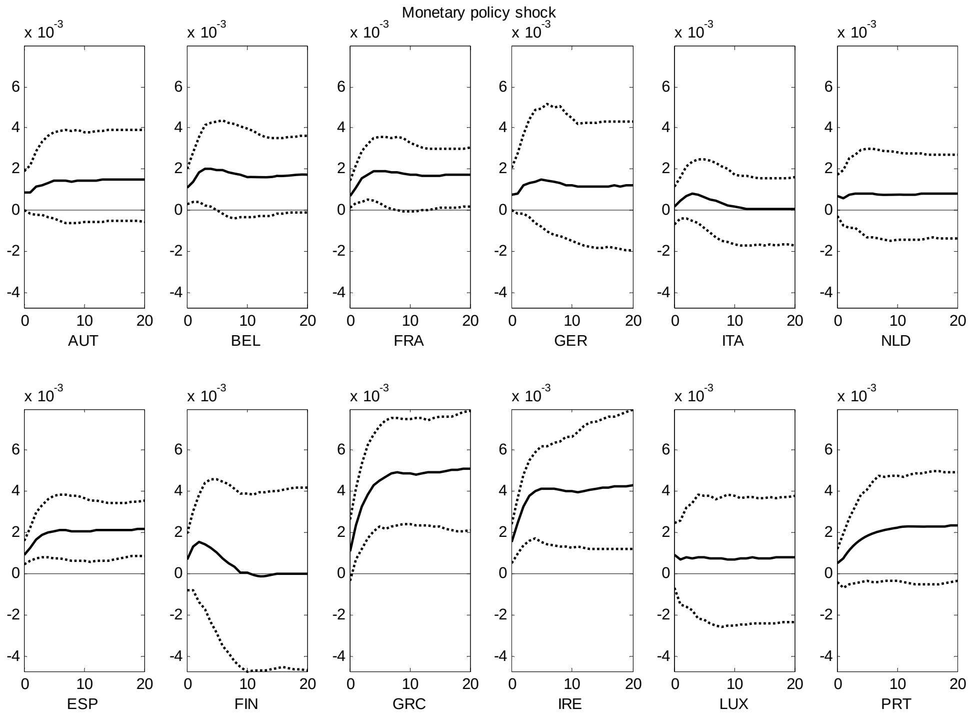


Demand shock 1

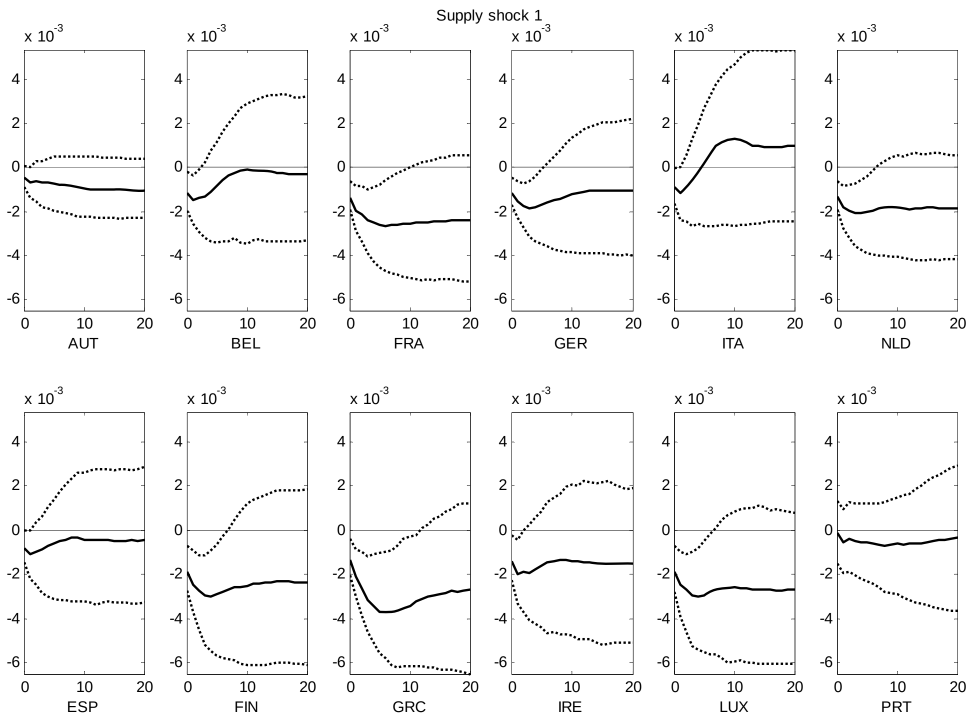


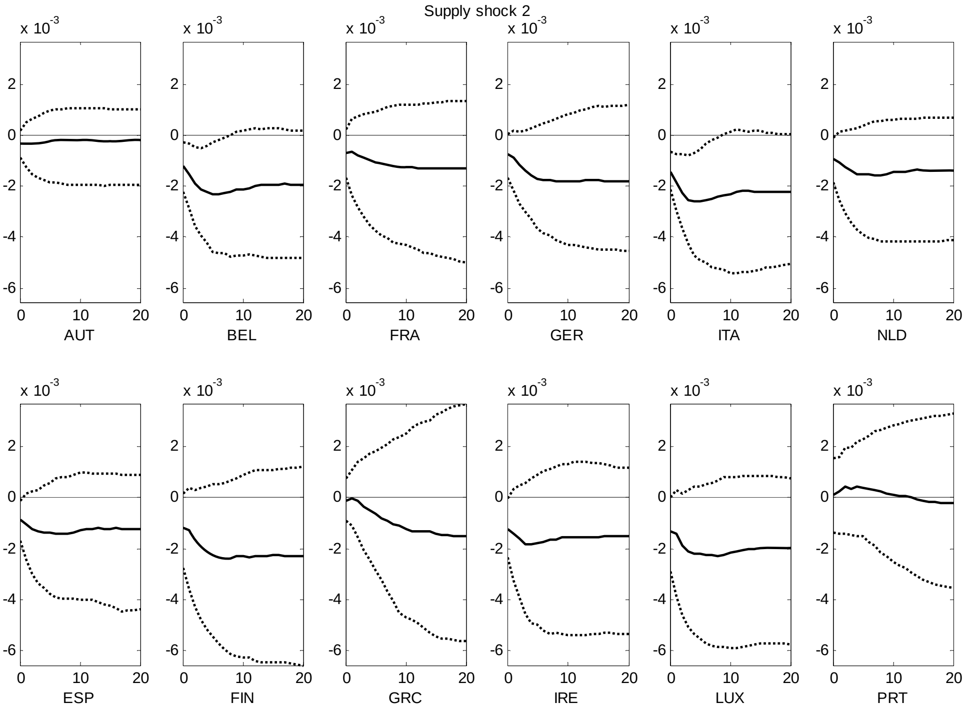
Demand shock 2

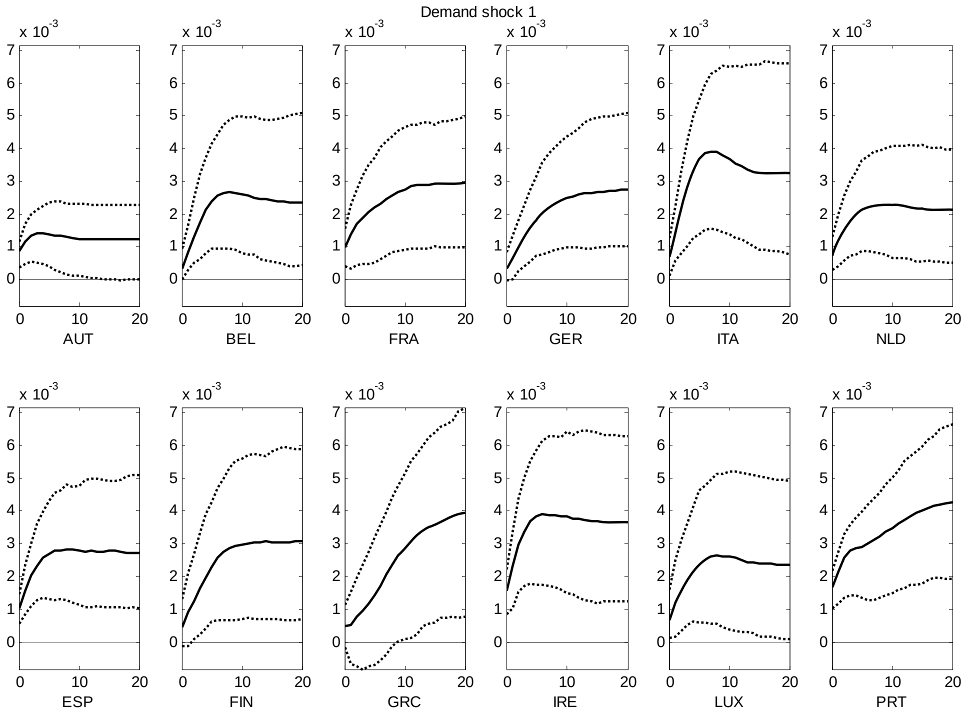


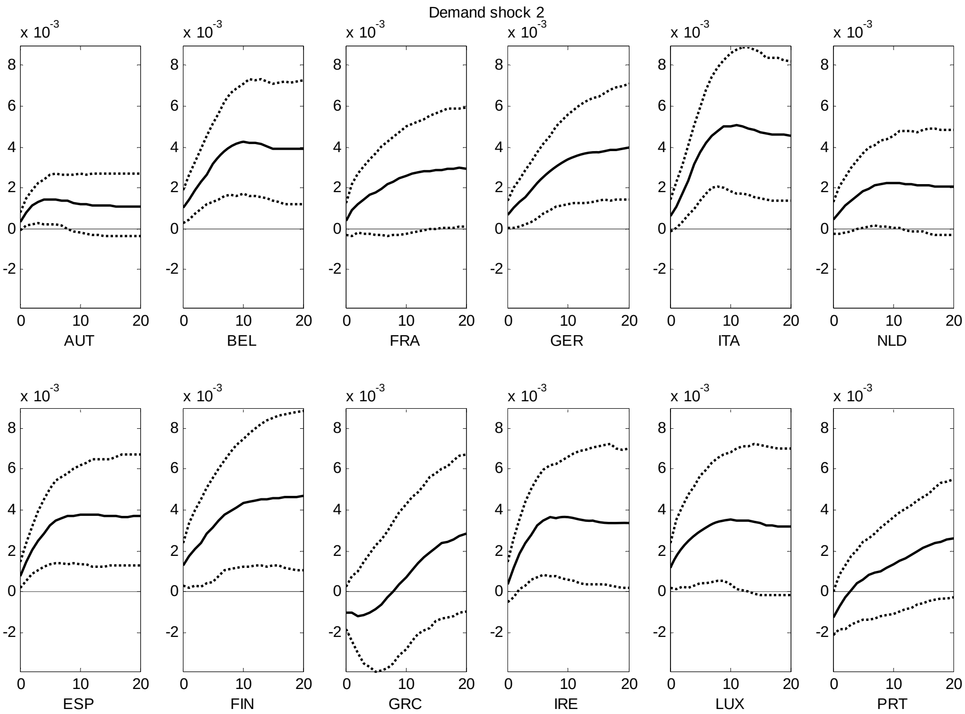

Monetary policy shock
x 10-3
AUT
BEL
FRA
GER
ITA
NLD
ESP
FIN
GRC
IRE
LUX
PRT

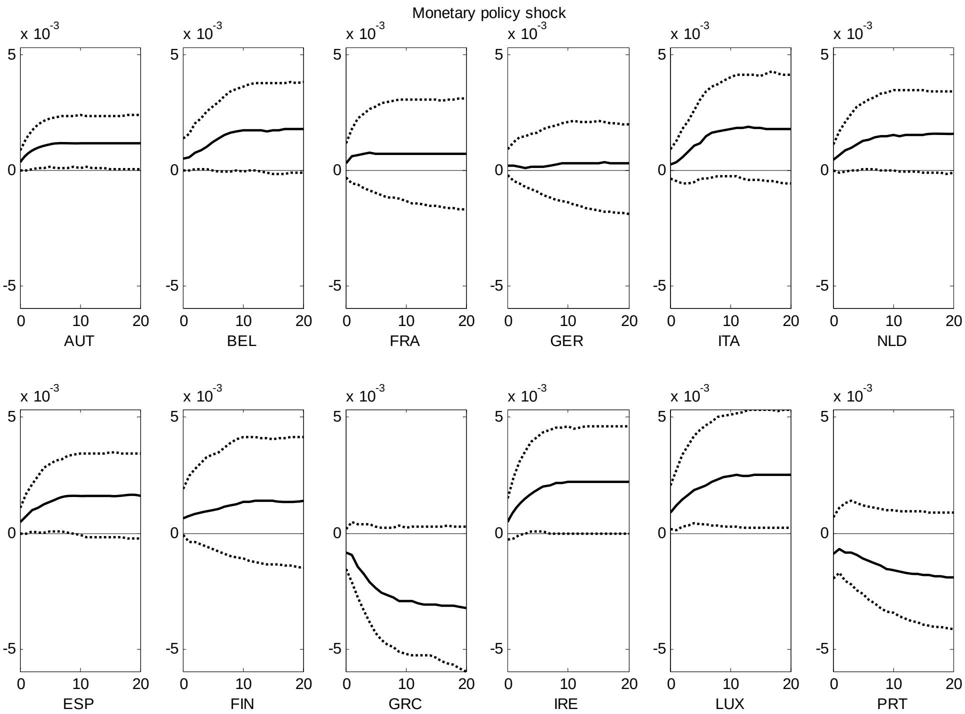
b) Prices


Supply shock 1
x 10-3
AUT
BEL
FRA
GER
ITA
NLD
ESP
FIN
GRC
IRE
LUX
PRT


Supply shock 2
x 10-3
AUT
BEL
FRA
GER
ITA
NLD
ESP
FIN
GRC
IRE
LUX
PRT


Demand shock 1
x 10-3
AUT
BEL
FRA
GER
ITA
NLD
ESP
FIN
GRC
IRE
LUX
PRT





Notes: Median (solid), 90% confidence bands (dotted).

with Sala (2003) who finds relatively weak output responses in the Netherlands and Italy to a common monetary policy shock. Sala (2003) and Ciccarelli and Rebucci (2006), in addition, find that economic activity in Germany and Spain tend to respond more strongly to the monetary policy shock than economic activity in France. This contradicts Clements et al. (2001); according to them, Portugal exhibits the weakest, and France and Finland the strongest response (cumulated after 5 years). We cannot confirm either of these findings.

Table 12: Forecast error variance decomposition of individual countries' macro variables

	Sup 1	Sup 2	Dem 1	Dem 2	Mon pol	Sup 1	Sup 2	Dem 1	Dem 2	Mon pol
			Output					Prices		
Forecast horizon of 0 to 1 year										
AUT	0.31	0.03	0.01	0.52	0.07	0.09	0.04	0.32	0.22	0.15
BEL	0.19	0.10	0.03	0.36	0.14	0.12	0.24	0.14	0.27	0.04
FRA	0.18	0.07	0.12	0.33	0.15	0.34	0.07	0.20	0.12	0.04
GER	0.45	0.05	0.02	0.36	0.03	0.29	0.16	0.12	0.21	0.02
ITA	0.43	0.24	0.07	0.11	0.03	0.06	0.27	0.28	0.23	0.03
NLD	0.42	0.06	0.08	0.27	0.04	0.28	0.13	0.18	0.12	0.06
ESP	0.18	0.04	0.04	0.40	0.20	0.07	0.10	0.29	0.31	0.07
FIN	0.08	0.52	0.11	0.06	0.09	0.32	0.13	0.08	0.21	0.03
GRC	0.08	0.03	0.34	0.18	0.24	0.49	0.03	0.06	0.11	0.14
IRE	0.06	0.13	0.09	0.11	0.46	0.12	0.11	0.35	0.16	0.07
LUX	0.36	0.04	0.15	0.25	0.05	0.27	0.12	0.10	0.18	0.08
PRT	0.55	0.02	0.04	0.22	0.05	0.06	0.06	0.57	0.09	0.10
Forecast horizon of 0 to 5 years										
AUT	0.32	0.03	0.04	0.36	0.09	0.12	0.05	0.23	0.22	0.17
BEL	0.13	0.21	0.09	0.18	0.16	0.06	0.13	0.16	0.41	0.07
FRA	0.13	0.19	0.06	0.20	0.22	0.22	0.07	0.24	0.23	0.04
GER	0.43	0.08	0.02	0.30	0.03	0.09	0.11	0.21	0.40	0.02
ITA	0.25	0.43	0.05	0.10	0.04	0.06	0.12	0.24	0.39	0.06
NLD	0.35	0.09	0.22	0.11	0.04	0.17	0.10	0.21	0.19	0.09
ESP	0.13	0.12	0.06	0.18	0.31	0.05	0.06	0.24	0.40	0.07
FIN	0.17	0.18	0.12	0.28	0.05	0.16	0.11	0.15	0.35	0.04
GRC	0.10	0.09	0.14	0.14	0.43	0.25	0.08	0.20	0.12	0.18
IRE	0.10	0.11	0.10	0.20	0.24	0.08	0.06	0.31	0.25	0.10
LUX	0.33	0.06	0.26	0.11	0.04	0.17	0.10	0.13	0.22	0.12
PRT	0.54	0.04	0.04	0.13	0.11	0.05	0.06	0.50	0.13	0.10

Notes: The median is shown. Output is real GDP, prices are consumer prices.

5. Conclusion

In this chapter, we have applied the non-stationary dynamic factor model of BN, complemented with the structural dynamic factor setup suggested by FR and FGLR, to a dataset of 173 stationary and non-stationary quarterly euro-area macroeconomic variables. The goal was to establish stylized facts on output and price comovements and heterogeneity across individual euro-area countries. The factor framework is particularly well suited to decompose heterogeneity into its components: idiosyncratic shocks and the asymmetric spread of common shocks. Moreover, with the BN approach, there is no need to impose restrictions on the persistence of series, common factors and, hence, comovements and heterogeneity *ex ante*.

We find that common permanent factors are important in explaining individual countries' output and price developments in the euro area. We also find that output and prices are not only hit by permanent common, but also by permanent idiosyncratic shocks. Idiosyncratic shocks and adjustments to them seem to be mainly responsible for cross-country heterogeneity during most of the sample. The asymmetric transmission of common shocks seems to play a minor role. We finally find no strong evidence that some common shocks lead to greater cross-country heterogeneity than others.

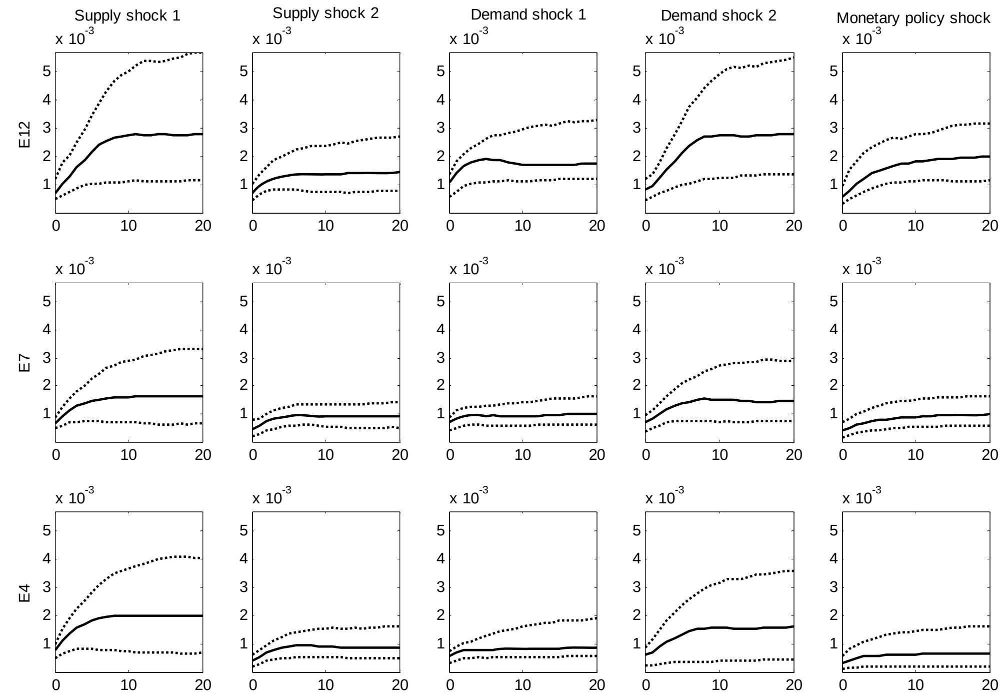
What are the policy implications? As we have explained above, not all observed heterogeneity, for example heterogeneity that goes along with the convergence process, leads to welfare losses and calls for policy intervention. However, even after (partly) prescinding from this type of heterogeneity, there seems to be considerable and persistent heterogeneity left. Given our finding that the remaining heterogeneity is, to a considerable extent, explained by idiosyncratic shocks which only slowly spread to individual countries' output and prices, national economic policies designed to carry out structural reforms to enhance factor mobility and to foster nominal flexibility would be well suited to speed up the adjustment to shocks and, in this way, to reduce such heterogeneities.

Comovements and heterogeneity in the euro area are, of course, also intensively studied in central banks, and there is a lively debate on the role of the ECB in the light of observed output and inflation differentials. Some papers find that monetary policy could improve overall welfare in a currency union if it gave a larger weight in the objective function to countries in which economic developments are more persistent (cf. Benigno, 2004; Benigno and López-Salido, 2006, who focus on inflation persistence). Other results suggest that heterogeneity could be lowered if cross-country differences in the transmission of a common monetary policy shock were exploited (Angelini et al., 2004). But those papers also acknowledge that such active and complex policies involve important risks not to be ignored.

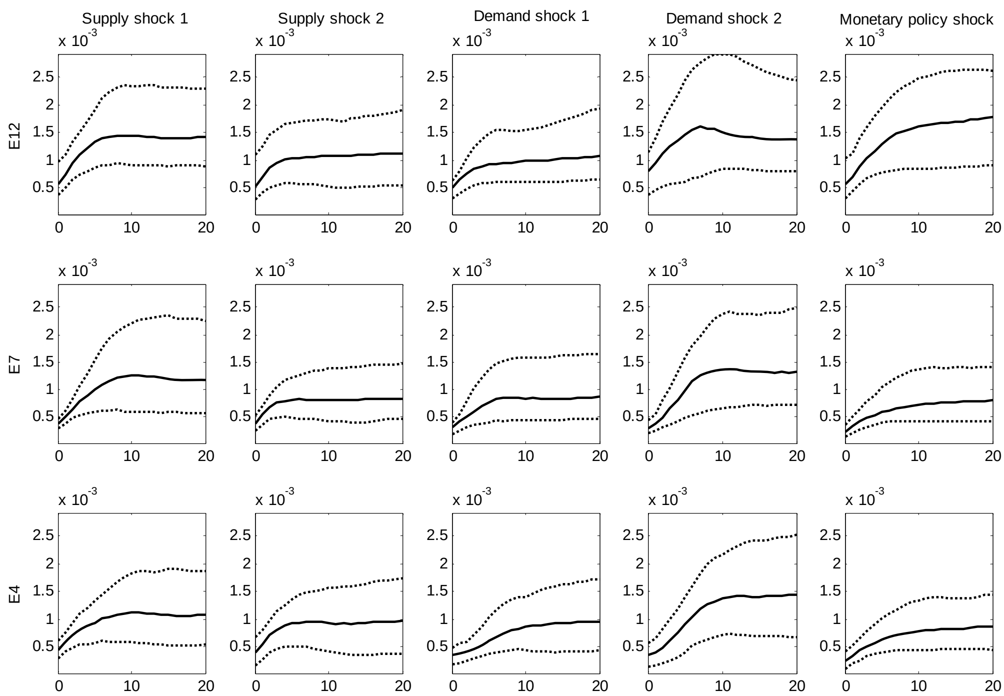
The first type of policy may reduce incentives for national governments to make necessary structural adjustments to increase flexibilities. The second type of policy would be difficult to implement given the relatively large uncertainty involved with the effects of monetary policy. This is even more true if monetary policy effects do not differ much across countries, one of our results in this study. The ECB tries to achieve price stability in the medium run. Insofar

Figure 9: Standard deviation of individual countries' impulse response functions

a) Output



b) Prices



Notes: Unweighted standard deviation of individual countries' GDP and CPI developments. Median (solid), 90% confidence bands (dotted). E12/7/4 refers to the groups of 12 euro-area countries, the 7 core euro-area countries and the 4 largest euro-area countries.

individual countries have enough time to adjust to shocks.[73] In addition, by aiming at keeping the inflation rate below, but close to 2%, the ECB ensures a safety margin which should avoid the possibility of individual countries encountering deflation.

Three possible extensions of this study come to our mind. First, in our framework, we could investigate heterogeneity at a more disaggregated level. For example, the analysis could be performed for the components of GDP, i.e. consumption, investment, government spending and external trade, which should help us to shed light on the determinants of dispersion. Second, as pointed out in Section 3, previous work suggests that the European integration process did not change considerably the transmission mechanism in the euro area. Nevertheless, future work could be devoted to estimating a time-varying parameters model such as the Bayesian dynamic factor model of Del Negro and Otrok (2005) to explicitly account for changes in economic comovements. As pointed out above, the framework proposed by the authors would also permit the inclusion of factors which affects only subsets of variables and, hence, distinguish between variable-specific and country-specific driving forces. Third, it would be interesting to fit alternative approaches such as the Factor Augmented VAR (FAVAR) model suggested by Bernanke et al. (2005), the Global VAR model developed by Pesaran et al. (2004) and the Bayesian Panel index VAR model developed by Canova and Ciccarelli (2004) to our dataset and to compare the outcomes. All these approaches are also suited to handle large datasets. The FAVAR is very similar to the model used in this chapter: the factors are estimated with principal component analysis which are then included in a VAR with observable variables. This model has, to our knowledge, not yet been applied to non-stationary datasets. The Global VAR and the Panel index VAR, by contrast, estimate the common factors as averages of observable variables, and the number of factors needs to be set. These two models can deal with non-stationary datasets. The Bayesian Panel index VAR model involves more complex and time-consuming estimation techniques than the other approaches, but can estimate time-varying parameters, just as the factor model suggested by Del Negro and Otrok (2005).

[73] See the speech by Otmar Issing at the ECB workshop on "Monetary policy implications of heterogeneity in a currency area", Frankfurt, 13-14 December 2004.

Appendix A (estimation of the shocks and construction of the confidence bands)

This appendix presents the structural analysis. The estimated vector of static stationary factors $\hat{f}_t = \Delta\hat{F}_t$ has the VAR(2) representation

$$\Psi(L)\hat{f}_t = u_t, \tag{A-A1}$$

with $\Psi(L) = I - \Psi_1 L - \Psi_2 L^2$. OLS is applied to each equation, yielding the reduced form VAR residuals $\hat{u}_t$. The q-vector of orthogonalized residuals v_t is estimated as

$$\hat{v}_t = \hat{M}^{-1/2}\hat{P}'\hat{u}_t, \tag{A-A2}$$

where $\hat{M}$ is a $q \times q$ matrix with the largest q eigenvalues of $\mathrm{cov}(\hat{u}_t)$ on the main diagonal and zeros elsewhere such that $\mathrm{cov}(\hat{v}_t) = I_q$. $\hat{P}$ is the corresponding $r \times q$ matrix of eigenvectors. The vector $\hat{v}_t$ is a consistent estimator of v_t. The estimated vector of structural shocks $\hat{w}_t$ is related to $\hat{v}_t$ through the $q \times q$ rotation matrix R:

$$\hat{w}_t = R\hat{v}_t, \tag{A-A3}$$

where $R'R = I_q$. Notice that, by construction, $\mathrm{cov}(\hat{w}_t) = I_q$. The matrix of impulse response functions at horizon h with respect to the structural shocks, $\partial y_{it+h} / \partial w_t' = \Theta_{ih}$, is obtained from

$$\Theta_i(L) = \Theta_{i0} + \Theta_{i1}L + \Theta_{i2}L + \ldots = \Lambda_i'\Psi(L)^{-1}PM^{1/2}R' \tag{A-A4}$$

(cf. FGLR). The rotation matrix R has to be chosen such that the identifying restrictions specified in the main text are satisfied.

Any q-dimensional rotation matrix can be parametrized as follows:

$$
\mathrm{R}(\theta) = \prod_{l,n} \begin{bmatrix} 1 & 0 & \cdots & & & & \cdots & 0 \\ 0 & \ddots & & & & & & 0 \\ \vdots & & \cos(\theta) & & & -\sin(\theta) & & \vdots \\ & & & \ddots & & & & \\ & & & & 1 & & & \\ & & & & & \ddots & & \\ & & \sin(\theta) & & & \cos(\theta) & & \vdots \\ \vdots & & & & & & \ddots & 0 \\ 0 & \cdots & & & & \cdots & 0 & 1 \end{bmatrix}, \qquad \text{(A-A5)}
$$

where only rows l and n are rotated by the angle θ_i, and there are $q(q-1)/2 = 10$ possible bivariate rotations. Hence, $\theta = \theta_1, ..., \theta_{q(q-1)/2}$ and each rotation angle is varied on a grid from 0 to π. The number of grids is chosen to be 24, and the rotation angles are fixed to satisfy the imposed restrictions.

Since $N \gg T$, the uncertainty involved with the factor estimation can be neglected (cf. Bernanke et al., 2005). In order to account for the uncertainty involved with the estimation of the VAR model on the factors, we construct confidence bands by means of the bootstrap-after-bootstrap techniques based on Kilian (1998). These techniques allow us to remove a possible bias in the VAR coefficients which can arise due to the small sample size of the VAR model. Most draws deliver not just one, but a set of shocks which all satisfy the restrictions. In this case, we follow Peersman (2005) and draw and save one of them. Some draws, however, do not deliver any shocks satisfying the restrictions. We draw until we have saved 300 shocks (1,269 draws on average were needed to get one shock which satisfied all restrictions). For more details on the identification, the reader is referred to Peersman (2005).

Appendix B (panel unit root tests of Harvey and Bates, 2003 and Breitung and Das, 2005)

This appendix describes the panel unit root tests of Harvey and Bates (2003, henceforth HB) and Breitung and Das (2005, henceforth BD) and how they were applied to the sets of idiosyncratic components of individual countries' output and prices. Let us consider a panel of $\widetilde{N} = 12$ idiosyncratic components Ξ_{it}, where $i = 1,..., \widetilde{N}$, and focus on the following model

$$\Delta\Xi_{it} = \phi\Xi_{it-1} + \sum_{l=1}^{\widetilde{p}_i} \alpha_{il}\Delta\Xi_{it-l} + \varepsilon_{it} \tag{A-B1}$$

for each unit. We then aim at testing the null hypothesis that all series have a unit root:

$$H_0 : \phi = 0. \tag{A-B2}$$

The short-term dynamics, i.e. the lags of $\Delta\Xi_{it}$ on the right hand side of equation (A-B1), are removed through a "pre-whitening" procedure suggested by BD. This involves estimating equation (A-B1) with OLS for each i, where $\widetilde{p}_i$ is determined with the Akaike criterion and is allowed to be specific for each i. The pre-whitened idiosyncratic components are then computed as

$$\widetilde{\Xi}_{it} = \Xi_{it} - \sum_{l=1}^{\widetilde{p}_i} \hat{\alpha}_{il}\Xi_{it-l} , \tag{A-B3}$$

where $\hat{\alpha}_{il}$ denotes the parameter estimate for unit i and lag l. The model can be written as a SUR system of equations

$$\begin{bmatrix} \Delta\widetilde{\Xi}_{1t} \\ \vdots \\ \Delta\widetilde{\Xi}_{\widetilde{N}t} \end{bmatrix} = \phi \begin{bmatrix} \widetilde{\Xi}_{1t-1} \\ \vdots \\ \widetilde{\Xi}_{\widetilde{N}t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{\widetilde{N}t} \end{bmatrix} \tag{A-B4}$$

or

$$\Delta\widetilde{\Xi}_t = \phi\widetilde{\Xi}_{t-1} + \varepsilon_t . \tag{A-B5}$$

The first test statistic we employ, t_{gls}, has been developed by HB and is the t-statistic based on a GLS regression:

$$t_{gls} = \frac{\sum_{t=1}^{\tilde{T}} \tilde{\Xi}'_{t-1}\,\Omega^{-1}\Delta\tilde{\Xi}_t}{\sqrt{\sum_{t=1}^{\tilde{T}} \tilde{\Xi}'_{t-1}\,\Omega^{-1}\tilde{\Xi}_{t-1}}}, \tag{A-B6}$$

where the unknown covariance matrix is replaced by its estimator

$$\hat{\Omega} = \frac{1}{T}\sum_{t=1}^{T}(\Delta\tilde{\Xi}_t - \hat{\phi}\tilde{\Xi}_{t-1})(\Delta\tilde{\Xi}_t - \hat{\phi}\tilde{\Xi}_{t-1})' \tag{A-B7}$$

and $\hat{\phi}$ denotes the OLS estimator of ϕ.

The second test statistic, t_{rob}, has been developed by BD (2005) and can be computed as

$$t_{rob} = \frac{\sum_{t=1}^{T} \tilde{\Xi}'_{t-1}\,\Delta\tilde{\Xi}_t}{\sqrt{\sum_{t=1}^{T} \tilde{\Xi}'_{t-1}\,\Omega\tilde{\Xi}_{t-1}}}. \tag{A-B8}$$

Critical values were obtained with Monte Carlo simulations. The series were simulated under the null for $\tilde{N}$ and T. The residuals were assumed to be normally distributed, and we rely on the empirical covariance matrix $\hat{\Omega}$. The simulated series were differenced, standardized and re-cumulated, as we did for the factor analysis. For this reason, our critical values may differ from the critical values reported in HB and BD. The number of replications was 5000. As for the structural factor analysis, we neglect the uncertainty involved with the estimation of the factors and, hence, the idiosyncratic components, since N is large (cf. Bernanke et al., 2005). Results are reported in Table 9.

V. Summary and Outlook

This dissertation both examines the forecast performance of large-scale factor models (chapter II) and employs these models to investigate international economic comovements (chapters III and IV). Over the last several years, factor models have gained great popularity and are widely applied by forecasters and by macroeconomic analysts. Of the advantages of factor models over previous empirical models, the possibility to exploit information from a large number of variables is probably the most important. The dissertation contributes in several dimensions which are potentially attractive for researchers dealing with factor models. It contains a systematic examination of the determinants of the forecast performance of factor models and has enabled us to reconcile the sometimes conflicting findings from the literature. Thus chapter II can hopefully be a useful guide for forecasters in policy and research institutions. Chapters III and IV contribute to interpreting these unobservable factors (or the shocks underlying the factors). The latter chapter shows how factor models applied to an international macroeconomic problem can be used to estimate stochastic trends. In what follows, we summarize chapters II to IV from a forecasters'/an economic point of view and give an outlook.

Chapter II investigates the forecast performance of factor models using a meta-analysis. We summarize the seemingly ambiguous results from existing studies which evaluate the forecast performance of factor models, and we identify the determinants of forecast quality in factor models using a meta-analysis. The main focus of chapter II has been the relative forecast performance of large-scale dynamic factor models for real economic activity and inflation. The relative factor forecast performance is approximated with the ratio of a root mean squared error (RMSE) obtained from a forecast based on a factor model and the RMSE obtained from a forecast based on a generally much smaller time series model. More than 50,000 relative RMSEs are taken from 52 studies. We first provide some descriptive statistics, and then derive possible determinants of the relative forecast performance of factor models from theory. These determinants affect the precision of factor estimates, the commonality of the target variable (i.e. the correlation between the variable to be predicted and the factors) and the specification of the forecasting equation. They also affect the forecast environment which has to be taken as given and the forecast design which can be influenced by the forecaster. The forecast environment determinants refer to the target variable – output or inflation of different countries/regions – to the benchmark model, and to the forecast horizon. The determinants of forecast design comprise the size of the dataset from which the factors are estimated and its characteristics (i.e. which criteria are used to determine inclusion of a variable), the frequency of the observations (monthly or quarterly), whether the forecaster relies on a balanced or an unbalanced panel, whether s/he makes rolling or recursive forecasts, the factor estimation techniques, whether the forecasting equation is estimated unrestricted or restricted (where the restrictions are provided by the factor model equation) and finally if direct or indirect multi-step forecasts are made. We estimate the impact of all these determinants on the relative factor forecast performance.

This analysis has the advantage of being less prone to subjectivity regarding the choice of papers and results than narrative survey articles (Stock and Watson, 2006; Reichlin, 2003; Breitung and Eickmeier, 2006). When compared to studies which concentrate on specific determinants of factor forecasts such as Kapetanios and Marcellino (2004), Schumacher (2007) and Boivin and Ng (2005, 2006), this meta-analysis is broader and can consider many possible determinants simultaneously.

Our analysis reaches several conclusions. First, forecasts can be improved if information from large datasets is exploited. This has been derived from our finding that factor models outperform smaller time series models. Alternative methods which are also able to exploit information from large datasets such as the combination of various smaller forecasting models ("pooled forecasts") even outperform or are comparable to factor models. This is also supported by the result that the size of the dataset from which the factors are extracted positively affects the predictions. The larger N and T are, the better the factor forecasts tend to be. Moreover, factor modellers are well advised to make recursive forecasts which are based on a longer estimation period, and to exploit information, not only from quarterly, but also from monthly data. Second, the target variable itself alters the quality of factor forecasts. According to our analysis, factor models are relatively better at predicting US than British variables. Moreover, factor models perform relatively better/worse than other models when predicting US output/inflation than when predicting euro-area output/inflation, whereas the opposite holds for inflation predictions. Third, it can pay off to carefully specify the model. More complex factor estimation techniques by Forni et al. (2000) and Kapetanios and Marcellino (2004) are shown to be better at predicting output than the Stock and Watson (1998, 2002a) approach which is less demanding on the specification. Fourth – and surprisingly –, a pre-selection of the variables included in the large dataset has not led to improvements in many of the past studies, although an improvement was shown by Boivin and Ng (2006) in simulations. A problem with these studies is that many of them used *ad hoc* methods to exclude or include variables in the dataset. In this respect, factor forecasts could certainly be improved further.

The meta-study also reveals further needs for research. Some cases are, up to now, characterized by only few observations; only 5% of all observations, for instance, are associated with rolling forecasts and less than 1% with the Kapetanios and Marcellino (2004) factor estimation technique. Results could change if new observations were added and the analysis was updated in the future. In addition, it would be interesting to verify if the factor models deliver relatively good forecasts in periods which are characterized by strong comovements or structural breaks, as is often claimed. So far, existing studies do not provide forecasts for periods which are clearly distinguished by strong or weak comovements of the variables, or by the presence of structural breaks. Finally, approaches which have been beyond the scope of the meta-analysis such as the combination of factor forecasts could be promising. This is derived from our finding that factor models as well as the combination of smaller models deliver good forecast results and is further suggested by studies which have already applied this ”factor forecast pooling” (Koop und Potter, 2004; Stock und Watson, 2006). A closer investigation in the future would certainly be of interest.

Chapter III estimates the extent and the dynamics of the transmission of macroeconomic shocks from the US to Germany between 1975 and 2002, with a particular focus on the second half of the 1990s and on 2001. In addition, the individual transmission channels are examined more thoroughly.

We construct a dataset containing almost 300 stationary German and US macroeconomic variables with variables covering real economic activity, prices, financial markets and external influences of both countries. We first estimate common factors by means of the structural dynamic factor model suggested by Forni and Reichlin (1998). We then separate US shocks from other common shocks using a method proposed by Uhlig (2005). This method consists in isolating shocks which explain the bulk of variation in the US economy. Short-term sign restrictions on impulse response functions (cf. Peersmann, 2005) help to identify a US demand shock and a US supply shock, and we investigate the impact of these shocks on German macroeconomic variables (GDP, consumption, investment, employment, prices etc.). We assess the importance of individual transmission channels through impulse response functions and variance decompositions of variables covering the transmission channels (trade, financial markets and confidence variables). We finally decompose the forecast error of German GDP for the second half of the 1990s and the recession phase in the US in 2001.

The analysis is one of the first applications of the structural factor model to an international macroeconomic topic. Previous articles use VAR models (Canova and Marrinan, 1998; SVR, 2001; Artis et al., 2006; Dees et al., 2007), large structural macroeconometric multi-country models (Dalsgaard et al., 2001; IMF, 2001; SVR 2001) or international general equilibrium models (Adjemian et al., 2004; De Walque und Wouters, 2004). Factor applications also exist, however, without an emphasis on the structural interpretation of the factors (Monfort et al., 2004; Kose et al. 2003a; Lumsdaine and Schlagenhauf, 1996). Only the articles by Sala (2003) and Eickmeier und Breitung (2006) (the latter article was written later than the study underlying chapter III) identify the structural shocks which drive the common factors and examine the transmission of a monetary policy shock on the core euro-area countries and the effect of the three structural euro-area shocks on the euro-area and the central and east European countries.

The focus on transmission channels is new. It is unclear theoretically whether the various transmission channels lead to a positive or a negative international propagation of shocks. Instead, this question needs to be answered empirically. An advantage of the factor approach is that all transmission channels can be examined simultaneously, while VAR models often run into scarce degrees of freedom problems. “New channels” are often not included in structural models due to a lack of consensus on the exact modelling of these channels.

Another contribution is the separation of the US shocks which are transmitted to Germany from other (external) common shocks such as oil prices shocks or global demand shocks. The latter common shocks and those shocks that originate in one country and spread to other countries are difficult to disentangle. Some studies assume that the former are transmitted simultaneously whereas the latter are transmitted with a lag (Monfort et al., 2004). The drawback of such an identification is that country-specific shocks which are transmitted to other countries within a period – often a quarter – are counted among the common shocks. In the present work, we follow a different approach. We identify the main driving forces of the US economy and label them the US shocks.

US shocks are found to affect the US and the German economies largely symmetrically. The US demand shock has a stronger effect on the German economy than the US supply shock. A “typical” (i.e. a temporary one standard deviation) US demand shock triggers an immediate increase of US GDP of roughly 3%. The impact rises slightly further, then declines and dies out after a year. Such a shock leads to an instantaneous increase of German GDP of slightly more than 2%; its effect is significant for about one year. Roughly 7% of German GDP fluctuations is explained with US demand shocks.[74] In contrast, German GDP goes up after a US supply shock, but the impulse response is insignificant. The trade channel dominates. We are not able to draw clear conclusions on the role of financial markets and the confidence channel. Interestingly, German confidence indicators are shown to be only affected by US shocks since the end of the 1990s. This may indicate that the confidence channel has become relevant only in recent years. A historical decomposition of German GDP finally shows that negative domestic factors overcompensated positive influences from the US between 1995 and 2000. In contrast, the US recession in 2001 was the main culprit for the German slump.

This chapter could be extended by applying a model with time-varying factor loadings to our dataset to account for the increased economic integration between Germany and the US. It uncovers two further research needs. First, the factor analysis does not replace a structural model which carefully models the individual transmission channels. Our factor approach allows us to assess the transmission of the shocks to the variables approximating the transmission channels. However, it does not allow us to make a statement on how fluctuations in these variables affect key macroeconomic variables such as output and prices in Germany.

[74] This is a rough approximation. Variance decompositions suggest that 9% of the common component of German GDP is explained by the US demand shock at the 5-year forecast horizon. The common component accounts for 79% of German GDP growth.

A deeper structural analysis would have this ability. Second, a more detailed analysis of the trade channel which would distinguish between partner countries and commodity groups would be of interest.

Chapter III can also be used to (tentatively) predict the impact of the current crisis in the US housing market on the German economy. Since the transmission has been shown to depend on the type of shock, it is useful to first characterize the current crisis in the US; see Dynan and Kohn (2007), Roubini (2007), Mishkin (2007), IMF (2007) for more elaborate discussions. Between 1995Q1 and 2005Q4, house prices, proxied by the Freddie Mac Conventional Mortgage Home Price Index, rose by 80% more than consumer prices, and studies trying to explain housing price developments in terms of fundamentals clearly suggest overevaluation of housing in recent years (cf. Girouard et. al., 2006). In the mid-1990s subprime lending expanded, leading to a relaxation of credit rationing for borrowers previously considered too risky by traditional lenders. This, together with overoptimistic households, led to a strong increase in housing demand and in households' indebtedness.[75] Since mid-2005, risks have been reassessed, and the moderation in economic growth and higher mortgage interest rates have made it more difficult for some borrowers to service their loans. Subprime delinquency rates doubled between mid-2005 and mid-2007 to 13%, and credit conditions tightened. Housing prices have strongly decelerated; they only rose by 0.1% in 2007Q2 compared to the previous quarter while consumer prices increased by 1.5%. Private residential investment has fallen by almost 20% since the beginning of 2006. Consumer confidence has also declined.

The impact of the burst of the housing price bubble has not fully materialized yet in the US. Its full effect is uncertain and depends on a number of circumstances. First, it seems likely that the deceleration of housing prices has not yet come to an end.[76] Second, private consumption growth which to date has remained relatively strong, is likely to be negatively affected in the coming months through the wealth effect and tightened credit conditions. Further declines in consumer confidence, oil prices which are at historically high levels, a labor market which has started to weaken and further tightening of monetary conditions are other downside risks for private consumption growth. Third, the extent to which the burst of the house price bubble will affect the US economy will further depend on the extent that the crisis will remain restricted to the housing market.[77]

To what extent can our analysis of chapter III help anticipate the impact on the German economy? The decline in residential investment and the possibly weaker personal

[75] The rate of homeowners rose from 64% in the mid-1990s to 69% in 2005. Their indebtedness increased by almost 4 percentage points to 18%.

[76] Roubini (2007) argue that excess supply in the housing market may increase further, and housing price inflation could further be dampened He also discusses the different channels.

[77] Vehicle sales and spending on consumer durables related to housing such as furniture and household equipment have already weakened. Credit tightening could extend beyond the subprime segment: total consumer credit are already at very growth low rates, which could further depress private consumption. Other parts of the financial market could be affected as equity markets experience greater volatility.

consumption expenditure represent a negative US demand shock.[78] It is very likely that this slowdown will have a negative effect on the German economy through the trade channel.

There are, however, two reasons to believe that the analysis of chapter III should be supplemented to fully estimate the international dimension of the current US crisis. First, the latter is not a "typical" demand shock for various reasons. Currently, the impact is widely concentrated on one segment of the economy, housing, and may remain so. Housing prices may also decelerate in countries which have experienced prior strong increases, such as in Ireland, Spain and the UK. In Germany, however, housing prices have stagnated and even have fallen in some recent periods. This might suggest *ceteris paribus* a smaller impact of the US crisis on the German economy than is estimated in existing studies. Most other arguments, however, point to a greater impact. The housing crisis is coupled with a crisis of the financial system and a loss in confidence, not only in the US, but also in other countries. Although we have included financial variables and confidence measures in our large dataset, our backward-looking study cannot fully grasp these "newer" channels and probably understates the impact on the German economy. Another channel not captured by our analysis is the transmission to foreign banks. Serious problems for the banking sector (leading to bank runs) due to the US housing crisis and financial turmoil have been observed only in the British banking system, but not in Germany so far. The crisis in the US is also not typical in the sense that it has been triggered by the burst of a bubble. Such a burst is a rare event and should be compared with similar events such as the strong stock market crashes in 1987 and 2001. Such comparisons are, however, beyond the scope of this thesis.

Second, it should be kept in mind that dynamic factor models are linear models which cannot account for transmission asymmetry. Studies employing non-linear empirical models find that negative real shocks are transmitted to a larger extent internationally than positive shocks (Artis et al., 2007; GCEE, 2001; Canova et al., 2007a; Osborn et al., 2005), and we have provided theoretical arguments that support these findings.[79]

Chapter IV establishes stylized facts about business-cycle and long-run comovements and dispersion in the euro area. This chapter is motivated by the observation that comovements at business-cycle and long-run frequencies among EMU-members are far from perfect, and there is still persistent heterogeneity, although countries are tightly linked through trade and financial markets. We establish stylized facts about comovements and heterogeneity of output and price developments in EMU member states and their determinants. We combine the recently developed PANIC ("Panel Analysis of Nonstationarity in Idiosyncratic and Common components") approach of Bai and Ng (2004) and the structural factor setup based on Forni and Reichlin (1998) and apply them to a newly constructed partly non-stationary dataset containing a total of 173 quarterly macroeconomic times series from 1981 to 2003, which

[78] Quarterly growth rates of domestic demand were less than 0.4% on average since 2006Q2 compared to over 0.8% since the beginning of the latest expansion.

[79] Cf. Ball and Mankiw (1994) and Peersman and Smets (2002).

capture economic developments in euro-area countries along with some external influences. The PANIC method allows estimation of the common and idiosyncratic components of individual countries' output and prices, while the structural factor setup allows a structural analysis, that is, an identification of common structural shocks along with their propagation to output and prices. We also decompose the heterogeneity, first assessing to what extent it is due to idiosyncratic shocks and adjustments to these shocks and to what extent it is due to the asymmetric spread of common shocks. We then decompose the latter determinant further and investigate whether some common shocks trigger more heterogeneity than others.

This study goes beyond existing studies which also examine economic linkages in the euro-area in two respects (Marcellino et al., 2000; Altissimo et al., 2004; Forni and Reichlin 2001; Beck et al., 2006; Altissimo et al., 2001; Sala, 2003; Eickmeier and Breitung, 2006). First, these studies fit stationary factor models to stationary datasets. The Bai and Ng (2004) approach allows us to examine comovements and heterogeneity without imposing restrictions on the persistence of the variables and their components (which are allowed to be non-stationary) and hence also on comovements and heterogeneity. This is a particularly favorable feature: it is not short-run but persistent heterogeneity that may indicate structural rigidities or inappropriate policies and that may be relevant for policy makers. Second, of the studies mentioned above, only Sala (2003) and Eickmeier and Breitung (2006) are concerned about the economic interpretation of the factors. They fit a structural dynamic factor model to the euro-area dataset and identify the structural shocks driving the common factors. While Sala (2003) focuses on a monetary policy shock, Eickmeier and Breitung (2006) identify aggregate euro-area supply, demand and monetary policy shocks. Chapter IV identifies a richer set of shocks and investigates their transmission.

The most important results can be summarized as follows. Common permanent factors are important in explaining individual countries' output and price developments in the euro area. Output and prices are, however, not only hit by permanent common, but also by permanent idiosyncratic shocks. Idiosyncratic shocks and adjustments to them seem to be mainly responsible for cross-country heterogeneity throughout most of the sample period. The asymmetric transmission of common shocks seems to play a minor role. We find no strong evidence that some common shocks lead to greater cross-country heterogeneity than others.

Future work could be devoted to estimating a model with time-varying parameters to better account for the European integration process or a model which permits the inclusion of factors which only affect specific variables or groups of variables and which could account for a "clustering" of certain countries. Heterogeneity could be investigated at a more disaggregated level (i.e. consumption, investment, government spending etc.) which should shed light on the determinants of dispersion. Moreover, an extended analysis could try to interpret the idiosyncratic shocks; fiscal policy is one candidate driver. Finally, it would be interesting to fit alternative approaches such as the Global VAR model of Pesaran et al. (2004) and the Panel index VAR -model of Canova and Ciccarelli (2004) to our dataset.

References

Adjemian, S., M. Darracq Parriès, F. Smets (2004), "Structural analysis of US and EA business cycles", mimeo. http://www.ecb.int/events/pdf/conferences/ecbimf/paper.darracq.paries.pdf.

Aguirre, A., L. F. Céspedes (2004), "Uso de análisis factorial dinámico para proyecciones macroeconómicas", Working Papers Central Bank of Chile 274.

Ahmed, S., B. W. Ickes, P. Wang, B. S. Yoo (1993), "International business cycles", *American Economic Review*, 83(3), 335-359.

Ahmed, S., J. H. Park (1993), "Sources of macroeconomic fluctuations in small open economies", *Journal of Macroeconomic*s, 16(1), 1-36.

Altig, D., L. Christiano, M., Eichenbaum, J. Linde (2002), "Technology shocks and aggregate fluctuations", mimeo. http://www.moneyworkshop.ch/documents/papers/christiano2.pdf.

Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Hallin, M. Lippi, L. Reichlin (2001), "EuroCOIN: a real time coincident indicator of the euro area business cycle", CEPR Working Paper 3108.

Altissimo, F., P. Benigno, D. R. Palenzuela (2004), "Inflation differentials in a currency area: facts, explanations and policy", mimeo. http://www.ecb.int/events/pdf/conferences/mpimphet/Altissimo_Benigno_RodriguezPalenzuela.pdf.

Amengual, D., M. Watson (2006), "Consistent estimators of the number of dynamic factors in a large N and T panel", mimeo. http://www.wws.princeton.edu/mwatson/papers/dyn_shocks_3.pdf.

Anderton, R., F. Di Mauro, F. Moneta (2004), "Understanding the impact of the impact of the external dimension of the euro area: trade, capital flows and other international macroeconomic linkages", ECB Occasional Paper 12.

Angelini, E., J. Henry, R. Mestre (2001), "Diffusion index-based inflation forecasts for the euro area", ECB Working Paper 61.

Angelini, P. , S. Siviero, D. Terlizzese (2004), "Monetary policy as a sheep dog", mimeo, http://www.ecb.int/events/pdf/conferences/mpimphet/Angelini_Siviero_Terlizzese.pdf

Angeloni, I., A. Kashyap, B. Mojon (eds.) (2003), “Monetary policy transmission in the euro area”, Cambridge University Press.

Artis, M. J., A. Banerjee, M. Marcellino (2005), “Factor forecasts for the UK”, *Journal of Forecasting*, 24(4), 279-298.

Artis, M. J., A. B. Galvão, M. Marcellino (2007), “The transmission mechanism in a changing world”, *Journal of Applied Econometrics*, 22(1), 39-61.

Artis, M., D. Osborn., P. J. Perez (2006), “The international business cycle in a changing world: volatility and the propagation of shocks in the G-7”, *Open Economics Review*, 17(3), 255-279.

Bai, J. (2004), “Estimating cross-section common stochastic trends in nonstationary panel data”, *Journal of Econometrics*, 122, 137-183.

Bai, J., S. Ng (2002), “Determining the number of factors in approximate factor models”, *Econometrica*, 70(1), 191-221.

Bai, J., S. Ng (2004), “A PANIC attack on unit roots and cointegration”, *Econometrica*, 72(4), 1127-1177.

Bai, J., S. Ng (2006), “Evaluating Latent and Observed Factors in Macroeconomics and Finance", *Journal of Econometrics*, 113(1-2), 507-537.

Bai, J., S. Ng (2007a), “Panel unit root tests with cross-section dependence: a further investigation”, mimeo, http://www-personal.umich.edu/~ngse/papers/newpanic.pdf.

Bai, J., S. Ng (2007b), “Determining the number of primitive shocks in factor models”, *Journal of Business and Economic Statistics*, 25(1), 52-60.

Bai, J., P. Perron (1998), “Estimating and testing linear models with multiple structural changes”, *Econometrica*, 66, 47-78.

Bai, J., P. Perron (2003), “Computation and analysis of multiple structural change models”, *Journal of Applied Econometrics*, 18, 1-22.

Ball, L., N. Mankiw (1994), “Asymmetric price adjustment and economic fluctuations”, *The Economic Journal*, 104, 247-261.

Banerjee, A., M. Marcellino (2006), “Are there any reliable leading indicators for US inflation and GDP growth?” *International Journal of Forecasting*, 22, 137-151.

Banerjee, A., M. Marcellino, I. Masten (2005), "Leading indicators for euro-area inflation and GDP growth.", *Oxford Bulletin of Economics and Statistics*, 67, 785-814.

Banerjee, A., M. Marcellino, I. Masten (2006a), "Forecasting macroeconomic variables for the new member states" in: Artis, M. J., A. Banerjee, M. Marcellino (eds.), The central and eastern European countries and the European Union, Cambridge University Press, Cambridge, Chapter 4, 108-134.

Banerjee, A., M. Marcellino, I. Masten (2006b): "Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples with Structural Change", mimeo. http://www.iue.it/Personal/Banerjee/papers/handbook_12_12_06_Elsevier%20Style.pdf.

Baxter, M., M. A. Kouparitsas (2005), "Determinants of business cycle comovement: a robust analysis", *Journal of Monetary Economics*, 52(1), 113-157.

Bayoumi, T., Helbling, T. (2003), "Are they all in the same boat? The 2000-2001 growth slowdown and the G-7 business cycle linkages", IMF Working Paper WP/03/46.

Beaton, A.E., J.W. Tukey (1974), "The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data", *Technometrics*, 16(2), 147-185.

Beck, G., K. Hubrich, M. Marcellino (2006), "Regional inflation dynamics within and across euro area countries and a comparison with the US", ECB Working Paper 681.

Benati, L., G. Kapetanios (2003), "Structural breaks in inflation dynamics", mimeo.

Benigno, P. (2004), "Optimal monetary policy in a currency area", *Journal of International Economics*", 63(2), 293-320.

Benigno, P., D. López-Salido (2006), "Inflation persistence and optimal monetary policy in the euro area", *Journal of Money, Credit and Banking*, 38(3), 587-614.

Bergman, M. (2004), "How similar are European business cycles?" EPRU Working Paper 2004-13.

Bernanke, B., J. Boivin, P. Eliasz (2005), "Measuring the effects of monetary policy: a Factor-Augmented Vector Autoregressive (FAVAR) approach", *The Quarterly Journal of Economics*, 120(1), 387-422.

Bernanke, B., J. Boivin (2003), "Monetary policy in a data-rich environment", *Journal of Monetary Policy*, 50(3), 525-546.

Boivin, J., S. Ng (2005), "Understanding and comparing factor-based forecasts", *International Journal of Central Banking*, 1, 117-151.

Boivin, J., S. Ng (2006), "Are more data always better for factor analysis", *Journal of Econometrics*, 132, 169-194.

Breitung, J., S. Das (2005), "Panel unit root tests under cross sectional dependence", *Statistica Neerlandica*, 59(4), 1-20.

Breitung, J., S. Eickmeier (2006), "Dynamic factor models", in: O. Hübler and J. Frohn (eds.), Modern econometric analysis, chapter 3, Springer.

Breitung, J., S. Eickmeier (2007), "Testing for structural breaks in dynamic factor models", mimeo.

Breitung, J., U. Kretschmer (2005), "Determining the number of dynamic factors in large macroeconomic panels", University of Bonn Working Paper.

Breitung, J., M. H. Pesaran (2006), "Unit roots and cointegration in panels", in: L. Matyas and P. Sevestre (eds.), The econometrics of panel data: fundamentals and recent developments in theory and practice, Kluwer Academic Publishers, forthcoming.

Brisson, M., B. Campbell, J.W. Galbraith (2003), "Forecasting some low-predictability time series using diffusion indices", *Journal of Forecasting*, 22, 515-531.

Bruneau, C., O. de Bandt, A. Flageollet (2003), "Forecasting inflation in the euro area", Banque de France NER 102.

Bruneau, C., O. de Bandt, A. Flageollet, E. Michaux (2007), "Forecasting inflation using economic indicators: the case of France", *Journal of Forecasting*, 26, 1-22.

Buisán, A., F. Restoy (2005), "Cross-country macroeconomic heterogeneity in EMU", Banco de España Documentos Ocasionales 0504.

Burns, A. F., W.C. Mitchell (1946), " Measuring business cycles", New York, NBER.

Busetti, F., L. Forni, A. Harvey, F. Venditti (2006), "Inflation convergence and divergence within the European Monetary Union", ECB Working Paper 574.

Camacho, M., I. Sancho (2003), "Spanish diffusion indexes", *Spanish Economic Review*, 5, 173-203.

Camba-Méndez, G., G. Kapetanios (2005), "Forecasting euro area inflation using dynamic factor measures of underlying inflation", *Journal of Forecasting*, 25, 491-503.

Canova, F. (2004), “Testing for convergence clubs: a predictive density approach”, *International Economic Review*, 45, 49-77.

Canova, F., M. Ciccarelli, E. Ortega (2007a), “Similarities and convergence in G-7 countries”, *Journal of Monetary Economics*, 54(3), 850-878.

Canova, F., M. Ciccarelli, E. Ortega (2007b), “Do political events affect business cycles? The Maastricht treaty, the creation of the ECB and the euro economy”, mimeo.

Canova, F., M. Ciccarelli (2004), “Formulation, estimation and testing of Bayesian Panel VAR models”, *Journal of Econometrics*, 120, 327-359.

Canova, F., H. Dellas (1993), “Trade interdependence and the international business cycle”, *Journal of International Economics*, 34, 23-47.

Canova, F., G. De Nicoló (2003), “On the sources of business cycles in the G-7”, *Journal of International Economics*, 59, 77-100.

Canova, F., J. Marrinan (1998), “Sources and propagation of international output cycles: common shocks or transmission?”, *Journal of International Economics*, 46, 133-166.

Carvalho, V. M., A. C. Harvey (2005). “Convergence in the trends and cycles of Euro-zone income”, *Journal of Applied Econometrics*, 20(2), 275-289.

Cristadoro, R., M. Forni, L. Reichlin, G. Veronese (2005), “A core inflation indicator for the euro area”, *Journal of Money, Credit and Banking*, 37(3), 539-560.

Chamberlain, G., M. Rothschild (1983), “Arbitrage, factor structure and mean-variance analysis in large asset markets”, *Econometrica*, 51, 1305-1324.

Cheung, C., F. Demers (2007), “Evaluating Forecasts from Factor Models for Canadian GDP Growth and Core Inflation”, Bank of Canada Working Paper 8.

Christiano, L., M. Eichenbaum, C. Evans (2005), “Nominal rigidities and the dynamic effects of a shock to monetary policy”, *Journal of Political Economy*, 113(1), 1-45.

Ciccarelli, M., A. Rebucci (2006), “Has the transmission mechanism of European monetary policy changed in the run-up to EMU”, *European Economic Review*, 50(3), 737-776.

Cimadomo, J. (2004), “The effects of systematic monetary policy on sectors: a factor model analysis”, mimeo. http://www.cepii.fr/anglaisgraph/pagepers/cimadomoecares.pdf.

Clark, T., M. McCracken (2001), “Tests of equal forecast accuracy and encompassing for nested models”, *Journal of Econometrics*, 105, 85-110.

Clements, B, Z. Kontolemis, J. Levy (2001), "Monetary policy under EMU: differences in the transmission mechanism", IMF Working Paper 01/102.

Corvoisier, S., B. Mojon (2005), "Breaks in the mean of inflation: how they happen and what to do with them", mimeo.

Cristadoro, R., M. Forni, L. Reichlin, G. Veronese (2005), "A core inflation indicator for the euro area", *Journal of Money, Credit and Banking*, 37(3), 539-560.

Croux, C., M. Forni, L. Reichlin (2001), "A measure for comovement for economic variables: theory and empirics", *The Review of Economics and Statistics*, 83(2), 231-241.

D´Agostino, A., D. Giannone (2006), "Comparing alternative predictors based on large-panel factor models", ECB Working Paper 680.

Dalsgaard, T., C. André, C., P. Richardson (2001), "Standard shocks in the OECD Interlink model", OECD Economics Department Working Paper 306.

Dees, S., F. Di Mauro, M.H. Pesaran, L.V. Smith (2007), "Exploring the international linkages of the euro area: a global VAR analysis", *Journal of Applied Econometrics*, 22(1), 1-38.

Del Negro, M., C. Otrok (2005), "A dynamic factor model with time-varying parameters", mimeo.

De Mol, C., D. Giannone, L. Reichlin (2006): "Forecasting using a large number of predictors. Is Bayesian regression a valid alternative to principal components?, ECB Working Paper 700.

Den Reijer, A. (2005), "Forecasting Dutch GDP using large scale factor models", DNB Working Paper 28.

De Walque, G., R. Wouters (2004), "An open economy DSGE model linking the euro area and the US", mimeo. http://www.bundesbank.de/download/vfz/konferenzen/20041126_27_eltville/vfz_20041126_4.pdf.

Diebold, F., R. Mariano (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics*, 13, 253-263.

Dornbusch, R., Y. C. Park, S. Claessens (2000), "Contagion: understanding how it spreads", The World Bank Research Observer, 15(2), 177-197.

Doucouliagos, C. (2005), “Publication bias in the economic freedom and economic growth literature”, *Journal of Economic Surveys*, 19(3), 367-387.

Doyle, B. M., J. Faust (2002), “An investigation of comovements among the growth rates of the G-7 countries”, *Federal Reserve Bulletin*, October, 427-437.

Doz, C., D. Giannone, L. Reichlin (2006), “A quasi-maximum likelihood approach for large approximate dynamic factor models“, ECB Working Paper 674.

Dynan, K.E., D.L. Kohn (2007), “The rise in U.S. household indebtedness: causes and consequences”, FRB Working Paper 2007-37, Finance and Economic Discussion Series.

ECB (2003), “Inflation differentials in the euro area: potential causes and policy implications”, mimeo.

ECB (2004), “The monetary policy of the ECB”, www.ecb.int/pub/pdf/other/monetarypolicy2004en.pdf.

Eickmeier, S. (2007), “Business cycle transmission from the US to Germany – a structural factor approach”, *European Economic Review*, 51(3), 521-551.

Eickmeier, S. (2005), “Common stationary and non-stationary factors in the euro area analyzed in a large-scale factor model”, Bundesbank Discussion Paper 02/2005, revised version.

Eickmeier, S. (2006), “Comovements and heterogeneity in the euro area analyzed in a non-stationary dynamic factor model”, Bundesbank Discussion Paper 31/2006, revised version.

Eickmeier, S., J. Breitung (2006), “How synchronized are new EU member states with the euro area? Evidence from a structural factor model”, *Journal of Comparative Economics*, 34, 538-563.

Eickmeier, S., C. Ziegler (2007), “How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach”, *Journal of Forecasting*, forthcoming.

Elliott, G., A. Fatás (1996), “International business cycles and the dynamics of the current account”, *European Economic Review*, 40(2), 361-387.

Faust, J. (1998), “The robustness of identified VAR conclusions about money”, *Carnegie-Rochester Conference Series in Public Policy*, 49, 207-244.

Favero, C., O. Ricchi, C. Tegami (2004), “Forecasting Italian inflation with large datasets and many models”, IGIER Working Paper 269.

Fidrmuc, J., I. Korhonen (2006), “Meta-analysis of the business cycle correlation between the euro area and the CEECs “, *Journal of Comparative Economics*, 34(3), 518-537.

Forni, M., D. Giannone, F. Lippi, L. Reichlin (2005a), “Opening the black box: structural factor models versus structural VARs”, CEPR Discussion Paper 4133.

Forni, M., M. Hallin, M. Lippi, L. Reichlin (2002), “The generalized dynamic factor model: one-sided estimation and forecasting”, CEPR Discussion Paper 3432.

Forni, M., M. Hallin, M. Lippi, L. Reichlin (2000), “The generalized dynamic factor model: identification and estimation”, *The Review of Economic and Statistics*, 82(4), 540-554.

Forni, M., M. Hallin, M. Lippi, L. Reichlin (2003), “Do financial variables help forecasting inflation and real activity in the euro area?”, *Journal of Monetary Economics*, 50, 1243-1255.

Forni, M., M. Hallin, M. Lippi, L. Reichlin (2005b), “The generalized dynamic factor model: one-sided estimation and forecasting”, *Journal of the American Statistical Association*, 100, 830-840.

Forni, M., L. Reichlin (1998), “Let’s get real: a factor analytic approach to disaggregated business cycle analysis”, *Review of Economic Studies*, 65, 453-474.

Forni, M., L. Reichlin (2001), “Federal policies and local economies: Europe and the US”, *European Economic Review*, 45(1), 109-134.

Frankel, J. A., A. K. Rose (1998), “The endogeneity of the optimum currency area criteria”, *The Economic Journal*, 108 (July), 1009-1025.

FRB (2007), “Monetary Policy Report to the Congress”, (July), www.federalreserve.gov/boarddocs/hh/2007/july/fullreport.htm.

Galí, J. (1999), “Technology, employment, and the business cycle: do technology shocks explain aggregate fluctuations?”, *American Economic Review*, 89(1), 249-271.

Gavin, W.T., K.L. Kliesen (2006), “Forecasting inflation and output: comparing data-rich models with simple rules”, Federal Reserve Bank of St. Louis, Working Paper 54A.

German Council of Economic Experts (2001), Annual Report 2001/2002, Verlag Metzler Poeschel, Stuttgart, p. 251-266.

Geweke, J. (1977), “The dynamic factor analysis of economic time series”, in: D.J. Aigner and A.S. Goldberger (eds.), Latent variables in socio-economic models, North-Holland, Amsterdam.

Giacomini, R., H. White (2006), “Tests of conditional predictive ability”, *Econometrica*, 74(6), 1545–1578.

Giannone, D., L. Reichlin (2006), “Trends and cycles in the euro area: how much heterogeneity and should we worry about it”, ECB Working Paper 595.

Giannone, D., L. Reichlin, L. Sala (2005), “Monetary policy in real time”, in: M. Gertler, and K. Rogoff (eds.), NBER Macroeconomics Annual 2004, No. 19, MIT Press.

Giannone, D., L. Reichlin, L. Sala (2002), ”Tracking Greenspan: systematic and unsystematic monetary policy revisited”, CEPR Discussion Paper 3550.

Giannone, D., T. Matheson (2006): “A new core inflation indicator for New Zealand”, Reserve Bank of New Zealand Discussion Paper 10.

Giannoni, M., J. Boivin (2006), “DSGE models in a data-rich environment”, NBER Working Paper 12772.

Gilbert, P.D., E. Meijer (2006), “Money and credit factors”, Bank of Canada Working Paper 2006-3.

Girouard, N., M. Kennedy, P. van den Noord, C. André (2006), “Recent house price developments”, OECD Working Paper 475.

Gosselin, M.-A., G. Tkacz (2001), “Evaluating factor models: an application to forecasting inflation in Canada”, Bank of Canada Working Paper 18.

Grenouilleau, D. (2004), “A sorted leading indicators dynamic (SLID) factor model for short-run euro-area GDP forecasting”, European Commission Economic Papers 219.

Harvey A., D. Bates (2003), “Multivariate unit root tests, stability and convergence”, University of Cambridge, DAE Working Paper 301, University of Cambridge.

Heathcote, J., F. Perri (2004), “Financial globalization and real regionalization” *Journal of Economic Theory*, 119(1), 207-243.

Hofmann, B. (2006): “Do monetary indicators (still) predict euro area inflation?”, ECB Working Paper 18.

Huber, P.J. (1964), “Robust estimation of a location parameter“, *Annals of Mathematical Statistics*, 35(1), 73-101.

Imbs, J. (2004), “Trade, finance, specialization and synchronization”, *Review of Economics and Statistics*, 86(3), 723-734.

IMF (2001a), “World Economic Outlook - October 2001”, p. 65-79.

IMF (2001b), “Monetary and exchange rate policies of the Euro Area”, Country Report 01/60.

IMF (2007), “United States: selected issues”, IMF Country Report 07/265.

Inoue, A., L. Kilian (2007), “How useful is bagging in forecasting economic time series? A case study of US CPI inflation”, *Journal of the American Statistical Association*, forthcoming.

Jeon, Y. (2004), “Combined forecasts, hidden common factors and international linkages”, mimeo, http://www.nd.edu/~meg/MEG2004/Jeon-Yongil.pdf.

Kabundi, A. (2004): “Estimation of Economic Growth in France Using Business Survey Data”, IMF Working Paper 69.

Kalemli-Ozcan, S., Sørensen, B. E., Yosha, O. (2003), “Risk sharing and industrial specialization: regional and international evidence”, *American Economic Review*, 93(3), 903-918(16).

Kapetanios, G., V. Labhard, S. Price (2005), “Forecasting using Bayesian and information theoretic model averaging: an application to UK inflation”, *Journal of Business and Economic Statistics*, forthcoming.

Kapetanios, G. (2004), “A note on modelling core inflation for the UK using a new dynamic factor estimation method and a large disaggregated price index dataset”, *Economics Letters*, 85, 63-69.

Kapetanios, G., M. Marcellino (2004), “A parametric estimation method for dynamic factor models of large dimensions”, Queen Mary University of London Working Paper 489 revised February 2004.

Kilian, L. (1998), “Small-Sample Confidence Intervals for Impulse Response Functions”, *Review of Economics and Statistics*, 80(2) (May), 218-230.

Knell, M., H. Stix (2005a), “The income elasticity of money demand: A meta-analysis of empirical results”, *Journal of Economic Surveys*, 19(3), 213-233.

Knell, M., H. Stix (2005b), “How robust are money demand estimations? A meta-analytic summary of findings about income elasticities”, *Kredit und Kapital*, 38 (4), 515-540.

Knell, M., H. Stix (2006), “Three decades of money demand studies: differences and similarities”, *Applied Economics*, 38, 805-818.

Koop, G., S. Potter (2004), Forecasting in dynamic factor models using Bayesian model averaging”, *Econometrics Journal,* 7, 550-565.

Kose, M. A., C. Otrok, C.H. Whiteman (2003a), “Understanding the evolution of world business cycles”, mimeo. http://www.imf.org/external/np/res/seminars/2002/global/preconf.htm.

Kose, M. A., C. Otrok, C.H. Whiteman (2003b), “International business cycles: world, region and country-specific factors”, *American Economic Review*, 93(4) (September), 1216-1239.

Kose, M. A., E. S. Prasad, M.E. Terrones (2003c), “Volatility and comovement in a globalized world economy: an empirical exploration”, IMF Working Paper WP/03/246.

Kose, M. A., E. S. Prasad, M.E. Terrones (2003d), “How does globalization affect the synchronization of business cycles?”, *American Economic Association Papers and Proceedings*, 93(2) (May), 57-62.

Kose, M. A., K. Yi (2001), “International trade and business cycles: is vertical specialization the missing link?”, *American Economic Association Papers and Proceedings*, 91(2), 371-375.

Lin, J.-L., R.S. Tsay (2005), “Comparison of forecasting methods with many predictors”, mimeo, http://www.atl-res.com/finance/LIN.pdf.

Lipsey, M.W., D. B. Wilson (2000), “Practical meta-analysis”, SAGE Publications Ltd.

Liu, D. (2004), “The effects of monetary policy and macroeconomic forecasting”, mimeo, http://econweb.tamu.edu/dliu/The%20Effects%20of%20Monetary%20and%20Forecast%20oct31.pdf.

Longhi, S., P. Nijkamp, J. Poot (2005), “ A meta-analytic assessment of the effect of immigration on wages”, *Journal of Economic Surveys*, 19(3), 451-477.

Luginbuhl, R., S. J. Koopman (2004), “Convergence in European GDP series: a multivariate common converging trends-cycle decomposition”, *Journal of Applied Econometrics*, 19(5), 661-636.

Lumsdaine, R. L., E. S. Prasad (2003), "Identifying the common component in international economic fluctuations: a new approach", *The Economic Journal*, 113(484), 101-127.

MacKinnon, R. (1963), "Optimum currency areas", *The American Economic Review*, 53(4), 717-725.

Marcellino, M., J. H. Stock, M. Watson (2000), "A dynamic factor analysis of the EMU", mimeo. http://www.igier.uni-bocconi.it/whos.php?vedi=1139&tbn=albero &id_doc=177.

Marcellino, M., J. Stock, M. Watson (2001), "Macroeconomic forecasting in the euro area: country-specific versus area-wide information", IGIER Working Paper 201.

Marcellino, M., J. Stock, M. Watson (2003), "Macroeconomic forecasting in the euro area: country-specific versus area-wide information", *European Economic Review*, 47 ,1-18.

Marcellino, M., J. Stock, M. Watson (2005), "A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series", *Journal of Econometrics*, 135, 499–526.

Matheson, T. D. (2006), "Factor model forecasts for New Zealand", *International Journal of Central Banking*, 2, 169-237.

Mishkin, F.S. (2007), "Housing and the monetary transmission mechanism", FRB Working Paper 2007-40, Finance and Economic Discussion Series.

Monfort, A., J. P. Renne, R. Rüffer, G. Vitale (2004), "Is economic activity in the G7 synchronised? Common shocks vs. spillover effects", CEPR Discussion Paper 4119.

Moser, G., F. Rumler, J. Scharler (2007), "Forecasting Austrian inflation", *Economic Modelling*, 24, 470–480.

Mumtaz, H., P. Surico (2006) "Inflation globalization and the fall of country-specific fluctuations", presented at the SCE conference on Computing in Economics and Finance in 2007 in Montreal.

Mundell, R. A. (1961), "A theory of optimum currency areas", *The American Economic Review* (November).

Norrbin, S. C., D. E. Schlagenhauf (1996), "The role of international factors in the business cycle: a multi-country study", *Journal of International Economics*, 40, 85-104.

Osborn, D. R., P. J. Perez, M. Sensier (2005), "Business cycle linkages for the G7 countries: does the US lead the world?", Centre for Growth and Business Cycle Research/University of Manchester Discussion Paper 050.

Otto G., G. Voss, L. Willard (2001), "Understanding OECD output correlations", Reserve Bank of Australia Research Discussion Paper 2001-05.

Peersman, G. (2005), "What caused the early millennium slowdown? Evidence based on vector autoregressions", *Journal of Applied Econometrics*, 20(2), 185-207.

Peersman, G., F. Smets (2002), "Are the effects of monetary policy in the euro area greater in recessions than in booms?" In: Mahadeve, L., P. Sinclair (Eds.), Monetary transmission in diverse economies, Chapter 2, Cambridge University Press, Cambridge, pp. 28-48.

Peersman, G., R. Straub (2006), "Putting the New Keynesian model to a test", mimeo. http://www.feb.ugent.be/Fineco/publications/gert/peersman-straub2-march2006.pdf.

Pericoli, M., M. Sbracia (2003), "A primer on financial contagion", *Journal of Economic Surveys*, 17(4), 571-608

Pesaran, M. H., T. Schuermann, S.M. Weiner (2004), "Modelling regional interdependencies using a global error-correcting macroeconometric model", *Journal of Business and Economic Statistics*, 22, 129-162.

Reichlin, L. (2003), "Factor models in large cross sections of time series", in: Dewatripont, M., L. Hansen, S. Turnovsky (eds.), Advances in econometrics, theory and applications, Eight World Congress of the Econometric Society, Vol. III.

Roubini, N. (2007), "The risk of a U.S. hard landing and implications for the global economy and financial markets", speech at the IMF seminar on September 13, 2007.

Rousseeuw, P. J., A. M. Leroy (1987), "Robust regression and outlier detection ", Jon Wiley & Sons Inc., New York.

Sala, L. (2003), "Monetary transmission in the euro area: a factor model approach", mimeo; http://www.igier.uni-bocconi.it/whos.php?vedi=1798&tbn=albero&id_doc=177.

Sargent, T.J., C.A. Sims (1977), "Business cycle modelling without pretending to have too much *a prior* economic theory", in: C.A. Sims (ed.), New methods in business research, Federal Reserve Bank of Minneapolis, Minneapolis.

Schneider, M., M. Spitzer (2004), "Forecasting Austrian GDP using the generalized dynamic factor model", OeNB Working Paper 89.

Schumacher, C. (2007), "Forecasting German GDP using alternative factor models based on large dataset", *Journal of Forecasting*, 26, 271-302.

Schumacher, C., J. Breitung (2006), "Real-time forecasting of GDP based on a large factor model with monthly and quarterly data", Bundesbank Discussion Paper, Series 1, 33/2006.

Schumacher, C., C. Dreger (2004), "Estimating large-scale factor models for economic activity in Germany: Do they outperform simpler models?“, *Jahrbücher für Nationalökonomie und Statistik*, 224, 731-750.

Smets, F., R. Wouters (2003), "An estimated stochastic dynamic general equilibrium model of the euro area", *Journal of the European Economic Association*, 1(5), 1123-1175.

Stanley, T. S. (2001), "Wheat from chaff: meta-analysis as quantitative literature review", *Journal of Economic Perspectives*, 15(3), 131-150.

Stavrev, E. (2006), "Measures of underlying inflation in the euro area: assessment and role for informing monetary policy", IMF Working Paper 197.

Stock, J., M. Watson (1998), "Diffusion indexes", NBER Working Paper 6702.

Stock, J., M. Watson (1999), "Forecasting inflation", *Journal of Monetary Economics*, 44, 293-335.

Stock, J., M. Watson (2002a), "Macroeconomic forecasting using diffusion indexes", *Journal of Business and Economic Statistics*, 20(2),147-162.

Stock, J., M. Watson (2002b), "Forecasting using principal components from a large number of predictors", *Journal of the American Statistical Association*, 97, 1167-1179.

Stock, J., M. Watson (2004), "Combination forecasts of output growth in a seven country dataset", *Journal of Forecasting*, 23(6), 405-430.

Stock, J., M. Watson (2005a), "An empirical comparison of methods for forecasting using many predictors", mimeo, http://ksghome.harvard.edu/~JStock/pdf/beb_4.pdf.

Stock, J., M. Watson (2005b), "Understanding changes in international business cycle dynamics", *Journal of the European Economic Association*, 3(5), 966-1006.

Stock, J., M. Watson (2005c), "Implications of dynamic factor models for VAR analysis", NBER Working Paper 11467.

Stock, J., M. Watson (2006), "Forecasting with many predictors", in: Elliott, G., C. W. J. Granger, A. Timmermann (eds.), Handbook of economic forecasting, Vol. 1, 515-554.

Tatiwa Ferreira, R., H. Bierens, I. Castelar (2005), "Forecasting quarterly Brazilian GDP growth rate with linear and nonlinear diffusion index models", EconomiA, Selecta, Brasilia (DF), 6(3), 261-292.

Uhlig, H. (2003), "What moves real GNP?", mimeo. http://eabcn.org/agenda/programme0302.htm.

Uhlig, H. (2005), "What are the effects of monetary policy on output? Results from an agnostic identification procedure", *Journal of Monetary Economics*, 52, 381-419.

Van Nieuwenhuyze, C. (2006), "A generalized dynamic model for the Belgian Economy – useful business cycle indicators and GDP growth forecasts", National Bank of Belgium Working Paper 80.

Watson, M. (2003), "Macroeconomic forecasting using many predictors", in M. Dewatripont, L. Hansen and S. Turnovsky (eds), Advances in economics and econometrics, theory and applications, Eight World Congress of the Econometric Society, 3, 87-115.

Weichselbaumer, D., R. Winter-Ebmer (2005), "A meta-analysis of the international wage gender gap", *Journal of Economic Surveys*, 19(3), 479-511.