Journal of Business & Economic Statistics
Publication details, including instructions for authors and subscription information:
http://amstat.tandfonline.com/loi/ubes20

Combining Disaggregate Forecasts or Combining Disaggregate Information to Forecast an Aggregate

David F. Hendry and Kirstin Hubrich
Department of Economics, Oxford University, Oxford OX1 3UQ, U.K.
Research Department, European Central Bank, Kaiserstrasse 29, D-60311, Frankfurt am Main, Germany

Version of record first published: 01 Jan 2012.

To cite this article: David F. Hendry and Kirstin Hubrich (2011): Combining Disaggregate Forecasts or Combining Disaggregate Information to Forecast an Aggregate, Journal of Business & Economic Statistics, 29:2, 216-227

To link to this article: http://dx.doi.org/10.1198/jbes.2009.07112

Combining Disaggregate Forecasts or Combining Disaggregate Information to Forecast an Aggregate

David F. Hendry
Department of Economics, Oxford University, Oxford OX1 3UQ, U.K. (*david.hendry@nuffield.ox.ac.uk*)

Kirstin Hubrich
Research Department, European Central Bank, Kaiserstrasse 29, D-60311, Frankfurt am Main, Germany (*Kirstin.hubrich@ecb.europa.eu*)

To forecast an aggregate, we propose adding disaggregate variables, instead of combining forecasts of those disaggregates or forecasting by a univariate aggregate model. New analytical results show the effects of changing coefficients, misspecification, estimation uncertainty, and mismeasurement error. Forecast-origin shifts in parameters affect absolute, but not relative, forecast accuracies; misspecification and estimation uncertainty induce forecast-error differences, which variable-selection procedures or dimension reductions can mitigate. In Monte Carlo simulations, different stochastic structures and interdependencies between disaggregates imply that including disaggregate information in the aggregate model improves forecast accuracy. Our theoretical predictions and simulations are corroborated when forecasting aggregate United States inflation pre and post 1984 using disaggregate sectoral data.

KEY WORDS: Aggregate forecasts; Disaggregate information; Forecast combination; Inflation.

1. INTRODUCTION

Forecasts of macroeconomic aggregates are employed by the private sector, governmental and international institutions as well as central banks. There has been renewed interest in the effect of contemporaneous aggregation in forecasting, and the potential improvements in forecast accuracy by forecasting the component indices and aggregating such forecasts, over simply forecasting the aggregate itself (see, e.g., Fair and Shiller 1990 for a related analysis for United States GNP; Zellner and Tobias 2000 for industrialized countries' GDP growth; Marcellino, Stock, and Watson 2003 for disaggregation across euro countries; and Espasa, Senra, and Albacete 2002 and Hubrich 2005 for forecasting euro area inflation). The aggregation of forecasts of disaggregate inflation components is also receiving attention by central banks in the Eurosystem (see, e.g., Benalal et al. 2004; Reijer and Vlaar 2006; Bruneau et al. 2007; and Moser, Rumler, and Scharler 2007). Similarly, for short-term inflation forecasting, staff at the Federal Reserve Board forecast disaggregate price categories (see, e.g., Bernanke 2007).

The theoretical literature shows that aggregating component forecasts is at least as accurate as directly forecasting the aggregate when the data generating process (DGP) is known, and lowers the mean squared forecast error (MSFE), except under certain conditions. If the DGP is not known and the model has to be estimated, the properties of the unknown DGP determine whether combining disaggregate forecasts improves the accuracy of the aggregate forecast. It might be preferable to forecast the aggregate directly. Contributions to the theoretical literature on aggregation versus disaggregation in forecasting can be found in, for example, Grunfeld and Griliches (1960), Kohn (1982), Lütkepohl (1984b, 1987), Granger (1987), Pesaran, Pierse, and Kumar (1989), Garderen, Lee, and Pesaran (2000), and Giacomini and Granger (2004); see Lütkepohl (2006) for a recent review on aggregation and forecasting. Since in practice the DGP is not known, it is largely an empirical question whether aggregating forecasts of disaggregates improves forecast accuracy of the aggregate of interest. Hubrich (2005), for example, shows that aggregating forecasts by component does not necessarily help to forecast year-on-year Eurozone inflation one year ahead.

In this paper, we suggest an alternative use of disaggregate information to forecast the aggregate variable of interest, namely including disaggregate variables in the model for the aggregate. This is distinct from forecasting the disaggregate variables separately and aggregating those forecasts, usually considered in previous literature. An alternative to including disaggregate variables in the aggregate model might be to combine the information in the disaggregate variables first, then include the disaggregate information in the aggregate model. This entails a dimension reduction, potentially leading to reduced estimation uncertainty and reduced MSFE. Bayesian shrinkage methods or factor models can be used for that purpose. We include the latter in our empirical analysis. A third alternative is to forecast the aggregate only using lagged aggregate information. Our analysis is relevant for policy makers and observers interested in inflation forecasts, since disaggregate inflation rates across sectors and regions are often monitored and used to forecast aggregate inflation. Many other applications of our results are possible, including forecasting other macroeconomic aggregates such as GDP growth, monetary aggregates, or trade, since our assumptions in a large part of the analysis are fairly general.

Journal of Business & Economic Statistics
April 2011, Vol. 29, No. 2
DOI: 10.1198/jbes.2009.07112

Our analysis extends previous literature in a number of directions outlined in the following.

First, our proposal of combining disaggregate information by including all, or a selected number of, disaggregate variables in the aggregate model is investigating the predictability content of disaggregates for the aggregate from a new perspective. Most previous literature focused on combining disaggregate forecasts rather than disaggregate information. Second, we present new analytical results for the forecast accuracy comparison of different uses of disaggregate information to forecast the aggregate. In contrast to Hendry and Hubrich (2006) focusing on the predictability of the aggregate in population, we investigate the improvement in forecast accuracy related to the sample information. From the analytical comparisons of the forecast error decompositions of the three different methods for forecasting an aggregate, we draw important conclusions regarding the effects of misspecification, estimation uncertainty, forecast origin mismeasurement, structural breaks, and innovation errors for their relative forecast accuracy. Instabilities have generally been found important for the forecast accuracy of different forecasting methods; see, for example, Stock and Watson (1996, 2007), Clements and Hendry (1998, 1999, 2006), and Clark and McCracken (2006). Therefore, an important extension of the previous literature on contemporaneous aggregation and forecasting is to allow for a DGP with an unknown break in the parameters and time-varying aggregation weights. Third, we investigate by Monte Carlo simulations the effect of the stochastic structure of disaggregates and their interdependencies, as well as structural breaks, estimation uncertainty, and misspecification, on the relative forecast accuracy of the different approaches to forecast the aggregate. Fourth, we examine whether our theoretical predictions can explain our empirical findings for the relative forecast accuracy of combining disaggregate sectoral information versus disaggregate forecasts or just using past aggregate information to forecast aggregate U.S. inflation. Note, that all empirical and simulation results discussed in the text of the current paper that are not presented explicitly in tables or graphs are available from the authors upon request.

The paper is organized as follows. In Section 2, we present new analytical results on the relative forecast accuracy of different approaches to forecast the aggregate. Section 3 presents Monte Carlo evidence. In Section 4, we investigate whether our analytical and simulation results are confirmed in a pseudo out-of-sample forecasting experiment for U.S. CPI inflation. Section 5 concludes.

2. COMBINING DISAGGREGATE FORECASTS OR DISAGGREGATE INFORMATION

In Hendry and Hubrich (2006) we presented analytical results on predictability of aggregates using disaggregates, as a property in population. In the following analytical derivations, we allow the model and the DGP to differ and the parameters need to be estimated. First, we extend previous literature including Lütkepohl (1984a, 1984b), Kohn (1982), and Giacomini and Granger (2004), by allowing for a structural break at the forecast origin and forecast origin uncertainties due to measurement errors. We assume that the break is not known to the forecaster who continues to use the previous forecast model based on in-sample information. It is of interest to know whether and how the relative forecast accuracy of different methods to forecast the aggregate is affected by an unmodeled structural break. Second, we compare our proposed approach of including and combining disaggregate information directly in the aggregate model with previous methods.

Relation to other literature on aggregation and forecasting. Allowing for estimation uncertainty introduces a trade-off between potential biases due to not specifying the fully disaggregated system, and increases in variance due to estimating an unnecessarily large number of parameters. Giacomini and Granger (2004) show that in the presence of estimation uncertainty, to aggregate forecasts from a space-time AR model is weakly more efficient than the aggregate of the forecasts from a VAR. They also show that if their poolability condition is satisfied, that is, zero coefficients on all included components are not rejected, it is more efficient to forecast the aggregate directly. Hernandez-Murillo and Owyang (2006) provide an empirical investigation of the Giacomini and Granger (2004) methodology. The space-time AR model implies certain restrictions on the correlation structure of the disaggregates. In contrast, in our proposal the type of restrictions depends on the model used. For instance, in a VAR our suggestion implies to impose zero restrictions on the coefficients of the disaggregates in the aggregate equation. In contrast to the empirical approach implemented in Carson, Cenesizoglu, and Parker (2007) and Zellner and Tobias (2000), who impose the parameters of all or most of the disaggregates to be identical across the individual empirical models, we suggest imposing either zero restrictions on disaggregate parameters in the aggregate model or imposing a factor structure, where the weights of the disaggregates in the factor maximize their explained variance. Granger (1980, 1987) considers correlations among the disaggregates due to a common factor. Granger (1987) suggests that the forecast of the aggregate is simply the factor component of the disaggregate expectations, so empirically derived disaggregate models may miss important factors, and are therefore misspecified. Our proposed approach extends this idea, formally investigating the effect of misspecification, estimation uncertainty, breaks, forecast-origin mismeasurement, innovation errors, changing weights, and a changing error variance-covariance structure. Another approach, implicit in Carson, Cenesizoglu, and Parker (2007) and Zellner and Tobias (2000), and explicitly analyzed in Hubrich (2005), is to impose the same variable selection across individual disaggregate models. Hubrich (2005) finds that allowing for different model specifications for different disaggregates does not improve forecast accuracy of the aggregate for euro area inflation.

In the following, we present *new analytical results* comparing forecast errors when forecasting the aggregate is the objective, for: (a) combining disaggregate forecasts (Section 2.1); (b) only using past aggregate information (Section 2.2); important conclusions comparing the analytics from (a) and (b) (Section 2.3); and (c) combining disaggregate information (Section 2.4). Unless otherwise stated, the following assumptions hold in this section:

Assumptions. The DGP of the disaggregates is stationary in-sample, but is unknown and has to be estimated. We allow for estimation uncertainty and model misspecification as well as structural change in the mean and slope parameters and measurement error at the forecast origin, unknown to the forecaster.

We also allow for (unknown) changes in aggregation weights out-of-sample. Our analytical derivations analyze the effects of those assumptions on the forecast error for the different methods to forecast an aggregate.

Let $\mathbf{y}_t$ denote the vector of n disaggregate price changes with elements $y_{i,t}$. The DGP for the disaggregates is assumed to be an $\mathsf{I}(0)$ VAR with unknown parameters that are constant in-sample and have to be estimated:

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t \quad \text{for } t = 1, \dots, T, \tag{1}$$

where $\boldsymbol{\epsilon}_t \sim \mathrm{ID}[\mathbf{0}, \boldsymbol{\Omega}]$ (ID: identically distributed), but with a break at the forecast origin T when:

$$\mathbf{y}_{T+h} = \boldsymbol{\mu}^* + \boldsymbol{\Gamma}^*\mathbf{y}_{T+h-1} + \boldsymbol{\epsilon}_{T+h} \quad \text{for } h = 1, \dots, H \tag{2}$$

although the process stays $\mathsf{I}(0)$. This break is assumed to be unknown to the forecaster. Such a putative DGP reflects the prevalence of forecast failure in economics by its changing parameters. Let $y_t^a = \boldsymbol{\omega}_t'\mathbf{y}_t$ be the aggregate price index with weights $\boldsymbol{\omega}_t$, which is the variable of interest.

2.1 Combining Disaggregate Forecasts: New Analytical Results

We first construct a decomposition of all the sources of forecast errors from aggregating the disaggregate forecasts using an estimated version of (1) when the forecast period is determined by (2).

This analysis follows the VAR taxonomy in Clements and Hendry (1998), but we only consider one-step forecasts although allow $\mathbf{y}_T$ to be subject to forecast-origin measurement errors (multistep ahead forecasts add further terms, which we omit for readability). Section 2.2 provides the corresponding taxonomy for the forecast errors from forecasting the aggregate directly from its past. In both cases, in-sample changes in the weights $\boldsymbol{\omega}_t$ make the analysis intractable, so we assume constant weights here, but refer to the implications of changing weights in Section 2.3. The weights are assumed to be positive and to lie in the interval [0, 1].

First, for the aggregated disaggregate forecast, taking expectations in (1) under stationarity—if the DGP is integrated, it must be transformed to a stationary representation—yields:

$$\mathsf{E}[\mathbf{y}_t] = \boldsymbol{\mu} + \boldsymbol{\Gamma}\mathsf{E}[\mathbf{y}_{t-1}] = \boldsymbol{\mu} + \boldsymbol{\Gamma}\boldsymbol{\phi}_y = \boldsymbol{\phi}_y,$$

which is the long-run mean, $\boldsymbol{\phi}_y = (\mathbf{I}_n - \boldsymbol{\Gamma})^{-1}\boldsymbol{\mu}$, referred to as the "equilibrium mean" by Clements and Hendry (1998, 2006), as it is the value to which the process converges in the absence of further shocks. Nevertheless, the long-run mean might shift from a structural break. Further,

$$\mathbf{y}_t - \mathsf{E}[\mathbf{y}_t] = \mathbf{y}_t - \boldsymbol{\phi}_y = \boldsymbol{\Gamma}(\mathbf{y}_{t-1} - \boldsymbol{\phi}_y) + \boldsymbol{\epsilon}_t. \tag{3}$$

(3) represents the deviation of the disaggregates from their long-run mean, which will facilitate the decomposition of the forecast errors below, and isolate terms that only affect the bias.

The forecasts from the estimated disaggregate model (1) at the estimated forecast origin $\hat{\mathbf{y}}_T$ are

$$\hat{\mathbf{y}}_{T+1|T} = \hat{\boldsymbol{\phi}}_y + \hat{\boldsymbol{\Gamma}}(\hat{\mathbf{y}}_T - \hat{\boldsymbol{\phi}}_y), \tag{4}$$

where from $T+1$ onwards, the aggregated 1-step forecast errors $\hat{\boldsymbol{\epsilon}}_{T+1|T} = \mathbf{y}_{T+1} - \hat{\mathbf{y}}_{T+1|T}$ are

$$\boldsymbol{\omega}'\hat{\boldsymbol{\epsilon}}_{T+1|T} = [\boldsymbol{\omega}'\boldsymbol{\phi}_y^* - \boldsymbol{\omega}'\hat{\boldsymbol{\phi}}_y] + \big[\boldsymbol{\omega}'\boldsymbol{\Gamma}^*(\mathbf{y}_T - \boldsymbol{\phi}_y^*) - \boldsymbol{\omega}'\hat{\boldsymbol{\Gamma}}(\hat{\mathbf{y}}_T - \hat{\boldsymbol{\phi}}_y)\big] + \boldsymbol{\omega}'\boldsymbol{\epsilon}_{T+1}. \tag{5}$$

Assuming the relevant moments exist, as is likely here, let $\mathsf{E}[\hat{\boldsymbol{\Gamma}}] = \boldsymbol{\Gamma}_e$ and $\mathsf{E}[\hat{\boldsymbol{\phi}}_y] = \boldsymbol{\phi}_{y,e}$.

The forecast-error taxonomy follows by decomposing each term in (5) into its components, namely, the parameter shifts, parameter misspecifications, and the estimation uncertainty of the parameters. As the DGP is $\mathsf{I}(0)$, the dependence of the estimated parameters on the last observation is $\mathsf{O}_p(T^{-1})$, as can be seen by terminating estimation at $T-1$, so is omitted below (in contrast, see Elliott 2007 for the case of a nonstationary DGP).

$\boldsymbol{\omega}'\hat{\boldsymbol{\Gamma}}(\hat{\mathbf{y}}_T - \mathbf{y}_T)$ can be decomposed (see the Appendix for details), yielding the taxonomy in (6). This forecast-error decomposition facilitates the analysis of the effects of structural change, model misspecification, estimation uncertainty, and forecast-origin mismeasurement, since specific terms involving these each vanish once no structural change, or a correctly specified model, etc., is assumed. Terms with nonzero means only affect the bias of the forecast and are shown in italics.

Aggregated disaggregate forecast-error decomposition

$$\begin{aligned}
&\boldsymbol{\omega}'\hat{\boldsymbol{\epsilon}}_{T+1|T} \\
&\quad = \boldsymbol{\omega}'(\mathbf{I}_n - \boldsymbol{\Gamma}^*)(\boldsymbol{\phi}_y^* - \boldsymbol{\phi}_y) && \text{(ia) \textit{long-run mean change}} \\
&\qquad + \boldsymbol{\omega}'(\boldsymbol{\Gamma}^* - \boldsymbol{\Gamma})(\mathbf{y}_T - \boldsymbol{\phi}_y) && \text{(ib) slope change} \\
&\qquad + \boldsymbol{\omega}'(\mathbf{I}_n - \boldsymbol{\Gamma}_e)(\boldsymbol{\phi}_y - \boldsymbol{\phi}_{y,e}) && \text{(iia) \textit{long-run mean misspecification}} \\
&\qquad + \boldsymbol{\omega}'(\boldsymbol{\Gamma} - \boldsymbol{\Gamma}_e)(\mathbf{y}_T - \boldsymbol{\phi}_y) && \text{(iib) slope misspecification} \\
&\qquad + \boldsymbol{\omega}'(\mathbf{I}_n - \boldsymbol{\Gamma}_e)(\boldsymbol{\phi}_{y,e} - \hat{\boldsymbol{\phi}}_y) && \text{(iiia) long-run mean estimation} \\
&\qquad - \boldsymbol{\omega}'(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}_e)(\mathbf{y}_T - \boldsymbol{\phi}_{y,e}) && \text{(iiib) slope estimation} \\
&\qquad - \boldsymbol{\omega}'\boldsymbol{\Gamma}_e(\hat{\mathbf{y}}_T - \mathbf{y}_T) && \text{(iv) \textit{forecast-origin mismeasurement}} \\
&\qquad + \boldsymbol{\omega}'(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}_e)(\hat{\boldsymbol{\phi}}_y - \boldsymbol{\phi}_{y,e}) && \text{(va) covariance interaction} \\
&\qquad - \boldsymbol{\omega}'(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}_e)(\hat{\mathbf{y}}_T - \mathbf{y}_T) && \text{(vb) mismeasurement interaction} \\
&\qquad + \boldsymbol{\omega}'\boldsymbol{\epsilon}_{T+1} && \text{(vi) innovation error.}
\end{aligned} \tag{6}$$

2.2 Forecasting the Aggregate Directly by Its Past

When forecasting the aggregate directly by its own past alone, and the weights are constant, $\boldsymbol{\omega}_t = \boldsymbol{\omega}$, premultiply (1) by $\boldsymbol{\omega}'$ to derive the aggregate relation:

$$\begin{aligned} y_t^a &= \boldsymbol{\omega}'\boldsymbol{\phi}_y + \boldsymbol{\omega}'\boldsymbol{\Gamma}(\mathbf{y}_{t-1} - \boldsymbol{\phi}_y) + \boldsymbol{\omega}'\boldsymbol{\epsilon}_t \\ &= \tau + \kappa(y_{t-1}^a - \tau) + \nu_t, \end{aligned} \tag{7}$$

where, in the second line, $\tau = \boldsymbol{\omega}'\boldsymbol{\phi}_y$, $y_{t-1}^a = \boldsymbol{\omega}'\mathbf{y}_{t-1}$, and (τ, κ) orthogonalize $(1, (y_{t-1}^a - \boldsymbol{\omega}'\boldsymbol{\phi}_y))$ with respect to ν_t. Hence

$$\nu_t = \boldsymbol{\omega}'(\boldsymbol{\Gamma} - \kappa\mathbf{I}_n)(\mathbf{y}_{t-1} - \boldsymbol{\phi}_y) + \boldsymbol{\omega}'\boldsymbol{\epsilon}_t, \tag{8}$$

where $\kappa = \boldsymbol{\omega}'\mathbf{Q}\boldsymbol{\Gamma}'\boldsymbol{\omega}/\boldsymbol{\omega}'\mathbf{Q}\boldsymbol{\omega}$. $\mathsf{E}[(\mathbf{y}_t - \boldsymbol{\phi}_y)(\mathbf{y}_t - \boldsymbol{\phi}_y)'] = \mathbf{Q}$ and $\mathbf{Q} = \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}' + \boldsymbol{\Omega}$ is the standardized sample second-moment matrix (about means) of the disaggregates $\mathbf{y}_t$.

The taxonomy of the sources of one-step-ahead forecast errors for y_{T+1}^a for a forecast origin at T from (8) highlight the potential gains from adding disaggregates to (7). The forecast-period DGP is (2) and $\boldsymbol{\omega}'\boldsymbol{\phi}_y^* = \tau^*$:

$$y_{T+1}^a = \boldsymbol{\omega}'\boldsymbol{\phi}_y^* + \boldsymbol{\omega}'\boldsymbol{\Gamma}^*(\mathbf{y}_T - \boldsymbol{\phi}_y^*) + \boldsymbol{\omega}'\boldsymbol{\epsilon}_{T+1},$$

with

$$\tilde{y}_{T+1|T}^a = \tilde{\tau} + \tilde{\kappa}(\tilde{y}_T^a - \tilde{\tau}),$$

where $\tilde{y}_T^a = \boldsymbol{\omega}'\hat{\mathbf{y}}_T$, matching (4). Let $\tilde{\nu}_{T+1|T} = y_{T+1}^a - \tilde{y}_{T+1|T}^a$, then in a similar notation to above:

$$\begin{aligned} \tilde{\nu}_{T+1|T} = (\tau^* - \tilde{\tau}) + \boldsymbol{\omega}'\boldsymbol{\Gamma}^*(\mathbf{y}_T - \boldsymbol{\phi}_y^*) \\ - \tilde{\kappa}(\tilde{y}_T^a - \tilde{\tau}) + \boldsymbol{\omega}'\boldsymbol{\epsilon}_{T+1}. \end{aligned} \tag{9}$$

The derivations of the corresponding taxonomy are similar, leading to (10). Where it helps to understand the relationship to (6), we have rewritten terms like $\tau^* - \tau$ as $\boldsymbol{\omega}'(\boldsymbol{\phi}_y^* - \boldsymbol{\phi}_y)$ which highlights the close similarities. Terms in italics letters again denote those that need not be zero under unconditional expectations, but would be zero if no shift in the mean occurred over the forecast period when a well-specified model was used from accurate forecast-origin measurements.

Direct aggregate forecast-error decomposition

$$\begin{aligned} &\tilde{\nu}_{T+1|T} \\ &\quad = \boldsymbol{\omega}'(\mathbf{I}_n - \boldsymbol{\Gamma}^*)(\boldsymbol{\phi}_y^* - \boldsymbol{\phi}_y) \\ &\qquad \text{(Ia) } \textit{long-run mean change} \\ &\qquad + \boldsymbol{\omega}'(\boldsymbol{\Gamma}^* - \boldsymbol{\Gamma})(\mathbf{y}_T - \boldsymbol{\phi}_y) \\ &\qquad \text{(Ib) slope change} \\ &\qquad + (1-\kappa)(\tau - \tau_e) \\ &\qquad \text{(IIa) } \textit{long-run mean misspecification} \\ &\qquad + \boldsymbol{\omega}'(\boldsymbol{\Gamma} - \kappa\mathbf{I}_n)(\mathbf{y}_T - \boldsymbol{\phi}_y) \\ &\qquad \text{(IIb) slope misspecification} \\ &\qquad + (1-\kappa)(\tau_e - \tilde{\tau}) \\ &\qquad \text{(IIIa) long-run mean estimation} \\ &\qquad + (\kappa - \tilde{\kappa})\boldsymbol{\omega}'(\mathbf{y}_T - \boldsymbol{\phi}_y) \\ &\qquad \text{(IIIb) slope estimation} \\ &\qquad - \kappa\boldsymbol{\omega}'(\hat{\mathbf{y}}_T - \mathbf{y}_T) \\ &\qquad \text{(IV) } \textit{forecast-origin mismeasurement} \\ &\qquad + (\kappa - \tilde{\kappa})(\tau - \tilde{\tau}) \\ &\qquad \text{(Va) covariance interaction} \\ &\qquad + (\kappa - \tilde{\kappa})\boldsymbol{\omega}'(\hat{\mathbf{y}}_T - \mathbf{y}_T) \\ &\qquad \text{(Vb) mismeasurement interaction} \\ &\qquad + \boldsymbol{\omega}'\boldsymbol{\epsilon}_{T+1} \\ &\qquad \text{(VI) innovation error.} \end{aligned} \tag{10}$$

2.3 Comparing Aggregated Disaggregate versus Aggregate Forecast Errors

Seven important conclusions follow from comparing the forecast errors from combining disaggregate forecasts in Equation (6), with forecast errors from forecasting the aggregate directly as in (10):

1. (Ia) is identical to (ia). This implies that unkown forecast-origin location shifts affect both methods of forecasting the aggregate in precisely the same way. This is an important and surprising result, since no matter how the long-run (or unconditional) means of the disaggregates shift, the two approaches suffer identically in terms of forecast accuracy. Therefore, *relative forecast accuracy* is not affected, while *absolute forecast accuracy* is affected. In contrast to previous literature on forecast combination of different forecast models for the same variable (see, e.g., Clements and Hendry 2004), there is no benefit in MSFE terms in the presence of unknown (and therefore unmodeled) forecast origin location shifts from combining disaggregate forecasts to forecast an aggregate.
2. Comparing (Ib) and (ib) shows that unknown slope changes at the forecast origin also do not affect the relative forecast accuracy of the different forecasting methods of the aggregate. Therefore, there are no gains or losses from aggregating disaggregates or directly forecasting the aggregate in the presence of all forms of unkown breaks at the forecast origin—this is a *relative comparison*, since both approaches can be greatly *affected absolutely* by such breaks.
3. Long-run mean misspecification in (IIa) and (iia) is unlikely in both taxonomies when the in-sample DGP is constant and the model is well specified.
4. The innovation error effects in (VI) and (vi) are also identical in population, irrespective of the covariance structure of the errors, and even if that were also to change at the forecast origin.
5. The impacts of forecast-origin mismeasurement in (iv) and (IV), namely $\boldsymbol{\omega}'\boldsymbol{\Gamma}_e(\hat{\mathbf{y}}_T - \mathbf{y}_T)$ versus $\kappa\boldsymbol{\omega}'(\hat{\mathbf{y}}_T - \mathbf{y}_T)$, are primarily determined by the relative slope misspecifications. In the empirical analysis in Section 4, we use factor models to deal with potential measurement errors.
6. The interaction terms (Va, Vb) and (va, vb) will be small due to the specification in terms of the long-run mean and its deviation therefrom. Therefore, the covariance interaction terms (Va) and (va) have zero mean. Also, terminating estimation at $T-1$ at a cost of $\mathsf{O}_p(T^{-1})$ should

induce a zero mean of the mismeasurement interaction terms (Vb) and (vb).

7. Thus, we conclude that slope misspecification (IIb) and (iib) and estimation uncertainty [(IIIa, IIIb) and (iiia, iiib)] are the primary sources of forecast error *differences* between these two approaches to forecast an aggregate. Mean and slope misspecification only affect the conditional expectations, and in practice depend on how close the aggregate model approximates the true DGP relative to the disaggregate model. Estimation uncertainty only affects the conditional variances and so depends on their respective data second moments (and hence on Ω). Thus, it is not possible to make general statements about whether differences in forecast accuracy are mainly due to the bias or variance of the forecast.

All our conclusions will remain true for small changes in the weights $\boldsymbol{\omega}_T$ over the forecast horizon. Changes of weights, or incorrect forecasts thereof, $(\boldsymbol{\omega}_{T+1} - \hat{\boldsymbol{\omega}}_{T+1})$, are additional sources of error. We leave more detailed investigation of that issue for future research.

2.4 Combining Disaggregate Information to Forecast the Aggregate: Variable Selection or Dimension Reduction

An alternative to the two methods for forecasting an aggregate considered in the previous sections, is to include disaggregate variables in the aggregate model. Since the DGP for the disaggregates is (1), from (7) and (8):

$$\begin{aligned} y_t^a &= \tau + \rho(y_{t-1}^a - \tau) + \boldsymbol{\omega}'(\boldsymbol{\Gamma} - \rho\mathbf{I}_n)(\mathbf{y}_{t-1} - \boldsymbol{\phi}_y) + \boldsymbol{\omega}'\boldsymbol{\epsilon}_t \\ &= \tau + \rho(y_{t-1}^a - \tau) + \sum_{i=1}^{n-1} \pi_i(y_{i,t-1} - \phi_{y,i}) + \nu_t \end{aligned} \tag{11}$$

say, where $\boldsymbol{\pi}' = \boldsymbol{\omega}'(\boldsymbol{\Gamma} - \rho\mathbf{I}_n)$ and ρ is the resulting autoregressive coefficient. In (11), the aggregate y_t^a depends on lags of the aggregate, y_{t-1}^a, and the lagged disaggregates $y_{i,t-1}$. Thus, if the DGP is (1) at the level of the components, an aggregate model could be systematically improved by adding disaggregates to the extent that $\pi_i \neq \mathbf{0}$. This could be ascertained by an F-test. Here, (11) is correctly specified given (1), so dropping $\mathbf{y}_{t-1}$ would induce dynamic misspecification when $\boldsymbol{\pi} \neq \mathbf{0}$, but offers a trade-off of misspecification versus estimation variation. If no disaggregates are included, so forecasting the aggregate by its past, the forecast error decomposition (10) applies. If all but one disaggregate variables are added to the aggregate model, the combined disaggregate model is recovered, so the forecast error taxonomy (6) applies. Our combination of disaggregate variables will inherit the common effects in Section 2.3, but will differ in terms of slope misspecification and estimation uncertainty. Selection of a subset of the most relevant disaggregates to add to the model might help improve forecast accuracy of the aggregate, largely by reducing estimation uncertainty. The forecast accuracy improvement depends on the explanatory power of the disaggregates in an R^2 sense. Consequently, our proposal could dominate both previous alternatives depending on the relative importance of the disaggregates. By including only one or a few disaggregates in the aggregate model, we impose restrictions on the large VAR which includes the aggregate and all but one disaggregate components. These restrictions can improve forecast efficiency, which should translate to MSFE improvement. The trade-off here is between a reduction in forecast error variance due to reduced estimation uncertainty, on the one hand, and increased bias due to potential (slope) misspecification, on the other. This results in a classical forecast model selection problem.

An alternative to including disaggregate variables in the aggregate model is to combine the information contained in the disaggregate variables first, and include this combined information in the aggregate. Relevant methods include factor models or shrinkage methods, entailing a dimension reduction which could lead to reduced estimation uncertainty.

3. MONTE CARLO SIMULATIONS

The simulation experiments are designed to compare forecasts of an aggregate by combining disaggregate and/or aggregate information with those based on combining disaggregate forecasts in small samples, when the orders and coefficients of the DGPs are unknown. Lütkepohl (1984b, 1987) and Giacomini and Granger (2004) present small-sample simulations on the effect of contemporaneous aggregation in forecasting. We complement and extend their simulations by including different DGPs, presenting results on our proposed method to forecast the aggregate, and allowing for a change in the parameters of the DGP.

3.1 Simulation Design

Constant Parameter DGPs. We construct two-dimensional and four-dimensional DGPs with different parameter values based on the following general structure:

$$\begin{pmatrix} 1+\gamma_{11}L & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{21} & 1+\gamma_{22}L & \gamma_{23} & \gamma_{24} \\ \gamma_{31} & \gamma_{32} & 1+\gamma_{33}L & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & 1+\gamma_{44}L \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \nu_{1,t} \\ \nu_{2,t} \\ \nu_{3,t} \\ \nu_{4,t} \end{pmatrix}, \tag{12}$$

where L is the backshift operator, γ_{ij} are the coefficients, and $y_1, \ldots, y_4$ are the disaggregates. The $\nu_{1,t}$ to $\nu_{4,t}$ are IN[0, 1] random numbers, where $\boldsymbol{\Sigma}_\nu = \mathbf{I}_4$. Table 1 summarizes the different DGPs employed in the simulations with constant parameter DGPs.

For DGP$_1$, the parameters in (12) are $\gamma_{11} = \gamma_{22} = 0.5$, $\gamma_{33} = \gamma_{44} = 0$ and $\gamma_{ij} = 0$ for $i, j = 1, \ldots, 4$ with $i \neq j$, so $(1 + 0.5L)y_t^a = 2 + \nu_t$ with $\sigma_\nu^2 = 2$ for the aggregate process. The eigenvalues of the dynamics in DGP$_1$ are equal, and the disaggregates y_1, y_2, and the aggregate y^a all follow an AR(1) process, so slope misspecification will have a small effect on the relative forecast accuracy. This is the first DGP used in Lütkepohl (1984b).

In DGP$_1$, the direct forecast of the aggregate and aggregating the disaggregate forecasts yield the same MSFE, since the components of the disaggregate multivariate process are independent and have identical stochastic structure. When the true

Table 1. Structure of DGPs for MC simulations: summary table

DGP	Disaggregate coefficients	True aggregate DGP	Eigenvalues
DGP_1	$\gamma_{11} = \gamma_{22} = 0.5$ and $\gamma_{12} = \gamma_{21} = 0$	AR(1)	$[-0.5, -0.5]$
DGP_2	$\gamma_{11} = -0.5, \gamma_{22} = -0.3, \gamma_{12} = 0.6, \gamma_{21} = 0.4$	ARMA(2, 1)	$[0.9, -0.1]$
DGP_3	$\gamma_{ii} = -0.5$ for $i = 1, \dots, 4$, $\gamma_{ij} = 0$ for $i, j = 1, \dots, 4$ and $i \neq j$	AR(1)	$[0.5, 0.5, 0.5, 0.5]$
DGP_4	$\gamma_{11} = -0.5, \gamma_{ii} = -0.3$ for $i = 2, 3, 4$, $\gamma_{12} = 0.6, \gamma_{21} = 0.4, \gamma_{ij} = 0$ for all other $i \neq j$	ARMA(4, 3)	$[0.9, -0.1, 0.3, 0.3]$

NOTE: DGP_1, DGP_2, and DGP_3, DGP_4 are two-dimensional and four-dimensional, respectively; population variances are $\sigma_{v_{i,t}} = 1$, $i = 1, \dots, 4$.

model is used for estimation, the MSFE differences result only from estimation uncertainty, and not from model misspecification. Therefore, we can isolate the effect of the estimation uncertainty. The four-dimensional DGP_3 is constructed in a similar way. DGP_2 differs from DGP_1 due to the mutual dependence of the disaggregates. Finally, we construct DGP_4 to approximate our empirical example of U.S. aggregate inflation: Two components are interdependent, whereas the two others behave quite differently.

The simulations were carried out based on $N = 1000$ repetitions. Additional simulations for other DGPs and different sample sizes did not change the qualitative conclusions. In the paper, we only consider results for $T = 100$ for all DGPs. All four DGPs are stationary in-sample. In DGP_1 and DGP_3, the aggregate process is an AR(p) model, in contrast to DGP_2 and DGP_4 where the DGPs of the aggregate are an ARMA(2, 1) and ARMA(4, 3) respectively. Consequently, in DGP_1 and DGP_3, the direct autoregressive (AR) forecast has higher accuracy relative to the other methods, in contrast to DGP_2 and DGP_4 where the AR(p) model is misspecified. DGP_1 and DGP_3 have a factor structure, where the factor is y^a with equal weights for the disaggregates.

As in Lütkepohl (1984b), we generate forecasts from independent samples. Possible extensions are to estimate the models recursively, or from a rolling sample. Results based on recursively expanding samples for DGP_1 and DGP_2 for an initial estimation sample of $R = 100$ ($R = 200$) and out-of-sample period of length $P = 40$ ($P = 100$) did not change the ranking of the different methods to forecast the aggregate and resulted in similar root MSFEs (RMSFEs) to independent samples (all additional results available on request).

Forecast Methods. We compare five different methods to forecast the aggregate:

1. direct forecast only using past aggregate information based on an AR model;
2. forecasting disaggregates with an AR model and aggregating those forecasts (indirect AR);
3. forecasting disaggregates with a VAR including all subcomponents, but no aggregate information, then aggregating those forecasts (indirect VAR^{sub});
4. forecasting disaggregates with a VAR including the aggregate and all subcomponents (except one to avoid collinearity when weights are constant) (direct $\text{VAR}^{agg,\text{sub}}$);
5. forecasting with a VAR including selected subcomponents y_i and the aggregate (direct VAR^{agg,y_i}).

All models are estimated by (multivariate) least squares, providing identical estimators to maximum likelihood under a normality assumption. All simulations assume constant aggregation weights, and are carried out using AIC for selecting the order of the model, since a model selection criterion would be employed in practice when the DGP is unknown.

Simulation With a Nonconstant DGP. To check analytical conclusions 1, 2, and 4 in Section 2.3 in small samples, we implemented a change in the mean and in the innovation variance as well as allowing for nonzero cross-correlations in the innovations in DGP_1 and DGP_2. The change in mean is implemented by changing the intercepts of both subcomponents over the out-of-sample period (a) in the same direction for both components, and (b) in opposite directions. The change in variance is implemented by a change in the variance of the innovation errors over the out-of-sample period. We also carry out simulations for DGPs with different innovation variance or with different cross-correlations of the errors for the entire sample period, including both in-sample and out-of-sample. In the latter experiments we allow for positive as well as negative correlations. Throughout, we investigate the impacts of the changes on the relative rankings of the different methods. For comparability, we consider $T = 100$, independent samples, and AIC is used for lag-order selection.

3.2 Simulation Results

Constant Parameter DGP. The results are presented in Tables 2 and 3 in terms of RMSFE relative to the direct AR benchmark model. Only for the direct AR benchmark actual RMSFEs are presented. Table 2 shows that for DGP_1, the direct forecast of the aggregate based only on aggregate information is best for a 1-step ahead horizon, while the indirect forecast of the aggregate using AR models for the component forecasts is ranked second. The VAR based forecast is worst for this particular DGP. The direct and indirect VAR models provide the same RMSFE because including one disaggregate component in the aggregate model when the DGP is two-dimensional is just a linear transformation of aggregating the disaggregate forecasts (see Section 2.4). The simulation results for DGP_1 are comparable to Lütkepohl (1987, table 5.2).

Investigating the RMSFE for all horizons between $h = 1$ and 12 showed that the differences for horizons larger than 3 were minor, in line with the results in Lütkepohl (1984b, 1987), who only presents results for $h = 1$ and $h = 5$. At forecast horizon $h = 12$, all forecasts are almost identical. At larger T (200

Table 2. Relative RMSFE for DGP_1, DGP_2 and DGP_3, $T = 100$

Horizon:	1		6		12	
Method:	Direct	Indirect	Direct	Indirect	Direct	Indirect
DGP_1						
AR	**1.408**	1.010	**1.634**	1.006	**1.656**	0.999
VAR^{sub}		1.011		1.002		1.001
$VAR^{agg,sub(y_1)}$	1.011		1.002		1.001	
DGP_2						
AR	**1.524**	1.113	**1.547**	1.061	**1.560**	1.035
VAR^{sub}		0.935		0.974		0.990
$VAR^{agg,sub(y_1)}$	0.935		0.974		0.990	
DGP_3						
AR	**2.026**	1.005	**2.314**	1.005	**2.422**	1.001
VAR^{sub}		1.018		0.998		0.998
VAR^{agg,y_3}	1.003		1.001		0.999	
VAR^{agg,y_2,y_3}	1.008		0.996		0.998	

NOTE: Actual RMSFE for AR model (direct) in bold. Superscripts indicate model. VAR^{sub}: VAR only including subcomponents; $VAR^{agg,sub(y_i)}$: VAR with aggregate and subcomponent y_i. Lag order selection for all models by Akaike criterion. $N = 1000$. See Table 1 for the DGPs.

and 400, not presented), the RMSFEs of the direct and indirect forecasts of the aggregate are closer: the DGP implies equal population MSEs, so a larger T leads to a decline in both estimation uncertainty and lag-order selection mistakes, and therefore higher forecast accuracy.

Table 2, second panel, shows that for DGP_2, in contrast to DGP_1, the VAR forecasts are most accurate and the direct AR forecast is second best. Even though that DGP is stationary, the two eigenvalues are substantially different. For DGP_2, including disaggregate information in the aggregate VAR model or forecasting the disaggregates from a VAR and aggregating their forecasts improves forecast accuracy over the other methods.

The simulation results for the four-dimensional DGP_3, with independent components that have the same stochastic structure (as in DGP_1), show that the direct AR forecast is again most accurate (as for DGP_1), but including just one disaggregate is second best for $h = 1$ (see Table 2, third panel). For DGP_4, instead, where the disaggregates are interdependent and follow different stochastic processes, Table 3 shows that including disaggregates in the aggregate model improves over the direct AR forecast. The indirect VAR^{sub} provides more accurate forecasts than the direct or indirect AR forecast for $h = 1$.

Nonconstant Parameter DGP. The simulations investigating the effects of a change in the mean and in the variance of the disaggregates on the relative forecast accuracy ranking of the different methods, yielded the following results for a one-step forecast horizon. First, a change in mean does not change the ranking of the different methods, whether the intercepts in the disaggregate components change in the same or in the opposite direction. This confirms our analytical results in conclusion 1 in Section 2.3. Second, a change in the error variance of the disaggregate components out-of-sample for DGP_1 still leads to the same ranking of the different methods, with the AR direct having highest forecast accuracy. For DGP_2 we get an unchanged ranking of the different methods. Third, changing the variances over the entire sample period, in-sample and out-of-sample, again does not alter rankings for DGP_2, while for DGP_1, all the RMSFEs are very close. Fourth, allowing for cross-correlations between innovation errors (instead of zero cross-correlations) alters the forecast accuracy ranking for

Table 3. Relative RMSFE for DGP_4, $T = 100$

Horizon:	1		6		12	
Method:	Direct	Indirect	Direct	Indirect	Direct	Indirect
AR	**2.134**	1.058	**2.163**	0.999	**2.214**	1.016
VAR^{sub}		0.962		0.979		1.000
VAR^{agg,y_1}	0.987		0.982		1.000	
VAR^{agg,y_2}	0.962		0.981		1.000	
VAR^{agg,y_3}	0.996		0.998		0.999	
VAR^{agg,y_1,y_2}	0.956		0.977		1.000	
VAR^{agg,y_1,y_3}	0.979		0.983		0.999	
VAR^{agg,y_2,y_3}	0.957		0.981		1.000	

NOTE: Actual RMSFE for AR model (direct) in bold. Superscripts indicate model. VAR^{sub}: VAR only including subcomponents; $VAR^{agg,sub(y_i)}$: VAR with aggregate and subcomponent y_i. Lag order selection for all models by Akaike criterion. $N = 1000$. See Table 1 for the DGP.

DGP_1, but not for DGP_2. Our analytical results show that the error covariance structure *per se* does not affect the rankings of the different methods directly, but it does affect it indirectly through estimation uncertainty, as pointed out in conclusion 7 in Section 2.3. Our simulation results confirm that in small samples the error covariance structure can affect the relative estimation uncertainty substantively.

Summary. Overall, including disaggregate variables in the aggregate model helps forecast the aggregate if the disaggregates follow different stochastic structures and are interdependent. The differences in forecast accuracy are less pronounced for higher horizons, since all the forecasts converge to the unconditional mean. In particular, we find that selecting disaggregates helps to improve forecast accuracy by reducing estimation uncertainty if the number of disaggregates is relatively large.

4. FORECASTING AGGREGATE U.S. INFLATION

In this section, we analyze empirically the relative forecast accuracy of the three methods to forecast the aggregate, investigated analytically and via Monte Carlo simulations in the previous sections, for forecasting aggregate U.S. CPI inflation.

Relation to other empirical studies of contemporaneous aggregation and forecasting. The Introduction discusses the large empirical literature on contemporaneous aggregation and forecasting, and the mixture of outcomes reported as to whether aggregation of component forecasts or forecasting the aggregate from past aggregate information alone provides the most accurate forecasts for aggregate inflation. For euro area countries, or the euro area as a whole, the results depend on the country analyzed, whether aggregation is considered across countries or disaggregate components, the forecasting methods or model selection procedures employed, the particular sample periods examined (e.g. before and after EMU), and the forecast horizons considered (see, e.g., Marcellino, Stock, and Watson 2003; Benalal et al. 2004; Hubrich 2005, and Bruneau et al. 2007).

For real U.S. GNP growth, Fair and Shiller (1990) find that disaggregate information helps forecast the aggregate, and Zellner and Tobias (2000), find for forecasting median GDP annual growth rates of 18 industrialized countries, that forecasts of the aggregate can be improved by aggregating disaggregate forecasts, provided an aggregate variable is included in the disaggregate model and all coefficients are restricted to be the same across countries.

We now consider empirically two very different sample periods for U.S. inflation (see, e.g., Atkeson and Ohanian 2001 and Stock and Watson 2007 for recent contributions to predictability changes in U.S. inflation). We investigate whether changes in aggregate U.S. inflation and its components over those different sample periods affects whether disaggregate information helps forecast the aggregate.

4.1 Data

The data employed in this study include the all items U.S. consumer price index (CPI) as well as its breakdown into four subcomponents: food (p^f), commodities less food and energy commodities (p^c), energy (p^e) and services less energy services prices (p^s) (Source: CPI-U for all Urban Consumers, Bureau of Labor Statistics). We employ monthly, seasonally adjusted data, except for CPI energy which does not exhibit a seasonal pattern. Seasonal adjustment by the BLS is based on X-12-ARIMA. We do not consider a real-time dataset, since revisions to the CPI index are extremely small. We consider a sample period for inflation from 1960(1) to 2004(12), where earlier data from 1959(1) onwards are used for the transformation of the price level. As observed by other authors (e.g., Stock and Watson 2007), there has been a substantial change in the mean and the volatility of aggregate inflation between the two samples. We document that the disaggregate components also exhibit a substantial change in mean and volatility. Aggregate as well as components of inflation, all exhibit high and volatile behavior until the beginning or mid-1980s and lower, more stable rates thereafter (see Table 4 for details).

In Sections 4.2 and 4.3, we present results of an out-of-sample experiment for two different forecast evaluation periods: 1970(1)–1983(12) and 1984(1)–2004(12). The date 1984 for splitting the sample coincides with estimates of the beginning of the great moderation, and is in line with what is chosen in Stock and Watson (2007) and Atkeson and Ohanian (2001). We use the same split sample for comparability of our results to those studies in terms of aggregate inflation forecasts.

Due to the mixed results of ADF unit-root tests for different CPI components and samples, we carry out the forecast accuracy comparisons for the level and the change in inflation. We present the results for the level of inflation, as results for the changes in inflation do not differ qualitatively from those for the level in terms of relative forecast accuracy of the different methods. We evaluate the 1-month-ahead and 12-month-ahead forecasts on the basis of the same forecast origin. The main criterion for the comparison of the forecasts here, as in a large part of the literature on forecasting, is RMSFE.

4.2 Combining Disaggregate Forecasts or Disaggregate Variables: AR and VAR Models

Forecasting Methods. We employed various forecasting methods, with different model selection procedures for both direct and indirect forecasts (forecasting inflation directly versus

Table 4. U.S., descriptive statistics, year-on-year CPI inflation

	All items	Energy	Commodities	Food	Services
1960–1983					
Mean	4.86	5.91	3.80	4.75	5.81
Std. deviation	3.41	8.17	2.89	4.11	3.40
1984–2004					
Mean	2.99	2.28	1.43	2.93	3.91
Std. deviation	1.06	8.26	1.65	1.26	0.99

Table 5. Relative RMSFE, U.S. year-on-year inflation (percentage points), 1970–1983

Horizon:	1		6		12	
Method:	Direct $\Delta_{12}\hat{p}^{agg}$	Indirect $\Delta_{12}\hat{p}^{agg}_{\text{sub}}$	Direct $\Delta_{12}\hat{p}^{agg}$	Indirect $\Delta_{12}\hat{p}^{agg}_{\text{sub}}$	Direct $\Delta_{12}\hat{p}^{agg}$	Indirect $\Delta_{12}\hat{p}^{agg}_{\text{sub}}$
AR	**0.294**	1.337	**1.358**	1.083	**2.985**	1.324
RW	1.031	1.378	1.053	1.048	1.045	1.061
MA(1)	1.395	1.198	1.899	1.828	1.695	1.318
VAR^{sub}		1.450		1.241		1.429
$\text{VAR}^{agg,\text{sub}}$	1.071	1.468	1.129	1.225	1.254	1.437
$\text{VAR}^{agg,f}$	1.046		0.992		0.936	
$\text{VAR}^{agg,c}$	1.017		0.991		0.974	
$\text{VAR}^{agg,s}$	1.027		0.962		0.939	
$\text{VAR}^{agg,e}$	1.028		1.065		1.180	

NOTE: Actual RMSFE (nonannualized) for direct AR model in percentage points in bold, for other models RMSFE relative to direct AR; recursive estimation samples 1960(1) to 1970(1), ..., 1983(12); lag order selection for all models [except MA(1) model with one lag] by Akaike criterion, maximum number of lags: $p = 13$; superscripts indicate model, VAR^{sub}: VAR only including subcomponents; $\text{VAR}^{agg,\text{sub}}$: VAR with aggregate and subcomponents; "direct": direct forecast of the aggregate, "indirect": aggregated subcomponent forecast.

aggregating subcomponent forecasts). Tables 5 and 6 present the comparisons of forecast accuracy measured in terms of RMSFE of year-on-year (headline) U.S. inflation for forecasting aggregate (all items) inflation using different approaches to forecast an aggregate.

The forecasting models include: (1) a simple autoregressive (AR) model; (2) the random walk (RW) implemented as inflation in $T + h$ being the simple average of the month-on-month inflation rate from $T - 12$ to T, as used in Stock and Watson (2007) referring to Atkeson and Ohanian (2001); (3) a subcomponent VAR^{sub} to indirectly forecast the aggregate by aggregating subcomponent forecasts; (4) VARs including the aggregate and all disaggregate components (perfect collinearity between aggregate and components does not occur due to annually changing weights in price indices), or a selected number of disaggregate components, $\text{VAR}^{agg,\text{sub}}$ and $\text{VAR}^{agg,\text{sub}_i}$; and (5) an MA(1) (as used in Stock and Watson 2007). Results for factor models are presented in the next section. Model selection procedures selecting the lag length in the various models employed above include the Schwarz (SIC) and the Akaike (AIC) criterion, respectively, with maximum lag order of 13. We find that the AIC-based models generally perform better for U.S. inflation and therefore present results for those models.

The benchmark model for the comparison is the (direct) forecast of aggregate inflation from the AR model, simply forecasting aggregate inflation from its own past (first entry in column labeled "direct" in Tables 5 and 6). This is compared to the indirect forecast from the AR model, that is, the aggregated AR forecasts of the subindices, as well as to the other methods of forecasting the aggregate directly (column labeled "direct") or indirectly (column labeled "indirect") using VARs (see above).

The combination of the disaggregate forecasts for all models is implemented by replicating the aggregation procedure employed by the BLS for the CPI disaggregate data. The data are aggregated in levels, taking into account the respective base year of the weights. Historical aggregation weights were provided to the authors by the BLS. For the aggregation of the forecasts, the current aggregation weights are used, since future weights would not be known to the forecaster in real time.

$\Delta_{12}\hat{p}^{agg}$ and $\Delta_{12}\hat{p}^{agg}_{\text{sub}}$ indicate that the forecast is evaluated on the basis of year-on-year inflation. The models are, however, specified in terms of month-on-month inflation. It should be noted that the ranking of the different forecast methods is

Table 6. Relative RMSFE, U.S. year-on-year inflation (percentage points), 1984–2004

Horizon:	1		6		12	
Method:	Direct $\Delta_{12}\hat{p}^{agg}$	Indirect $\Delta_{12}\hat{p}^{agg}_{\text{sub}}$	Direct $\Delta_{12}\hat{p}^{agg}$	Indirect $\Delta_{12}\hat{p}^{agg}_{\text{sub}}$	Direct $\Delta_{12}\hat{p}^{agg}$	Indirect $\Delta_{12}\hat{p}^{agg}_{\text{sub}}$
AR	**0.190**	1.528	**0.685**	1.024	**1.261**	1.021
RW	1.000	1.617	0.994	1.095	0.955	0.997
MA(1)	1.037	1.508	1.129	1.134	1.116	1.021
VAR^{sub}		1.627		1.155		1.102
$\text{VAR}^{agg,\text{sub}}$	1.044	1.610	1.107	1.179	1.074	1.111
$\text{VAR}^{agg,f}$	0.995		0.903		0.871	
$\text{VAR}^{agg,c}$	1.053		1.091		1.078	
$\text{VAR}^{agg,s}$	1.037		1.120		1.177	
$\text{VAR}^{agg,e}$	1.048		1.173		1.201	

NOTE: As Table 5, but recursive estimation samples 1960(1) to 1984(1), ..., 2004(12).

not invariant to the selected transformations (see, e.g., Clements and Hendry 1998, p. 68). We found that models formulated in terms of year-on-year inflation provided the same ranking and less accurate forecasts than those for monthly changes in inflation evaluated at year-on-year inflation. Iterative multistep ahead forecasts are based on the following model (only including one lag of inflation and no other macroeconomic variables as predictors for expositional purposes): $\hat{\pi}_{T+h} = \hat{\alpha}\sum_{i=0}^{h-1}\hat{\beta}^i + \hat{\beta}^h\pi_T$, where inflation π_t is specified in first differences as $(P_t - P_{t-1})/P_{t-1}$. Results for the change in inflation were not qualitatively different from the results for the level of inflation. In the tables values below unity for the relative RMSFE indicate an improvement in that forecast over the direct AR forecast.

Results. The RMSFE results indicate, first, that the direct forecast is generally more accurate than the indirect forecast of the aggregate, irrespective of whether disaggregate information is included in the aggregate model or not. Second, for the high inflation sample in the 1970s, including one disaggregate in the aggregate model might improve over the direct AR model forecast for longer horizons as well as over the MA(1). The MA(1) is less accurate than the AR(p) in the first sample period and similar to it in the second (see Stock and Watson 2007, who analyze four different price measures, for similar results for quarterly CPI inflation). Including disaggregate variables in most cases also dominates combining disaggregate forecasts in RMSFE terms. For the latter sample 1984–2004, including food inflation in the aggregate model improves forecast accuracy over the direct AR model for all horizons. Interestingly, including food inflation in the aggregate model also improves over the RW model that performs better in RMSFE terms for the second sample period for a one-year horizon. It also improves over the MA(1) for all horizons. We apply the Clark and West (2007) test of equal forecast accuracy for the food inflation model against an AR benchmark for a horizon of one month, and find that this RMSFE improvement is significant at the 10% level. It should be noted, however, that the improvement is not significant when using appropriate critical values for testing a set of four models including different disaggregates against the benchmark AR model for the aggregate (see Hubrich and West 2010, also for similar results for other macroeconomic regressors). Overall, the results suggest that variable selection is important in reducing the impact of parameter uncertainty here.

4.3 Disaggregate Information in Dynamic Factor Models

We now compare combining disaggregate information by including factors estimated from the disaggregate components in the aggregate model with forecasting the aggregate by the benchmark AR model. The analytical investigation showed estimation uncertainty to be an important determinant of the relative forecast accuracy of the different methods to forecast an aggregate. Factor models can reduce estimation uncertainty in comparison with a VAR with many parameters. We employ factor models averaging away idiosyncratic variation in the disaggregate series, and include the factors, estimated by principal components from disaggregate price information, in the aggregate model.

Under the assumptions in Stock and Watson (2002a, 2002b) the model is identified and the factors and loadings can be estimated. Related studies of approximate factor models have shown consistency of principal components estimators of the factor space, for example, Bai (2003), Bai and Ng (2002), and Forni et al. (2000, 2005). Treatments of classical factor models when the cross-sectional dimension n is small can be found in, for example, Anderson (1984), Geweke (1977), Sargent and Sims (1977), and Stock and Watson (1991). A larger cross-section relative to T improves asymptotic performance, in that consistency is achieved at a faster rate compared to a small cross-section (see Stock and Watson 1998). To keep our information set comparable with that in the forecast experiments with VAR models, we retained the same disaggregate variables.

Little is known so far how the size and the composition of the data affect the factor estimates (see, e.g., Boivin and Ng 2006). We are concerned with how factors from disaggregate information affect forecast accuracy of the aggregate economic variable. Since the models considered here are more parsimonious than many VARs considered above, forecast accuracy may be less affected by estimation uncertainty.

The results from the factor analysis are not directly comparable across all horizons with previous tables except for $h = 12$, since here direct multistep ahead forecasts are carried out and forecast accuracy is evaluated for annualized inflation in line with Stock and Watson (1999, 2007) (instead of year-on-year inflation as above). We compute the direct h-step factor forecasts and single predictor forecasts, and consider forecast combinations of all single predictor models based on the respective disaggregate component with equal weights. The results are presented in Table 7.

For the first sample, disaggregate information helps forecast aggregate U.S. inflation one and 12 months ahead. The improvements over the AR model are up to 6.5% in RMSFE terms

Table 7. U.S., RMSFE ratios

	1970–1983		1984–2004	
Horizon:	1	12	1	12
RMSFE AR^{SIC}	**0.280**	**2.660**	**0.193**	**1.296**
RMSFE ratios over AR^{SIC}				
FM($f1$)	0.969	1.000	0.996	0.980
FM($f2$)	0.964	0.976	1.007	1.009
FM($f3$)	0.972	0.979	1.010	0.999
FM($f1$)$_1$	0.957	1.003	1.002	0.972
FM($f2$)$_1$	0.936	0.964	1.010	1.017
FM($f3$)$_1$	0.948	0.969	0.992	0.975
p^f	1.007	0.986	0.999	0.971
p^c	0.982	1.014	1.006	0.980
p^s	1.005	0.995	0.998	1.044
p^e	1.004	0.997	1.013	1.017
p^{comb}	0.995	0.995	1.000	0.995

NOTE: RMSFE (not annualized) for AR$^{(SIC)}$ model in percentage points; SIC: lag order selection by Schwarz criterion; Recursive estimation samples 1960(1) to 1970(1), ..., 1983(12) and 1960(1) to 1984(1), ..., 2004(12); FM(fi): factor models with $i = 1, 2, 3$ static factors; FM(fi)$_1$: factor models with $i = 1, 2, 3$ factors with 1 lag; principal component estimators of static factors; p^f, p^c, p^s, p^e: single predictor models with respective subcomponent as predictor; p^{comb}: simple average of the forecasts from the four disaggregate component models.

(up to 12.5% in MSFE terms). Including one factor is statistically significant for the first sample period for a one-month horizon by the Clark and West (2007) test of equal forecast accuracy. However, the improvement using factor models is lower in the second sample period. This is in line with what Stock and Watson (2007) find for including real variables in an inflation model.

4.4 Summary of Empirical Results

To summarize our empirical results, overall the direct forecast of the aggregate, either using only past aggregate or using disaggregate information, is more accurate than combining disaggregate forecasts. Therefore, combining disaggregate information helps over combining disaggregate forecasts. Further, including a selected number of disaggregate variables or factors summarizing disaggregate information tends to improve forecast accuracy over forecasting the aggregate directly by only using past aggregate information, in particular in samples with sufficient variability in the aggregate.

5. CONCLUSIONS

We presented new analytical results on the relative forecast accuracy of forecasting an aggregate by (1) combining disaggregate forecasts (forecasting the disaggregate variables then aggregating those forecasts), (2) using only lagged aggregate information, and (3) to combine disaggregate information by including a subset of disaggregate components (or a combination thereof) in the aggregate model.

In the analytical derivations we investigated the effects of misspecification and estimation uncertainty on the relative forecast accuracy of the three different approaches to forecasting an aggregate, and we extended previous results by allowing for a change in the parameters of the DGP unknown to the forecaster, forecast origin uncertainty and time-varying weights. Decompositions of the sources of forecast errors led us to conclude that relative forecast accuracy is not affected by forecast-origin location shifts and slope changes, whereas absolute accuracy is. This is in contrast to the forecast combination literature, which focuses on combining forecasts of the same variable, where combination helps in the presence of mean shifts in opposite directions. Our second main result, in addition to a number of other important conclusions, is that slope misspecification and estimation uncertainty are the primary sources of differences in forecast accuracy between the different methods.

In the Monte Carlo simulations we find that including disaggregate variables in the aggregate model helps forecast the aggregate if the disaggregates follow different stochastic structures, the components are interdependent, and only a selected number of components is included to reduce estimation uncertainty. Unknown and unmodeled structural change in the mean does not affect relative forecast error of the different forecast methods, even though it has major effects on absolute forecast accuracy.

The empirical results for U.S. CPI inflation before and after the Great Moderation confirmed our analytical and simulation findings that estimation uncertainty plays an important role in relative forecast accuracy across the different approaches to forecast an aggregate. Consequently, we recommend model selection procedures for choosing the disaggregates to be included in the aggregate model, or methods to combine disaggregate information, and careful modeling of location shifts. Alternative methods for reducing estimation uncertainty, such as Bayesian or shrinkage methods, are beyond the scope of the paper, but are an interesting direction of further research in this context.

APPENDIX: FORECAST ERROR DECOMPOSITION—ADDITIONAL DERIVATIONS

The decomposition for the long-run mean, the first bracketed term in (5), is

$$\boldsymbol{\omega}'\boldsymbol{\phi}_y^* - \boldsymbol{\omega}'\hat{\boldsymbol{\phi}}_y = \boldsymbol{\omega}'(\boldsymbol{\phi}_y^* - \boldsymbol{\phi}_y) + \boldsymbol{\omega}'(\boldsymbol{\phi}_y - \boldsymbol{\phi}_{y,e}) + \boldsymbol{\omega}'(\boldsymbol{\phi}_{y,e} - \hat{\boldsymbol{\phi}}_y). \quad \text{(A.1)}$$

Decomposing $\boldsymbol{\omega}'\hat{\boldsymbol{\Gamma}}(\hat{\mathbf{y}}_T - \hat{\boldsymbol{\phi}}_y) = \boldsymbol{\omega}'\hat{\boldsymbol{\Gamma}}(\hat{\mathbf{y}}_T - \mathbf{y}_T) + \boldsymbol{\omega}'\hat{\boldsymbol{\Gamma}}(\mathbf{y}_T - \hat{\boldsymbol{\phi}}_y)$ to separate the measurement error, the second bracketed term in (5), becomes $\boldsymbol{\omega}'\boldsymbol{\Gamma}^*(\mathbf{y}_T - \boldsymbol{\phi}_y^*) - \boldsymbol{\omega}'\hat{\boldsymbol{\Gamma}}(\mathbf{y}_T - \hat{\boldsymbol{\phi}}_y)$ and can be decomposed as

$$\begin{aligned} &-\boldsymbol{\omega}'(\boldsymbol{\Gamma}^* - \boldsymbol{\Gamma})(\boldsymbol{\phi}_y^* - \boldsymbol{\phi}_y) - \boldsymbol{\omega}'(\boldsymbol{\Gamma} - \boldsymbol{\Gamma}_e)(\boldsymbol{\phi}_y^* - \boldsymbol{\phi}_y) \\ &\quad - \boldsymbol{\omega}'\boldsymbol{\Gamma}_e(\boldsymbol{\phi}_y^* - \boldsymbol{\phi}_y) + \boldsymbol{\omega}'(\boldsymbol{\Gamma}^* - \boldsymbol{\Gamma})(\mathbf{y}_T - \boldsymbol{\phi}_y) \\ &\quad + \boldsymbol{\omega}'(\boldsymbol{\Gamma} - \boldsymbol{\Gamma}_e)(\mathbf{y}_T - \boldsymbol{\phi}_y) - \boldsymbol{\omega}'\boldsymbol{\Gamma}_e(\boldsymbol{\phi}_y - \boldsymbol{\phi}_{y,e}) \\ &\quad - \boldsymbol{\omega}'(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}_e)(\mathbf{y}_T - \boldsymbol{\phi}_y) - \boldsymbol{\omega}'(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}_e)(\boldsymbol{\phi}_y - \boldsymbol{\phi}_{y,e}) \\ &\quad + \boldsymbol{\omega}'\boldsymbol{\Gamma}_e(\hat{\boldsymbol{\phi}}_y - \boldsymbol{\phi}_{y,e}) + \boldsymbol{\omega}'(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}_e)(\hat{\boldsymbol{\phi}}_y - \boldsymbol{\phi}_{y,e}). \end{aligned} \quad \text{(A.2)}$$

$\boldsymbol{\omega}'\hat{\boldsymbol{\Gamma}}(\hat{\mathbf{y}}_T - \mathbf{y}_T)$ can be decomposed, such that collecting terms from (A.1) and (A.2) above yields the taxonomy in (6).

ACKNOWLEDGMENTS

We thank Marcel Bluhm and Eleonora Granziera for excellent research assistance. We thank two anonymous referees, an associate editor, and the editor Arthur Lewbel for valuable comments. We are grateful to Olivier de Bandt, Domenico Gianonne, Lutz Kilian, Helmut Lütkepohl, Serena Ng, Lucrezia Reichlin, Kenneth West, and seminar participants at Bocconi University, Duke University, the Federal Reserve Bank of Atlanta, the Federal Reserve Board, the University of Michigan, the University of Wisconsin, participants of the EUI Conference 2006, ISF 2007, NASM 2007, and EABCN-CEPR conference 2007 for useful comments. We thank Walter Lane and Patrick Jackman (BLS) for providing the historical weights for the U.S. CPI data. Financial support from the ESRC under grant RES 051 270035 is gratefully acknowledged. The views expressed in this paper are not necessarily those of the European Central Bank.

[Received May 2007. Revised April 2009.]

REFERENCES

Anderson, T. W. (1984), *An Introduction to Multivariate Statistical Analysis* (2nd ed.), New York: Wiley. [225]

Atkeson, A., and Ohanian, L. E. (2001), "Are Phillips Curves Useful for Forecasting Inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 25 (1), 2–11. [223,224]

Bai, J. (2003), "Inference on Factor Models of Large Dimensions," *Econometrica*, 71 (1), 135–172. [225]

Bai, J., and Ng, S. (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70 (1), 191–221. [225]

Benalal, N., del Hoyo, J. L. D., Landau, B., Roma, M., and Skudelny, F. (2004), "To Aggregate or Not to Aggregate? Euro Area Inflation Forecasting," Working Paper 374, European Central Bank. [216,223]

Bernanke, B. (2007), "Inflation Expectations and Inflation Forecasting," speech at the Monetary Economics Workshop of the NBER Summer Institute. Available at *http://www.federalreserve.gov/newsevents/speech/Bernanke20070710a.htm*. [216]

Boivin, J., and Ng, S. (2006), "Are More Data Always Better for Factor Analysis?" *Journal of Econometrics*, 132, 169–194. [225]

Bruneau, C., De Bandt, O., Flageollet, A., and Michaux, E. (2007), "Forecasting Inflation Using Economic Indicators: The Case of France," *Journal of Forecasting*, 26, 1–22. [216,223]

Carson, R., Cenesizoglu, T., and Parker, R. (2007), "Aggregation Issues in Forecasting Aggregate Demand: An Application to US Commercial Air Travel," unpublished manuscript, University of California. Available at *http://ssrn.com/abstract=1401453*. [217]

Clark, T. E., and McCracken, M. W. (2006), "Forecasting With Small Macroeconomic VARs in the Presence of Instabilities," Research Working Paper 06-09, The Federal Reserve Bank of Kansas City. [217]

Clark, T. E., and West, K. D. (2007), "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models," *Journal of Econometrics*, 138 (1), 291–311. [225,226]

Clements, M. P., and Hendry, D. F. (1998), *Forecasting Economic Time Series*, Cambridge, U.K.: Cambridge University Press. [217,218,225]

——— (1999), *Forecasting Non-Stationary Economic Time Series*, Cambridge, MA: MIT Press. [217]

——— (2004), "Pooling Forecasts," *Econometrics Journal*, 7, 1–31. [219]

——— (2006), "Forecasting With Breaks in Data Processes," in *Handbook of Economic Forecasting*, eds. G. Elliott, C. W. J. Granger, and A. Timmermann, Amsterdam, Netherlands: Elsevier, pp. 605–657. [217,218]

Elliott, G. (2007), "Unit Root Pre-Testing and Forecasting," mimeo, University of California, San Diego. Available at *http://www-leland.stanford.edu/group/SITE/archive/SITE_2006/Web%20Session%205/Elliott.pdf*. [218]

Espasa, A., Senra, E., and Albacete, R. (2002), "Forecasting Inflation in the European Monetary Union: A Disaggregated Approach by Countries and by Sectors," *European Journal of Finance*, 8 (4), 402–421. [216]

Fair, R. C., and Shiller, J. (1990), "Comparing Information in Forecasts From Econometric Models," *The American Economic Review*, 80 (3), 375–389. [216,223]

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000), "The Generalized Factor Model: Identification and Estimation," *Review of Economics and Statistics*, 82, 540–554. [225]

——— (2005), "The Generalized Factor Model: One-Sided Estimation and Forecasting," *Journal of the American Statistical Association*, 100 (471), 830–840. [225]

Garderen, V. K. J., Lee, K., and Pesaran, M. H. (2000), "Cross-Sectional Aggregation of Non-Linear Models," *Journal of Econometrics*, 95, 285–331. [216]

Geweke, J. F. (1977), "The Dynamic Factor Analysis of Economic Time Series," in *Latent Variables in Socio-Economic Models*, eds. D. J. Aigner and A. S. Goldberger, Amsterdam: North-Holland. [225]

Giacomini, R., and Granger, C. W. J. (2004), "Aggregation of Space-Time Processes," *Journal of Econometrics*, 118, 7–26. [216,217,220]

Granger, C. (1987), "Implications of Aggregation With Common Factors," *Econometric Theory*, 3, 208–222. [216,217]

Granger, C. W. J. (1980), "Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, 14, 227–238. [217]

Grunfeld, Y., and Griliches, Z. (1960), "Is Aggregation Necessarily Bad?" *The Review of Economics and Statistics*, 42 (1), 1–13. [216]

Hendry, D. F., and Hubrich, K. (2006), "Forecasting Aggregates by Disaggregates," Working Paper 589, European Central Bank. [217]

Hernandez-Murillo, R., and Owyang, M. T. (2006), "The Information Content of Regional Employment Data for Forecasting Aggregate Conditions," *Economics Letters*, 90, 335–339. [217]

Hubrich, K. (2005), "Forecasting Euro Area Inflation: Does Aggregating Forecasts by HICP Component Improve Forecast Accuracy?" *International Journal of Forecasting*, 21 (1), 119–136. [216,217,223]

Hubrich, K., and West, K. (2010), "Forecast Evaluation of Small Nested Model Sets," *Journal of Applied Econometrics*, to appear. [225]

Kohn, R. (1982), "When Is an Aggregate of a Time Series Efficiently Forecast by Its Past?" *Journal of Econometrics*, 18, 337–349. [216,217]

Lütkepohl, H. (1984a), "Linear Transformations of Vector ARMA Processes," *Journal of Econometrics*, 26, 283–293. [217]

——— (1984b), "Forecasting Contemporaneously Aggregated Vector ARMA Processes," *Journal of Business & Economic Statistics*, 2 (3), 201–214. [216,217,220,221]

——— (1987), *Forecasting Aggregated Vector ARMA Processes*, Berlin, Germany: Springer-Verlag. [216,220,221]

——— (2006), "Forecasting With VARMA Processes," in *Handbook of Economic Forecasting*, eds. G. Elliott, C. W. J. Granger, and A. Timmermann, Amsterdam, Netherlands: Elsevier. [216]

Marcellino, M., Stock, J. H., and Watson, M. W. (2003), "Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-Wide Information," *European Economic Review*, 47, 1–18. [216,223]

Moser, G., Rumler, F., and Scharler, J. (2007), "Forecasting Austrian Inflation," *Economic Modelling*, 24, 470–480. [216]

Pesaran, M. H., Pierse, R. G., and Kumar, M. S. (1989), "Econometric Analysis of Aggregation in the Context of Linear Prediction Models," *Econometrica*, 57, 861–888. [216]

Reijer, A., and Vlaar, P. (2006), "Forecasting Inflation: An Art as Well as a Science!" *De Economist*, 127 (1), 19–40. [216]

Sargent, T. J., and Sims, C. A. (1977), "Business Cycle Modeling Without Pretending to Have Too Much a-priori Theory," in *New Methods in Business Cycle Research*, ed. C. Sims, Minneapolis: Federal Reserve Bank of Minneapolis. [225]

Stock, J. H., and Watson, M. W. (1991), "A Probability Model of the Coincident Economic Indicator," in *The Leading Economic Indicators: New Approaches and Forecasting Records*, eds. G. Moore and K. Lahiri, Cambridge, U.K.: Cambridge University Press, pp. 63–90. [225]

——— (1996), "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business & Economic Statistics*, 14 (1), 11–30. [217]

——— (1998), "Diffusion Indexes," Working Paper 6702, NBER. [225]

——— (1999), "Forecasting Inflation," *Journal of Monetary Economics*, 44, 293–335. [225]

——— (2002a), "Forecasting Using Principal Components From a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167–1179. [225]

——— (2002b), "Macroeconomic Forecasting Using Diffusion Indices," *Journal of Business & Economic Statistics*, 20 (2), 147–162. [225]

——— (2007), "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking*, 39, 3–34. [217,223-226]

Zellner, A., and Tobias, J. (2000), "A Note on Aggregation, Disaggregation and Forecasting Performance," *Journal of Forecasting*, 19, 457–469. [216,217,223]