working paper

15 02

Lessons for Forecasting Unemployment in the U.S.: Use Flow Rates, Mind the Trend

Brent Meyer and Murat Tasci

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 15-02 February 2015

Lessons for Forecasting Unemployment in the U.S.:
Use Flow Rates, Mind the Trend

Brent Meyer and Murat Tasci

This paper evaluates the ability of autoregressive models, professional forecasters, and models that leverage unemployment flows to forecast the unemployment rate. We pay particular attention to flows-based approaches—the more reduced-form approach of Barnichon and Nekarda (2012) and the more structural method in Tasci (2012)—to generalize whether data on unemployment flows is useful in forecasting the unemployment rate. We find that any approach that leverages unemployment inflow and outflow rates performs well in the near term. Over longer forecast horizons, Tasci (2012) appears to be a useful framework, even though it was designed to be mainly a tool to uncover long-run labor market dynamics such as the "natural" rate. Its usefulness is amplified at specific points in the business cycle when unemployment rate is away from the longer-run natural rate. Judgmental forecasts from professional economists tend to be the single best predictor of future unemployment rates. However, combining those guesses with flows-based approaches yields significant gains in forecasting accuracy.

Key words: Unemployment Forecasting, Natural Rate, Unemployment Flows, Labor Market Search.

JEL classification: E24, E32, J64, C53.

Suggested citation: Meyer, Brent and Murat Tasci, 2015. "Lessons for Forecasting Unemployment in the U.S.: Use Flow Rates, Mind the Trend," Federal Reserve Bank of Cleveland, working paper no 15-02.

Brent Meyer is at the Federal Reserve Bank of Atlanta (Brent.Meyer@atl.frb.org), and Murat Tasci is at the Federal Reserve Bank of Cleveland (Murat.Tasci@clev.frb.org). The authors thank participants of the Midwest Economic Association Meetings (Evanston, 2013). The views expressed herein are the authors and not necessarily those of the Federal Reserve Banks of Cleveland or Atlanta, or the Federal Reserve System.

1 Introduction

The unemployment rate has been the primary summary statistic for the health of the labor market for quite some time. Recently, however, forecasts of the unemployment rate have come to the forefront, as monetary policy makers are trying to formulate a way of conditioning expectations in the new and extraordinary policy environment. For instance, in September 2012, the Federal Open Market Committee (FOMC) decided to tie its asset purchases to a "substantial improvement" in labor market conditions and in December 2012, it made the tightening of the policy rate conditional on the level of the unemployment rate.[1]

Hence, the progression of the unemployment rate became a central issue in the policy debate. Furthermore, the behavior of the unemployment rate over the course of the Great Recession and subsequent recovery has left researchers and policy makers puzzled over whether there was a significant change in the long-run trend in the unemployment rate[2]. Given the new-found policy focus and potential for a shift in the dynamics of the unemployment rate since the Great Recession, we compare the forecast performance of different approaches to forecasting the series. In addition to considering the forecasts of professional forecasters (The Federal Reserve Board's *Greenbook,* The Federal Reserve Bank of Philadelphia's *Survey of Professional Forecasters,* and the *Blue Chip* panel of economists) and a few well-known autoregressive models of the unemployment rate, we pay special attention to new research that focuses on unemployment flows (job-finding and separation rates, in particular) and their role in accounting for unemployment fluctuations.

A novel method that leverages data on unemployment flows to forecast the unemployment rate was recently put forth by Barnichon and Nekarda (2012). Using a simple vector autoregression (VAR) for unemployment flows to predict unemployment rate in quasi-real-time, along with certain leading indicators such as initial claims for unemployment insurance and job vacancies, they report forecasts that dramatically outperform the Survey of Professional Forecasters, the Federal Reserve Board's Greenbook Forecast, and basic univariate time-series models over near-term forecast horizons in their sample. The

[1] In particular the FOMC Statement read: "... In particular, the Committee decided to keep the target range for the federal funds rate at 0 to 1/4 percent and currently anticipates that this exceptionally low range for the federal funds rate will be appropriate at least as long as the unemployment rate remains above 6-1/2 percent, inflation between one and two years ahead is projected to be no more than a half percentage point above the Committee's 2 percent longer-run goal, and longer-term inflation expectations continue to be well anchored." - FOMC Statement, December 12, 2012.

[2] This issue often took the form of a debate about the nature of the high unemployment rate after the Great Recession. That is, whether the high unemployment refelected purely cyclical factors or structural change (Bernanke (2012), Kocherlakota (2010)).

exercise is in quasi-real-time, as the true real-time data on initial claims and job vacancies were not used.[3] Another approach we investigate that leverages flows data is Tasci (2012), which uses a simple econometric model of comovement between the aggregate economic activity and unemployment flows to uncover the unobserved trend components of the underlying flow rates. These unobserved trends then pin down the long-run trend of the unemployment rate in a way that is consistent with modern theory of unemployment. In this paper, our focus will be on the forecasting performance of that model, recognizing that it yields a natural forecasting framework that is also consistent with a well defined long-run trend for the unemployment rate.

Our paper is related to the line of research that aims to address the forecasting challenges of macro-aggregates in general and the unemployment rate in particular, such as Montgomery et. al. (1998) and Rothman (1998), among others. Most of the focus in these early studies were on the asymmetric nature of the unemployment rate over the business cycle and the adequacy of linear models to address this. As in Barnichon and Nekarda (2012), we also rely on linear models, but the underlying equation of motion for the unemployment rate and the focus on the flows in and out of it accommodates the non-linear nature of the unemployment movements with ease and results in substantial forecast performance improvements. Our focus on flow rates is also related to the recent literature on the importance of flow rates in explaining unemployment fluctuations in the U.S., such as Shimer (2005, 2012), Elsby, Michaels, and Solon (2009), and Fujita and Ramey (2009). Our baseline model, Tasci (2012), is closely related to studies of measuring the cyclical component of economic aggregates, as in Clark (1987, 1989) and Kim and Nelson (1999). In the next section, we describe the model in some detail, closely following Tasci (2012) and the forecasting approach taken in Barnichon and Nekarda (2012).

We compare the forecasting performance of three approaches to predicting the unemployment rate (non-linear autoregressive models, flows-based models, and professional forecasters) to a simple linear autoregressive benchmark. Not only do we evaluate these in terms of relative root mean-squared errors (RMSEs), but also we attempt to determine statistical significance based on a variant of the Diebold and Mariano (1995) equality-of-prediction test. Additionally, we employ a few regression-based and simple-average-forecast combinations. Other tests include a conditional forecasting exercise, where we leverage the structure of Tasci (2012) by augmenting some of the embedded forecasts to back out different paths for the unemployment flow rates. While the paper is a straight-

[3] Both series are subject to seasonal adjustment factors, which Barnichon and Nekarda (2012) claim to be inconsequential. However, our analysis shows that a great deal of the forecast improvement is due these variables.

forward "horse-race," we also report a few "practitioners' issues" that we uncovered in the course of our analysis –choices which seemed trivial on the surface but which led to material differences in some cases, such as using forecasts that have been rounded to the nearest tenth, the timing of forecasts within a month, and differing sample periods.

In general, we find leveraging data on unemployment flows yields a "nowcast" (current-quarter forecast) superior to professional forecasts over most samples we investigate, but those gains disappear (and usually reverse) relative to professional forecasts beyond a 1-quarter-ahead forecasting horizon. Combining unemployment rate forecasts from professional forecasters and the two flows-based models using regression weights or simple averaging was superior to any single approach we investigated. In contrast to Montgomery et. al. (1998) and Rothman (1998), we find little support for non-linear time-series methods. This also holds true for the Barnichon and Nekarda (2012) flows-based approach, as we find the simple (linear) VAR model they employ tends to outperform their "official" approach. Perhaps the most disappointing aspect of our investigation–and one that merits further discussion–is that, while professional forecasters and flows-based models tend to *significantly* outperform our simple autoregressive benchmark through the near-term (current-quarter to 1-year ahead), no single approach we investigate *significantly* improves on that benchmark over longer forecast horizons (8-quarters ahead), over our full sample period.

2 Approaches to Forecasting the Unemployment Rate

We evaluate the forecasting performance of three distinct approaches to predicting the unemployment rate. The first group consists of a set of univariate autoregressive models, including a benchmark AR(6) model. The second group includes professional forecasts that are available at different sample periods and varying forecast horizons. The thrid group consists of models that incorporate unemployment flows as a forecasting tool and includes the rather structural and parsimonious model of Tasci (2012).

2.1 Univariate Autoregressive Models

We chose three simple autoregressive statistical models to compare to the flows-based forecasts and professional forecasters. The motivation for using these models comes from the literature on forecasting the unemployment rate–namely Montgomery et al. (1998) and Rothman (1998). The simplest version, the AR model, is a standard benchmark across most of the forecasting literature, used for its parsimony and its ability to project the persistent part of a series. Unfortunately, it often becomes a measure of economists' collective ignorance, as it is hard to beat (see Atkeson and Ohanian (2001) among others). The other two models, the generalized autoregressive (GAR) and self-exciting threshold autoregressive (SETAR) models, were chosen for a couple of reasons. First, as Rothman (1998) puts it, these models are "state dependent" in that their behavior changes given the recent past behavior of the series. Second, these two approaches attempt to model the asymmetry observed in the unemployment rate, which has been long documented (Neftci (1984) and Rothman (1991), among others). During recessions, the unemployment rate rises rapidly, but as the recovery takes hold, it declines only gradually. This feature of the unemployment rate can become troublesome for linear models that are unable to incorporate those dynamics.

2.1.1 Autoregressive model (AR)

We would like to perform our forecast evaluation across different frameworks at the highest possible frequency possible. Hence, we chose a monthly baseline AR(6) specification for this exercise, as it corresponds to the quarterly statistical models used in Montgomery et. al. (1998) and Rothman (1998).[4]

$$U_t = \beta_0 + \sum_{i=1}^{6} \beta_i U_{t-i} + \epsilon_t \tag{1}$$

This specification, expressed in equation (1), will serve as our benchmark forecasting equation. Forecast improvements across different frameworks will be compared to this basic statistical benchmark.

2.1.2 Generalized autoregressive model (GAR)

As we described above, earlier literature identified potential gains from non-linear specifications, because they could capture the asymmetric behavior of the unemployment rate over business cycles. Following this, we chose a GAR(6) specification for the monthly data. This model performed well in out-of-sample forecast tests in Rothman (1998). In his quarterly GAR(2) model, the second lag of the unemployment rate also enters into the equation with a cubic term.

$$U_t = \beta_0 + \sum_{i=1}^{6} \beta_i U_{t-i} + \sum_{i=4}^{6} \beta_i U_{t-i}^3 + \epsilon_t \tag{2}$$

[4] We need 6 lags in our baseline estimation period to soak up all the excess serial correlation, obtaining a DW stat of nearly 2.0.

Since we focus on a monthly model, lags 4-6 enter in levels and with a cubic term in our GAR(6) specification expressed in equation (2).

2.1.3 Self-exciting threshold autoregressive model (SETAR)

A version of the SETAR model was used in both Montgomery et al. (1998) and Rothman (1998). This model follows two sets of dynamics given a specific threshold, allowing us to exploit the asymmetry observed in the unemployment rate more explicitly. We take our threshold directly from Montgomery et. al. (1998), which is equal to 0.1 percentage point of the change in the previous quarter's unemployment rate.

$$\Delta U_t = \begin{cases} \beta_0 + \sum_{i=1}^{6} \beta_i \Delta U_{t-i} + \epsilon_t, \text{ if } \sum_{i=4}^{6} \Delta U_{t-i} \leq 0.1 \\ \alpha_0 + \sum_{i=1}^{6} \alpha_i \Delta U_{t-i} + \varsigma_t, \text{ otherwise} \end{cases}$$

Hence, when the unemployment rate starts to rise (as during a recession), a different dynamic will govern the forecasts relative to more stable times, when the unemployment rate is either declining or has recorded only a small increase.

2.2 Professional Forecasts

We focus on three different, commonly available professional forecasts for our analysis. First, we collect unemployment forecasts from the Federal Reserve Board's *Greenbook Part I* (now called *Tealbook Part I*). The *Greenbook* (GB) provides the Fed Board staff's summary of economic conditions and forecasts and is distributed to Federal Open Market Committee (FOMC) participants roughly a week prior to an FOMC meeting. These forecasts are released to the public after a five-year period (the last forecast year we have available is 2007). The FOMC usually meets 8 times a year, and the timing of these meetings (and information available to the staff at the time of a forecast) varies somewhat from year to year. We will return to this complication in the following section.

Our second set of professional forecasts comes from the Federal Reserve Bank of Philadelphia's *Survey of Professional Forecasters* (SPF). The Philadelphia Fed took over the survey from the American Statistical Association (ASA) and National Bureau of Economic Research (NBER) in 1990. Participants are surveyed quarterly, usually at the end of the first month of each quarter (timed to concur with the release of the BEA's advance GDP report). They are asked to provide forecasts for 32 economic variables for the current quarter through 4-quarters ahead. The forecasts are tabulated and released to the public usually by the middle of the following month.

Lastly, we also put together a time series of the consensus (mean) forecast from the

Blue Chip Economic Indicators (BC) panel of roughly 50 forecasters. Participants for this panel are surveyed monthly, and the results are tabulated and released on the 10th of every month. This survey asks for quarterly forecasts from the current quarter through the fourth quarter of the next calendar year (which creates its own set of sampling issues) as well as annual averages for the current and following calendar years. We will review some of the important timing and forecast horizon issues later in the Data section.

2.3 Forecasts Relying on Unemployment Flows

We use two recent unemployment flows-based models in this forecasting exercise. The first is from Tasci (2012). The other one is a vector-autoregression based forecasting exercise proposed in Barnichon and Nekarda (2012). Since our focus will be evaluating the role of incorporating unemployment flows in forecasting the unemployment rate, we provide a detailed discussion of these two forecasting approaches in this section.

Although Tasci (2012) focuses on the long-run behavior of the unemployment rate and underlying flow rates, we focus on the forecasting performance of the model using real-time data. This simple econometric model incorporates the comovement of flows into and out of unemployment with aggregate output and delivers a theoretically meaningful long-run unemployment trend. The main premise is of our use of Tasci (2012) is that by disciplining the long-run unemployment rate—the so-called natural rate—we can improve forecasting performance in the short to medium term.

In our implementation of this framework, we assume that real GDP has both a stochastic trend and a stationary cyclical component, but these components are not observed by the econometrician. We also assume that both flow rates, F_t and S_t, (job-finding and separation rate, respectively) have a stochastic trend as well as a stationary cyclical component. The stochastic trend follows a random walk, but the cyclical component in the flow rates depends on the cyclical component of real GDP. More specifically, let Y_t be log real GDP, $\bar{y}_t$ a stochastic trend component, and y_t the stationary cyclical component. Similarly, let F_t (S_t) be the quarterly job finding (separation) rate, $\bar{f}_t$ ($\bar{s}_t$) its stochastic trend component, and f_t (s_t) the stationary cyclical component. Then we consider the following unobserved components model:

$$Y_t = \bar{y}_t + y_t; \quad \bar{y}_t = g_{t-1} + \bar{y}_{t-1} + \varepsilon_t^{yn}; \quad g_t = g_{t-1} + \varepsilon_t^{g}; \quad y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t^{yc} \tag{3}$$

$$F_t = \bar{f}_t + f_t; \quad \bar{f}_t = \bar{f}_{t-1} + \varepsilon_t^{fn}; \quad f_t = \rho_1 y_t + \rho_2 y_{t-1} + \rho_3 y_{t-2} + \varepsilon_t^{fc} \tag{4}$$

$$S_t = \bar{s}_t + s_t; \quad \bar{s}_t = \bar{s}_{t-1} + \varepsilon_t^{sn}; \; s_t = \theta_1 y_t + \theta_2 y_{t-1} + \theta_3 y_{t-2} + \varepsilon_t^{sc} \tag{5}$$

where g_t is a drift term in the stochastic trend component of output which is also a random walk, following Tasci (2012). All the error terms, ε_t^{yn}, ε_t^{g}, ε_t^{yc}, ε_t^{fn}, ε_t^{fc}, ε_t^{sn}, ε_t^{sc}, are independent white-noise processes.

Equation (3) is very conventional and governs the movement in real output. We impose a stochastic trend, which might be subject to occasional drifts, and a persistent but stationary cyclical component. The comovement in the rates of job finding and separation expressed in (4) and (5) is more unconventional. However, Tasci (2012) argues that one can map this empirical representation to a simple extension of the textbook search model with endogenous job destruction and shocks to aggregate productivity, as in Mortensen and Pissarides (1994).[5]

The trend of the unemployment rate in this model is pinned down by the stochastic trend components of the job-finding and separation rates, which is the main focus of Tasci (2012). We can estimate this model and use the Kalman filter to back out the underlying trends to get an estimate of a time-varying trend. More importantly for us, however, we can use our estimates for the model at any point and we can generate forecasts for the underlying flow rates and the unemployment rate. To start with, first we write down the system of equations (3)-(5), in the following state-space representation:

$$\begin{bmatrix} Y_t \\ F_t \\ S_t \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \rho_1 & \rho_2 & \rho_3 & 0 & 1 & 0 \\ 0 & \theta_1 & \theta_2 & \theta_3 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{y}_t \\ y_t \\ y_{t-1} \\ y_{t-2} \\ g_t \\ \bar{f}_t \\ \bar{s}_t \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon_t^{fc} \\ \varepsilon_t^{sc} \end{bmatrix} \tag{6}$$

[5]The low-frequency movements in the trends, $\bar{f}_t$ and $\bar{s}_t$, are assumed to capture the effects of institutions, demographics, tax structure, labor market rigidities, and the long-run matching efficiency of the labor markets, which will be more important in determining the steady state of unemployment. The cyclical components, f_t and s_t, on the other hand, move in response to purely cyclical changes in output. In this class of models, market tightness—hence the job-finding rate—increases during expansions and declines during recessions. Similarly, when aggregate productivity is temporarily low, there will be a surge of separations, resulting in higher unemployment, because some existing matches cease to be productive enough in the recession. Hence, the assumed relationship of (4) and (5) is in line with the predictions of the search theory of unemployment.

$$
\begin{bmatrix} \bar{y}_t \\ y_t \\ y_{t-1} \\ y_{t-2} \\ g_t \\ \bar{f}_t \\ \bar{s}_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \phi_1 & \phi_2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{y}_{t-1} \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ g_{t-1} \\ \bar{f}_{t-1} \\ \bar{s}_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^{yn} \\ \varepsilon_t^{yc} \\ 0 \\ 0 \\ \varepsilon_t^{g} \\ \varepsilon_t^{fn} \\ \varepsilon_t^{sn} \end{bmatrix} \tag{7}
$$

where all error terms come from an i.i.d. normal distribution, with zero mean and variance σ_i such that $i = \{yn, g, yc, fn, fc, sn, sc\}$. Once we estimate this model using US data, we can back out an estimate of a time-varying unemployment rate trend by using the estimates of the unobserved trend components. In particular, $\bar{u}_t = \frac{\bar{s}_t}{\bar{s}_t + \bar{f}_t}$ will give us the desired unemployment rate trend, which the trend in the flows will predict in the long-run.

When we simulate the model forward from the current quarter, this unemployment trend will be the one that the forecast will converge to in the long run. This feature of the model provides the necessary gravitational force on the unemployment rate when it deviates from its steady state. Moreover, any secular change in the average growth rate of output, g_t, or the flow rates will be reflected in the forecast. Tasci (2012) shows that there has been a significant slowdown in labor market turnover, which has manifested itself as declining trends in both outflow and inflow rates. Therefore, for instance, our model will produce a distinctly different forecast path from the same initial unemployment rate during the recent recovery as opposed to the aftermath of the 1981-82 recession. Lower turnover will imply more persistence in the unemployment rate now, as the degree of labor market churning determines how quickly unemployment adjusts. All these important characteristics of current labor market trends can easily be incorporated into the forecasting exercise due to our structure. We believe this is a very important contribution of our paper, and will discipline the forecast in the right direction.

Note that equations (6)-(7) do not have observed unemployment in them. Therefore, once we have j-period-ahead forecasts of the outflow rate, $\hat{F}_{t+j}$, and the inflow rate, $\hat{S}_{t+j}$, we use the implied equation of motion for the unemployment rate forward, starting from the last available unemployment rate in real-time, u_{t-1}. This equation of motion is an integral part of the measurement of the flow rates in the data and is very standard in the literature (Elsby et. al. (2011) and Shimer (2012)). We can then use the following recursive equation to generate our unemployment forecasts based on the forecasts of the flow rates from the model.

$$
\hat{u}_t = \left(1 - e^{-\hat{F}_{t-1} - \hat{S}_{t-1}}\right) \frac{\hat{S}_{t-1}}{\hat{S}_{t-1} + \hat{F}_{t-1}} + e^{-\hat{F}_{t-1} - \hat{S}_{t-1}} \hat{u}_{t-1} \tag{8}
$$

As a result, the system of equations (6) through (8) constitutes our forecasting framework.[6] Given the trend estimates for the flow rates, $\bar{s}_t$, $\bar{f}_t$, and the implied natural rate, $\frac{\bar{s}_t}{\bar{s}_t+\bar{f}_t}$, our framework enables us to project a certain path for the observed counterparts, $\hat{S}_t, \hat{F}_t$ and $\hat{u}_t$. We will refer to this unemployment rate forecast as the FLOW-UC case. Notice that deviations from $\bar{s}_t$, $\bar{f}_t$ will disappear as the cyclical component of output, y_t, converges to zero. This will in turn manifest itself as a gradually declining gap between the observed unemployment rate and the natural rate estimate, as equation (8) describes.

As mentioned in the previous section, our paper is not the first to use labor market flow rates to forecast the unemployment rate. Barnichon and Nekarda (2012) propose using flow rates to forecast the aggregate unemployment rate. Conceptually it is the closest approach to Tasci (2012) and relies on a vector autoregression to forecast $\hat{F}_{t+j}$, and $\hat{S}_{t+j}$. Their preferred specification has information including leading indicators, such as unemployment insurance claims and vacancy data. More specifically, they run the following vector autoregression that includes additional sources of contemporaneous information

$$z_t = c + \Phi_1 z_{t-1} + \Phi_2 z_{t-2} + v_t \tag{9}$$

where $z_t = (\ln S_{t-1}, \ln F_{t-1}, \Delta \ln u_t, \ln uic_t, \ln vac_t)'$, uic_t is the monthly average of weekly unemployment insurance claims, and vac_t is Barnichon's (2010) composite help-wanted index. Because flow rates are lagged by one month, the last data point, z_t only contains the flow rate between month $t-1$ and month t. Once this representation in (9) is used to obtain forecasts for future flow rates, $\hat{F}_{t+j}$, and $\hat{S}_{t+j}$, they use the same equation of motion, (8), to generate the final unemployment rate forecasts. Notice that in principle, this resulting unemployment rate forecast might be different from the implied unemployment rate forecast in equation (9), since $\Delta \ln u_t$ is part of the vector z_t. In section 4, we compare our model's forecast with both of these unemployment rate forecasts. Hereafter, we will refer to the implicit forecast as the VAR case and the one generated by taking the flow rates from (9) into the equation of motion for unemployment, (8), as the FLOW-VAR case.

[6] There are two principal problems that need to be tackled in estimating the model. First, we need data on job-finding and separation rates for the aggregate economy, which are not readily available. This measurement issue is described in the Data section. Second, the model, as spelled out in equations (6)-(7), is subject to an identification problem. Even though we have only three observables, we are estimating parameters for seven shocks. We resolve these estimation issues using the approach discussed in Tasci (2012).

2.4 Conditioning the FLOW-UC Model

When comparing the forecasts from the FLOW-UC model to FLOW-VAR or VAR, one needs to recognize the information disadvantage that the FLOW-UC model has due to the quarterly nature of the real GDP data. For instance, in the first month of any quarter we will not only not have any GDP data for the current quarter but also for the previous quarter. By contrast, FLOW-VAR and VAR incorporate a lot of the contemporaneous information through monthly estimation that relies on high-frequency and timely observations of UIC and vacancy data. Hence, we also look at an extension of the FLOW-UC model, where we relied on the same VAR structure in (9) for the current quarter flow rate forecasts that are not observed. Essentially, we take the "nowcast" from the VAR and let the rest of the forecast horizon still be governed by the FLOW-UC model. This case, referred to as the FLOW-UC|VAR below, produces the same current-quarter unemployment rate forecast as the FLOW-VAR case, as they both use the same equation, (9), to get the flow rate forecasts and exploit (8) to generate the final unemployment rate forecasts. The implied steady state unemployment rate that forecasts converge to, as well as the evolution of it from quarter $t+1$ onwards, will be still different.

For the same reasons, when comparing the FLOW-UC model forecast to the professional forecasts, we also check whether the real GDP forecast from those professional forecasters could improve the forecast accuracy of the FLOW-UC model. Hence, we also have an additional set of forecasts where we condition the GDP forecast to follow one of the professional forecasts, GB, SPF or BC, and analyze the unemployment rate forecast.

3 Data and Some Practical Issues

We use real-time data on monthly labor market flows, based on a publicly available data set, the Current Population Survey (CPS). Using data on the levels of the labor force and the unemployment pool, as well as the number of short-term unemployed, we construct our measures for the observables, F_t, and S_t, following Shimer (2005, 2012). Our measure of real GDP is from the real-time dataset compiled by the Federal Reserve Bank of Philadelphia's Real-Time Data Research Center.[7] We use monthly vintages of all the data, starting from the first vintage as of January 1976. All of the data are available at a monthly frequency with the exception of GDP. In order to compare our results to the set of forecasts proposed in Barnichon and Nekarda (2012), we replicate their exercise using additional data. Their methodology relies on the same flow rates we

[7] Data can be downloaded from website: http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files/, <last accessed on May 6, 2013>.

have in addition to data on initial claims of unemployment benefits and vacancy data. Weekly unemployment claims data are provided by Department of Labor's Employment and Training Administration. The measure of aggregate vacancies is a composite index presented in Barnichon (2010). Unfortunately, these two data sources are not real-time, but only the former is subject to some revisions.[8]

All the real time data, starting from the January 1976 vintage, span the period between 1948:M1 through 2012:M12 for monthly data, and 1948:Q1 through 2012:Q4 for quarterly data. Because of the timing of the data releases, the information set at every point might be slightly different. For instance, real-time GDP data become available in the second month of every quarter for the previous quarter. Worker flows from the CPS will have a lag of two months, due to the way we construct this series.[9] Hence, it is important to keep in mind that in our real-time forecasting exercise, the information set evolves accordingly.

3.1 Constructing the Unemployment Flow Rates

Flow rates are not readily available for the aggregate economy. However, recent research on the cyclical features of unemployment, led by Shimer (2005, 2012) and Elsby, Michaels, and Solon (2009), provide us with a simple method to measure these rates using CPS data. The method infers continuous time hazard rates into and out of unemployment by using readily available short-term unemployment, aggregate unemployment, and labor force data. We briefly describe the method used to infer these rates here.

Let U_t be the number of unemployed workers in month t in the CPS, U_t^s the number who are unemployed less than five weeks in month t, and L_t the size of the labor force in month t. At the heart of the measurement is a simple equation determining the evolution of unemployment over time in terms of flows into and out of unemployment:

$$\frac{dU_t}{dt} = S_t(L_t - U_t) - F_t U_t. \tag{10}$$

Given this simple accounting equation, we start with a typical unemployed worker's probability of leaving unemployment. As Shimer (2012) and Elsby, Michaels, and Solon

[8] The composite index for the period between 1951 through 1995 entirely depends on the Conference Board's help-wanted print advertising index, which was not regularly revised.

[9] Labor force stocks, such as employment and unemployment, will only have one lag. However, as we describe below, solving for the separation rates implicitly requires an additional lag, as is evident in equation (12).

(2009) show, the job-finding probability will be given by the following relationship:

$$\hat{F}_t = 1 - \left[\left(U_{t+1} - U^s_{t+1}\right)/U_t\right] \tag{11}$$

which maps into an outflow hazard, the job-finding rate, $F_t = -\log(1 - \hat{F}_t)$. This formulation in (11) computes the job-finding probability for the average unemployed person by implicitly assuming that contraction in the pool of unemployed, net of newcomers to the pool (U^s_{t+1}), results from unemployed workers finding jobs. The next step is to estimate the separation rate S_t. This step involves solving the continuous-time equation of motion for unemployment forward to get the following equation, which uniquely identifies S_t.

$$U_{t+1} = \frac{\left(1 - e^{-F_t - S_t}\right) S_t}{F_t + S_t} L_t + e^{-F_t - S_t} U_t \tag{12}$$

Given the outflow hazard, F_t, measured through (11), and data on U_t and L_t, we can solve for S_t numerically for each month t. The equation of motion for the unemployment *rate*, (8), follows from the measurement equation for the separation rate, (12), as long as $\frac{L_{t+1}}{L_t} \simeq 1$, which holds in the data at a monthly frequency.

One potential problem that could bias the estimates is the redesign of the CPS in 1994. As discussed by Shimer (2012) and Elsby, Michaels, and Solon (2009), the CPS redesign deflated the actual number of short-term unemployed by changing the way this number is computed for every rotation group except the first and the fifth[10]. To correct for this bias, we follow Elsby, Michaels, and Solon (2009) and use the average fraction of short-term unemployed among the unaffected first and fifth rotation groups to inflate the aggregate short-term unemployment number. This reduces to multiplying every month's u^s_{t+1} by 1.1549 from February 1994 through the end of the sample period. Implementing this correction finally provides us with the data we need to compute unemployment flow rates. Note that we compute the flow rates in real time. Even though the micro data from the CPS are not subject to revision, periodic corrections in population adjustments might induce changes in weights, which might slightly affect aggregate stocks.

3.2 Some Practical Issues

The timing of the forecasts stand out as the most important practical issue that we have to pay attention, in our analysis. For model-based forecasts, conditioning on the same information set does not pose a serious problem, as the requisite data come from the same sources, i.e. the CPS. It gets a bit tricky when we compare the results from the

[10] See Polivka and Miller (1998) and Abraham and Shimer (2001) for more detail.

model forecasts to the professional forecasts. The SPF and BC have historically had a regular release calendar so it is easier to match to the corresponding model timing. However, for some observations in the GB sample, the timing is such that the corresponding month's model-based forecast is not compatible, as the GB forecast predates the BLS's employment release by a few days. We describe this issue in more detail and explore the potential bias in the robustness section.

Recall that we analyze the forecast performance for all models up to 8 quarters ahead, starting from 1976:M1, giving us a sample size of 420. This reduces the size of the comparable professional forecasts for horizons beyond 4 quarters. For instance, we start with 114 sample points in the GB sample for the current-quarter horizon, and that diminishes in horizons beyond $t+4$, ending with only 20 data points for $t+8$. The SPF sample only includes forecasts up to $t+4$, giving us a sample size of 140 for all forecast horizons. The BC sample, on the other hand, starts from 371 observations and declines to 90 observations by $t+7$.[11] Another seemingly innocuous issue is that professional forecasts for the unemployment rate are rounded to the nearest tenth. Since we have the exact stocks to generate the unemployment rate for each month very precisely, we refrain from rounding for the model forecasts and the benchmark data counterparts. In the robustness section, we explore the potential impact of this on our results.

4 Forecast Performance

In this section, we compare the professional forecasts (GB, SPF, and BC), the autoregressive model forecasts (GAR and SETAR), and the flows-based unemployment rate forecasts (FLOW-UC, VAR, and FLOW-VAR) to the simple AR benchmark. The forecasts generated from the unobserved components–FLOW-UC–model (equations (6) through (8)), rely heavily on Kalman filtering and smoothing, and are estimated recursively with maximum likelihood.[12] On the other hand, Barnichon and Nekarda's (2012) model and its forecasts (both FLOW-VAR and VAR) from the underlying vector autoregressive representation in (9) are estimated from samples of rolling windows. Barnichon and Nekarda (2012) note in their paper that the forecasting performance of the rolling-window estimation is superior to the recursively estimated alternative. Therefore, we will focus on the 15-year rolling window approach that they adopted in their paper.

The baseline estimation period runs from 1948Q1 to 1975Q4 (or 1948M1 to 1975M12

[11] In the tabulations hereafter, we will explicitly note the sample size, as appropriate.

[12] Any past data are important in the estimation of the FLOW-UC to better identify the unobserved trend. Hence, we do not use a rolling-window approach for this model.

for monthly data) and is expanded by one month at each iteration.[13] Our forecast evaluation period runs from 1976M1-2010M12.[14] Even though Barnichon and Nekarda (2012) focus on the near-term forecasting performance (t, $t+4$), we extend the evaluation period by 4 quarters (t, $t+8$) when available, to be more consistent with a policy-relevant horizon.[15,16]

We evaluate model performance using relative RMSFEs (using the AR model forecasts as the benchmark in each case as possible) and the equality-of-prediction tests (using MSFEs). We use a variant of the Diebold-Mariano test that attempts to control for serially correlated error terms. We employ the Harvey, Leybourne, and Newbold (1996), hereafter HLN, adjustment to the Diebold-Mariano test statistic to determine whether a candidate forecast delivers a statistically different forecast from the AR model forecast.[17,18]

4.1 Model-based Approaches

Table 1 presents the relative RMSFEs for the current quarter and the next 8 quarters for all the model-based forecasts over the full forecast evaluation period, 1976M1-2010M12.[19] In general, the flows-based approaches (FLOW-VAR, VAR, FLOW-UC, and FLOW-UC|VAR) significantly outperform the AR benchmark in the near term and maintain a relative advantage in forecasting accuracy throughout all horizons we evaluate. Moreover, they outperform the asymmetric autoregressive models (GAR and SETAR) in terms of RMSFEs, and in some instances, reduce forecasting error by roughly 20 percent. These results stand in stark contrast to the earlier literature, especially Montgomery et al.

[13] Note that this means that for the applications of the FLOW-UC model where the aggregate variable driving the cycle, Y, is quarterly data on GDP, not every month will expand the information set. For instance, from the third month of a quarter to the first month of the next quarter, there is no new quarterly GDP information, but the early vintages of the data might be subject to revisions.

[14] This leaves us with a consistent sample of 420 monthly observations. For the sample prior to the Great Recession, which we explore in the robustness section, the sample size is 360.

[15] We also suspect that the benefit of uncovering the long-run trends in the unemployment flows data will become apparent as the forecasting horizon increases.

[16] When we estimate the extensions of the FLOW-UC model on monthly basis, we aggregate the monthly forecasts to produce quarterly unemployment rate forecasts and compare them to data counterparts at a quarterly frequency.

[17] The HLN test statistic is a variant of the Diebold and Mariano (1995) test that employs a rectangular kernel to estimate the long run error variance and adjusts the t-test statistic by $\sqrt{(n+1-2*t+(t*(t-1))/n)n}$. The simple Diebold-Mariano test statistic uses a Bartlett kernel and $h-1$ lags.

[18] Clark and McCracken (2011) highlight some issues with various equality-of-prediction test statistics. Therefore, we also computed the Diebold-Mariano and Andrews-Monahan (1991) test statistics. The results were qualitativly similar to the HLN test; hence, we do not report them.

[19] For a more direct comparison, we report the absolute RMSFEs in Table 2. We will reference these later in the paper.

(1998) and Rothman (1998), as they highlighted the forecast improvements made by asymmetric auto-regressive models. The results of this exercise highlight the usefulness of the flow rates for forecasting purposes. These flows have different time-series properties from the underlying stock itself, therefore enabling us to capture the asymmetric dynamics in a natural way.

The Barnichon and Nekarda (2012) variants, FLOW-VAR and VAR, both significantly improve on the AR benchmark in the near term (through 2 quarters ahead). In comparing the FLOW-UC forecast to the FLOW-VAR approach, it appears that the superior forecasts from the VAR and FLOW-VAR dissipate and as the forecast horizon increases the FLOW-UC model's performance improves. This pattern suggests that pinning down the longer-run trends in the flows data can be useful in forecasting the unemployment rate over longer time horizons. Estimated unobserved trend levels for the underlying flow rates constrain the movement of the unemployment rate while still keeping its inertial nature due to the basic equation of motion for the unemployment rate.

While, alone, the FLOW-UC (trends-based) forecasting approach fails to deliver a significant reduction in forecast error relative to the AR benchmark, when we condition the FLOW-UC on the "nowcast" of flow rates from the VAR specification, the forecasting accuracy improves dramatically. This hybrid specification, FLOW-UC|VAR, which addresses the informational disadvantage of the FLOW-UC model, attains the absolute minimum RMSFE for six quarters, ($t+3$ through $t+8$), and significantly outperforms the AR benchmark through the first 5 forecast horizons. Still, all of the model-based approaches we consider fail to significantly outperform the AR benchmark throughout longer forecast horizons ($t+5$ through $t+8$).

It is interesting to observe from Table 1 that the FLOW-VAR forecast never attains the absolute minimum RMSFE among the specifications we consider, even when we ignore the hybrid case, FLOW-UC|VAR. For all near-term forecast horizons (up to 3-quarters ahead), the basic VAR forecast yields the smallest RMSFE. Relative to the VAR forecast, the only value-added for the FLOW-VAR model is the use of the non-linear equation of motion. Barnichon and Nekarda (2012) report this to be the main reason behind the superior performance of this case in comparison to the GB and SPF forecasts. We defer the full evaluation of the forecast performance relative to professional forecasters to the next section; however, results in Table 1 clearly show that the equation of motion is *not* the main contributor *per se*. As long as one uses unemployment flow rates (as in VAR), the equation of motion for the unemployment rate does not seem to have more information content to improve the forecast accuracy.

We are intrigued by the overall performance of the simple VAR in Table 1, and especially curious about the influence of the leading indicators on the VAR's performance. Relative to the autoregressive approaches and the FLOW-UC model, these leading indicators provide more contemporaneous information. Moreover, since data on these variables are essentially quasi-real time, it might give the model some ex-post advantage. To explore these issues, Table 3 reports the *absolute* RMSFEs for alternative specifications of the VAR and FLOW-VAR forecasting models. We investigate three alternative specifications to disentangle the relative performance of the leading indicators included in the benchmark VAR in (9). The second and the sixth columns in Table 3 correspond to the benchmark VAR and the FLOW-VAR, and repeat what is already reported in Table 2. Then we present the results for variants of the model with specific exclusions from the VAR; one column excludes the Help Wanted Index (HWI), and the other column reintroduces the HWI but excludes the data on unemployment initial claims (UIC). Finally we also report the RMSFEs from a VAR that only includes the lagged unemployment rate and the flows series. Rather than comparing to the AR benchmark as is the case throughout most of the paper, the benchmarks for the equality-of-prediction (HLN) tests are the VAR on the left panel and the FLOW-VAR on the right panel. This comparison is intended to further highlight the informational content of the leading indicators. All the results cover the full sample period.

In general, the results suggest that the leading indicators are vital to the success of the overall VAR, at least in the near term, where the FLOW-VAR model holds its greatest advantage over the SPF and Fed's Greenbook according to Barnichon and Nekarda (2012). This is true whether one uses the unemployment rate forecasts directly from the VAR specification or the FLOW-VAR model using the added equation of motion for unemployment. Excluding both the HWI and initial claims data from the VAR leads to a statistically significant deterioration in forecasting performance relative to the VAR benchmark, for t through $t+6$. It also appears that the data on initial claims are more useful than the HWI to the forecasting performance. This result is very much in line with the work of Montgomery et. al. (1998). They find that initial claims help improve the forecast accuracy for the univariate linear models, especially around business cycle contractions.

Interestingly, excluding the HWI appears to worsen the performance of the VAR in the near term, but improves it in the out-quarters. This pattern is robust to the various sample periods we employ. While its cause is a puzzle, it does suggest a potential gain between using leading indicators to pin down the near-term forecasts and allowing information on trends in flows to dominate in the out-quarters.

Note also that ignoring both the HWI and the UIC, and relying only on the flows data and $\Delta \ln u_t$, not only significantly increases the RMSFEs relative to the respective benchmarks using either approach, but also brings the performance in both VAR and FLOW-VAR closer to each other.[20] This feature once again highlights our doubt about the usefulness of utilizing the equation of motion for forecasting in Barnichon and Nekarda (2012). The VAR specification uses the same information but does not infer the unemployment rate from the equation of motion, yet we find the forecasting performance between the two approaches to be nearly indistinguishable.

So far, our findings suggest: a) models that leverage flow rates perform better than univariate autoregressive models, even when they are allowed to have an explicit asymmetric behavior over the cycle, b) among models with flow rates, the VAR has the best near-term performance, which is mostly driven by the quasi-real time data on UIC and HWI, c) incorporating a good "nowcast" for the flow rates into FLOW-UC improves its performance significantly on all horizons, making it the best alternative for longer horizons. In the next section we compare the most successful approaches to the forecasts of professionals.

4.2 Model-based Approaches versus Professional Forecasts

Having explored the forecast performance across different sets of models, the next natural task is to understand whether professional forecasts fare better relative to the model-based approaches. Our comparison includes comparing the RMSFEs for each one relative to the AR counterpart. To save some space and redundant discussion, we omit the GAR and SETAR models as well as the FLOW-VAR. Recall from Table 1 that the former two do not perform significantly better than the AR model at any horizon, and the latter one does not improve over the basic VAR. We keep the novel approaches, the FLOW-UC and the FLOW-UC|VAR forecasts, in our comparison. Hence, we compare the forecasts from FLOW-UC, FLOW-UC|VAR, and VAR to a particular professional forecast each time.

In each case, we also have a FLOW-UC model forecast that is conditioned on the GDP growth path implied by the professional forecast. This alternative provides a different forecast for the aggregate output in the FLOW-UC model from the one that is implied by the estimated model. We think that this could potentially improve the forecast for the unemployment in the FLOW-UC model too, as professional forecasters use a large set

[20] Over the full sample, the RMSFEs for these two specifications were nearly identical (differing only in the out quarters and by less than 4 basis points). More formally, the results of the HLN equality-of-prediction test yielded no statistically distinguishable differences between these competing forecasts.

of information to generate their GDP forecasts, which the FLOW-UC model naturally ignores. These conditional forecasts are referred to as FLOW-UC|GB, FLOW-UC|SPF, and FLOW-UC|BC. For each professional forecast, whether it is from the Federal Reserve Board staff (GB), the Survey of Professional Forecasters from the Federal Reserve Bank of Philadelphia (SPF) or the Blue Chip panel of economists (BC), we pick the corresponding calendar time for the model forecast, making sure that the information set is the same. Because of the timing issues related to each one, the comparison sample varies depending on the professional forecast we have at hand.

Table 4 presents the results of this forecast evaluation in terms of relative RMSFEs for each set of professional forecasts included. The top panel in Table 4 shows the RMSFEs for the GB sample relative to the forecast accuracy of the benchmark AR model. Our sample contains 114 observations for the t through $t+4$ horizons, but that rapidly declines as the horizon increases, leaving us with only 20 data points for the $t+8$ forecast horizon. The SPF and BC samples do not provide us with more than 4-period- and 7-period-ahead forecasts, respectively.

Several important points stand out from Table 4. We start by observing that for all relevant professional forecast samples, the VAR attains the lowest RMSFE for the current period and one-quarter-ahead forecast horizon. It is also statistically significantly different from the benchmark AR forecast for those horizons, by almost 20 percent. Beyond 1-quarter ahead though, the SPF and GB forecasts both beat the alternatives, judged by the relative RMSFE. They are not all significant.

The primacy of the Greenbook forecast–at least in terms of relative forecasting error–can be seen in Table 4. In the current quarter, the GB carries a RMSFE that is 12 percent less than the simple AR model, and that gap widens to roughly 35 percent at the 4- and 5-quarter-ahead horizons. Interestingly, despite the seemingly tremendous gains in forecast accuracy, the HLN test fails to find significant differences in forecast accuracy. This is due to some wild misses (relative to the AR model) in the late 1970s and early 1980s on the part of the GB.[21] Another important observation from Table 4a is the lack of improvement in the FLOW-UC model from conditioning. Conditioning the FLOW-UC model with the GDP growth forecast from the GB does not improve the unemployment rate forecast accuracy relative to GB itself. Moreover, it does not improve the forecast accuracy of the baseline FLOW-UC model, either.

The SPF sample is about 10 percent longer than the GB sample and occurs at a somewhat more regular frequency (the second month of every quarter). The superior

[21] We suspect that this is probably due to an adherence to an implicit Phillips curve-approach in their thinking at a time which we now characterize as a "stagflationary" period.

performance of the VAR approach for the immediate near term is confirmed for the SPF sample. As the middle panel in Table 4 shows, this performance changes with 2-quarter-ahead forecast and beyond. In addition, the SPF is about 20 percent more accurate than the benchmark AR model throughout the forecast horizon, but only significantly outperforms it at the 10 percent threshold over the $t+2$ through $t+4$ horizon. Using the output growth forecasts from the SPF improves the forecast accuracy for the FLOW-UC model, but not beyond the baseline SPF forecasts for unemployment. This stands in contrast to the rather poor performance of the FLOW-UC|GB case earlier.

The bottom panel of Table 4 reports the results for the BC sample and highlights an interesting challenge for the BC consensus forecast. Contrary to the GB and SPF samples, the Blue Chip consensus forecast never attains the minimum relative RMSFE for any horizon. The VAR model continues to provide the best forecast for the near term, and beyond 3 quarters ahead, the FLOW-UC|BC yields the best forecast. Curiously, the BC's current-quarter forecast for the unemployment rate is exceptionally bad relative to the AR benchmark and is significantly different at the 1 percent threshold.

Results in Table 4 once again confirm the exceptionally good forecast accuracy of the basic VAR model for the near term, which utilizes unemployment flow rates. Beyond the 1-quarter-ahead horizon though, professional forecasts start to catch up and gradually outperform the VAR model. This is especially true for the GB and the SPF cases. The BC sample provides the exception for us, where the BC forecasts for longer horizons do not seem to outperform the model alternatives.[22]

5 Forecast Combination

So far our results indicate that the VAR model delivers the most accurate forecasts for up to 2 quarters ahead, and the FLOW-UC model presents the most potential for the farther horizons, especially when conditioned on a good "nowcast" as in FLOW-UC|VAR (see Table 1). However, as the length of the forecast horizon increases, besting the simple AR benchmark becomes more tenuous. In this section, we investigate whether certain combinations of forecasts can yield an improved prediction of the unemployment rate. In particular, we want to understand whether particular forecast combinations can beat the best single forecast for different horizons. Forecast combination is fairly common in the literature and, as Wright (2003) puts it, "...the basic idea that forecast combination outperforms any individual forecast is part of the folklore of economic forecasting, going

[22] However, conditioning the FLOW-UC model on the consensus forecasts from the Blue Chip panel does yield the lowest RMSFEs from the $t+3$ to the $t+8$ horizon.

back to Bates and Granger (1969)."

As in the previous section, we look at the model-based forecast combinations and the alternatives that contain professional forecasts separately. We are particularly interested in combinations that include the AR model (which seems to be statistically indistinguishable from the best alternative for some horizons), the VAR model (which yields the best near-term forecast), the FLOW-UC model (which has desirable longer-run forecasting properties), and professional forecasters (which can leverage the collective wisdom and judgment of a large group of economists).

First, we analyze the performance of the forecast combinations for a variety of model-based configurations. Since the AR model performed as well, if not better than the GAR and SETAR models, we ignore these latter two in our simple forecast combinations. Similarly, since the FLOW-VAR model did not improve over the VAR model, we do not highlight it either. These restrictions let us economize on the number of different combinations. Hence, we present results for three different combinations: AR/VAR, AR/FLOW-UC, and AR/VAR/FLOW-UC. All the combinations are generated using regression weights.[23] In addition, we created a particular forecast combination that includes all the model-based forecasting frameworks, including the GAR, SETAR, FLOW-VAR, and the FLOW-UC|VAR. We present two different combinations of this case, one where the forecasts are weighted by regression coefficients, REG_W, and one where they are summed up using equal weights, EQ_W. These forecast combinations are presented in Tables 6 and 7, along with the configuration we call the BEST. BEST refers to the best relative RMSFE in Table 1 (or best absolute RMSFE in Table 2) for a particular forecast horizon and effectively means the VAR forecast for the first three forecast periods and the FLOW-UC/VAR forecast for the rest of the forecast horizon.[24]

Table 6 reveals that combining different forecasts from a variety of models improves the forecast performance. Combinations that include simultaneously both AR and VAR models yield lower RMSFEs relative to the best alternative single-model forecast. Using all the alternative models according to their regression weights incorporates different comparative advantages and yields smaller RMSFE for every forecast horizon. The improvements gradually increase from around 1 percent to more than 7 percent as the forecast horizon increases, albeit the gains in accuracy were only significant at the $t + 3$ horizon for the AR/VAR/FLOW-UC combination and the REG_W combination. We also ran forecast-encompassing tests for these forecast combinations and found that, in

[23] Available upon request.

[24] We present the RMSFEs in absolute terms in this section. We use the lowest RMSFE forecasts (denoted BEST) as the benchmark in the HLN equality-of-prediction tests in Table 6. As the benchmark comparison forecast in Table 7, we use the appropriate professional forecast.

almost every instance, there is NOT one forecast that encompasses the combined result. The exception is the VAR forecast for the current period. In other words, for the forecast combination cases AR/VAR and AR/VAR/FLOW-UC, the VAR model's forecast is not statistically different from the combined forecast; hence the other models do not add any forecasting improvement.[25]

We should also note that, while the equally weighted (EQ_W) forecast combination never attains the minimum RMSFE in our tests, it significantly outperforms the BEST approach over forecast horizons greater than 1 year. We attribute this to the consistent nature in which EQ_W outperformed the BEST benchmark. The forecast errors stemming from the REG_W combination suffered from a slightly more volatile performance relative to the EQ_W combination.

We can conduct a similar analysis for forecast combinations involving the professional forecasts. Admittedly, these professional forecasts (GB, SPF, and BC), are already representative of a form of collective judgement. However, the success of each individual source of professional forecasts we discussed in section 4.2 was mixed. Thus, we think forecast combination might still improve forecast accuracy over and beyond those reported in Table 4 or Table 5. To gauge whether this is the case, we follow a similar approach and look for combinations with one professional forecast at a time. Once again, to economize on the number of potential combinations, we focus on pairs with the AR, VAR, and FLOW-UC models as well as regression-weighted and equal-weighted specifications, where we include all model-based forecasts and the professional forecast in question. For instance, for the GB sample this implies looking at the following alternatives: GB (benchmark), GB/AR, GB/VAR, and GB/FLOW-UC. In addition, we have REG_W and EQ_W, combining alternative specifications with the regression and equal weights, respectively.[26]

Table 7 presents our results for this exercise and highlights the benefits of forecast combination for unemployment predictions, even for professional forecasts. Combining the forecasts of professionals with model-based approaches, especially the AR or VAR model, improves the forecast accuracy. In general, combining the VAR model with professional judgment appears to yield a statistically signifcant improvement in forecasting accuracy in the near term.

[25] Forecast-encompassing results available upon request.

[26] Note that in this case, regression weights and equal weights also include the GAR, SETAR, FLOW-VAR and FLOW-UC|VAR models as well.

6 Business Cycle Turning Points

In this section we present some evidence on the forecasting performance of the models at different points in the business cycle. Figure (1) plots forecasts from the FLOW-VAR and FLOW-UC models as well as the actual data around the turning points of the business cycle during the last three episodes. Univariate models and the VAR model are omitted, as they do not perform better than the ones presented. We first present the picture for the 1990-91, 2001, and Great Recession periods. These three jobless-recovery episodes present a challenge for the forecasting models. The upper panel shows model forecasts for the recessionary episode, starting from the two quarters prior to the (ex-post) start of the recession as dated by the NBER. The lower panel repeats the same exercise for the subsequent recoveries, where recovery start dates follow NBER dates for the respective troughs.

Figure (1) makes it tragically clear that these models do not perform well during recessions. That said, the FLOW-VAR model (and, thus, the VAR model) picks up the relatively sharp increases in the unemployment rate from $t-1$ to $t+2$ in the 2001 recession and performs admirably. However, during other peaks, the FLOW-VAR model performs as poorly as the FLOW-UC model. The forecasting performance of the FLOW-UC model, which relies on the underlying trends in the flow rates, and the FLOW-VAR are all tightly grouped at business cycle peaks for the first and the last jobless recovery episodes. In both cases, they predict close to no change in the unemployment rate going forward.

When we look at the recovery episodes, the picture changes, somewhat dramatically. For all of the jobless recoveries, the FLOW-UC model picks up on the general dynamics of the unemployment very well. In fact, for the last two jobless recovery episodes, FLOW-UC predicts the unemployment rate evolution incredibly well. For the 1990-91 episode, the FLOW-VAR (hence VAR) model seems to do a better job early on, but the predicted path remains quite persistent even after 6 quarters into the recovery, when the actual unemployment rate turned around. Figure (1) highlights that the FLOW-UC model more closely captures the dynamics of the labor market as the unemployment rate returns to its long-run trend, especially when the cyclical gap in unemployment is substantially large. The FLOW-VAR model misses the last two recoveries entirely. The main reason for this miss is the lack of structure in the VAR specification to discipline the convergence to a well-defined empirically consistent long-run average for the flow rates. The force of the last few observations on the observed flow rates dominates and creates a lot of momentum going forward. Moreover, if there are low-frequency variations in

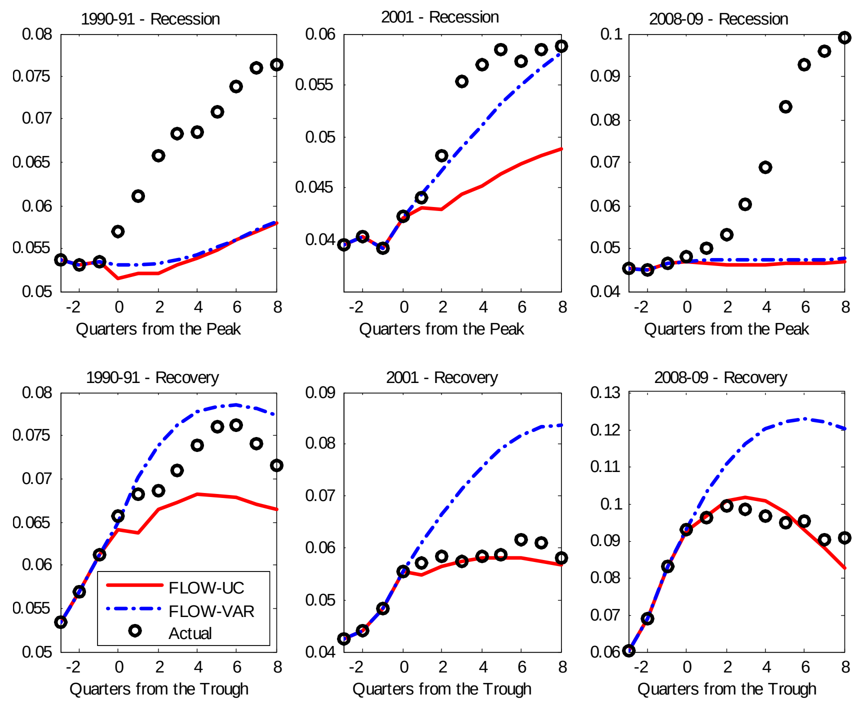

Figure 1: Unemployment forecast from FLOW-UC and FLOW-VAR and the data during the last three recessions.

those trend rates, as Tasci (2012) shows, then the sample average from a 15-year-rolling-window estimation may not be sufficient to pull it back to a sensible new trend value, thereby implying a lot of persistence in the unemployment rate.

Our conclusions do not change for the prior two business cycle episodes in our sample. Figure (2) plots the evolution of the actual unemployment rate and the two model forecasts for the earlier two episodes in our sample. The FLOW-VAR model's success for the rescessionary periods remains mixed for the 1980 and 1981-82 episodes, missing the latter entirely and predicting the early rise in unemployment rate very well for the former. Even though FLOW-UC fails to predict the rescessionary surges, it catches the evolution of the unemployment rate after the local through remarkably well, especially in the 1981-82 episode. Note that the relative failure of the FLOW-UC model in figure (2) to predict the latter part of the recovery in the 1980 episode is due to the beginning of the 1981-82 recession before the forecast horizon ends.

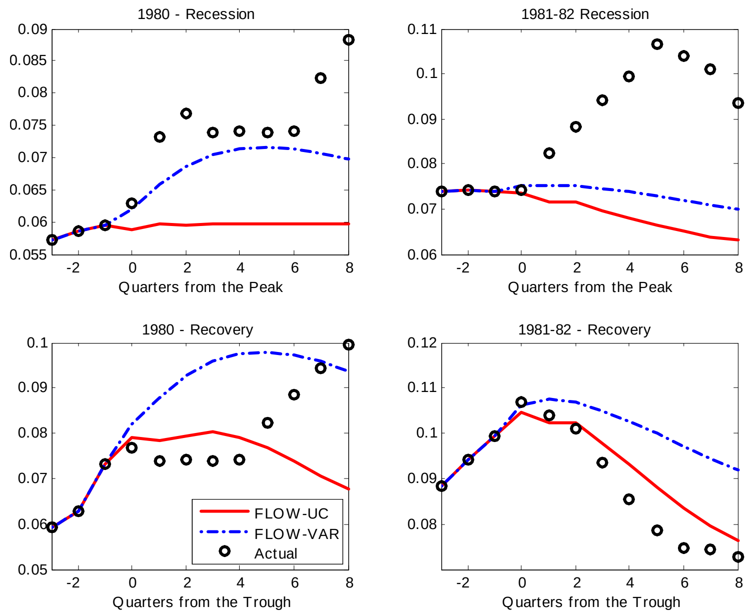

Figure 2: Unemployment forecast from FLOW-UC and FLOW-VAR and the data during the recessions of 1980 and 1981-82.

7 Robustness

Looking into the forecasting performance of the different models and the professional forecasts, we identified several interesting observations. In general, using flow rates improved forecast performance, and incorporating an explicit role for trends helped beyond the short-term horizon. In this section, we address whether our results are robust to alternative sample periods, assumptions about the information set, and the rounding of the forecasts to the nearest tenth.

7.1 Before the Great Recession

The Great Recession and the subsequent recovery constitute only about 12 percent of our model-based forecast evaluation sample, 36 months out of 420. However, the recession and recovery stand out as an episode with exceptionally high and persistent rates of unemployment by historical standards. The unemployment rate sharply increased from 4.7 percent at the end of 2007 to 10 percent in 2009:Q4 and stayed above 8 percent for another 12 quarters until 2012:Q4, the end of the out-of-sample forecast horizon in our exercises. Even though the FLOW-UC model performed exceptionally well for this period, it was not universally the case for other model-based and professional forecasts. Thus, we find it very natural to ask whether the forecasting accuracy changes if we focus only on the sample period prior to the Great Recession. To get at this, we restrict our forecast evaluation sample to the pre-2006 period, making December 2005 our last estimation point for forecast evaluation. Hence, forecast evaluation ends by December 2007, the onset of the Great Recession.

Table 8 presents our results for this exercise and provides a stark comparison to Table 1. All of the models relying on flow rates improve substantially over the AR benchmark beyond the 2-period-ahead forecast horizon. Perhaps most notabe is the performance of the FLOW-UC|VAR model, which significantly outperforms the AR specification over every forecast horizon. In addition, the competition between the VAR and FLOW-UC|VAR modesl tips in favor of the VAR, making it marginally superior across all forecast horizons (though that difference is a mere 4 basis points relative to the FLOW-UC model at the longest horizon ($t+8$). The improvements in the models with flow rates are broad based, relative to the AR benchmark, even at the t+8 forecast horizon. The improvements in RMSFEs in absolute terms are even more pronounced than the relative improvements. For instance, comparing Table 9 to Table 2 shows that all of the models using flow rates (FLOW-VAR, VAR, FLOW-UC, FLOW-UC|VAR) for $t+8$, have RMSFE around 1.20, registering between 30 to 40 percent declines relative to the full sample results. This represents a disproportionately larger drop in the RMSFE than the reduction in the sample size, though given that we're excluding a large recession, these gains aren't all that surprising.

Restricting our sample to the pre-2006 sample period affects the forecast performance of the professional forecasts as well. In terms of sample size, this restriction minimally changes the GB sample (given that Board forecasts are sequestered for 5 years). However, for the BC and SPF, depending on the forecast horizon, this restriction amounts to a reduction of the sample size by 15 to 20 percent. Thus, a priori, we would expect a significant decline in the RMSFE, at least in the levels. A comparison between Table 5 and Table 11 highlights the effects of this sample restriction. We see that uniformly, every forecast in GB, SPF, or BC samples declines, sometimes by a large margin. The largest impact is registered for the GB forecast in the GB sample, prior to 2006. The omission of two forecasts from the GB sample reduces the sample size from 20 to 18 for the 8-quarter-ahead forecast horizon, but leads to a RMSFE of 0.914, down from 1.489. Two observations, each from the October FOMC meetings of 2006 and 2007, stand out as big misses for the GB. The improvement in the RMSFEs relative to the AR model for professional forecasts seems to be somewhat muted compared to the improvement for the model-based forecasts. Comparing Table 10, which includes relative RMSFEs for professional forecasts in the absence of the Great Recession, with Table 4, reveals that even the ranking of the best-forecast framework for each horizon did not change.[27] However, after excluding just 4 observations, the GB forecast was able to significantly outperform the AR benchmark at the 10 percent level at the current quarter and 2 quarter-ahead horizons.

7.2 Differences in the Information Set

We tried to pay a lot of attention to making sure that when we compare the forecasting performance, the information set relied upon in real-time by the different methods we analyze are the same, to the extent possible. One issue we discussed in this vein in the preceding sections is the quasi-real time nature of the VAR and FLOW-VAR approach, as they rely upon UIC and HWI data. We know that UIC is only subject to some seasonal adjustments, but the history of the HWI is not as clear. Our discussion in section 4.1 highlighted the major contribution of these leading indicators to the performance of the VAR models.

When comparing forecasts across alternative frameworks and sources, we discovered another issue about the information set as it pertains to the GB sample. In particular, we found out that, in 13 different GB releases, the employment report from the BLS for the prior month would not have been available. In each instance, the BLS report came out couple of days later.[28] To the extent that Board staff incorporated a "good" forecast for the yet-to-be-released employment report, the GB forecast may not suffer from this informational disadvantage. All of our analysis so far accounted for this discrepancy between the information sets and focused on the smaller sample when we referred to the GB sample results.[29] Table 12 presents a simple comparison between these two samples, and tries to understand the extent of the potential bias. Even though the GB's forecast performance mostly improves qualitatively in the "same information" sample,[30] it is far from being significant.

7.3 Rounding

In our computations of the unemployment forecast from different models as well as the data, we do not round the resulting numbers to the nearest tenth, which is the case

[27] The only exception is the $t+3$ horizon forecast in the BC sample.

[28] Out of those 13 releases, only 9 had data available at the $t+6$ forecast horizon, 6 had data at the $t+7$ horizon, and 2 had data at $t+8$ forecast horizon.

[29] The Barnichon and Nekarda (2012) results seem to focus on the larger sample, where there is a difference between the information sets of the GB and the FLOW-VAR or VAR models.

[30] With the exception of the $t+1$ and the $t+8$ forecast horizons.

for the professional forecasts and the published unemployment series. We opted for not rounding, since the equation of motion we use for forecasting, (8), will be affected by rounding, adding some potential spurious bias. Moreover, since we rely on the real-time data on the actual levels of unemployment and the labor force, we can pin down the official unemployment rate to a higher precision. However, this potentially can distort the comparisons between model-based forecasts and the professional forecasts as well as creatie larger wedge between the official and model-generated unemployment forecasts.

To understand the potential bias, consider the following naive example. Suppose for every forecast, the rounding for the model brings the resulting forecast "up" one-tenth relative to the actual one: i.e., a model forecast of 9.56 versus an actual one of 9.44. Then the RMSFE will be 0.20 with rounding and 0.12 without rounding, potentially creating a significant difference. Certainly, as we aggregate these differences across longer forecast horizons, this bias will diminish, as rounding becomes less of a concern farther out. And this happens to be exactly what we find in our analysis of the effects of rounding. We present the RMSFEs for the model forecasts and professional forecasts in Table 13 and Table 14, respectively. In the latter table, only the data we compare to are different relative to the baseline example without rounding (Table 5). Table 13, on the other hand, reports both the computed model forecasts and the respective data counterparts with rounding. It shows that rounding leads to a deterioration in the current-period and 1-quarter-ahead forecast horizons, though not significantly. Similarly, rounding seems to lower forecast accuracy in the near term for the professional forecasts as well. For both cases, the effects of rounding beyond $t+1$ are close to non-existent.

8 Conclusion

Barnichon and Nekarda (2012) demonstrates the usefulness of unemployment flows in forecasting the unemployment rate, yielding forecasts that significantly outperform private forecasters and in some cases the Federal Reserve Board's "Greenbook" forecast in the near term. In this paper, we explore the forecasting performance of another model that leverages unemployment flows data -Tasci (2012)- in addition to several univariate time series models. We find that this approach performs about as well as Barnichon and Nekarda's forecasting model, which employs an equation of motion for the unemployment rate to leverage VAR forecast output on the job-finding and separation rates. While the difference in forecasting performance is, on balance, modest, the more structured trends-based approach of Tasci (2012) is parsimonious and doesn't require additional contemporaneous information on leading indicators.

Interestingly, we find that, in a broad sense, it really doesn't matter how you leverage the data on worker flows, as long as you use them to begin with. The rolling-window estimates of the VAR model, which uses data on flows, the unemployment rate, and a few leading indicators of the labor market, performs as well (if not better) than the Barnichon-Nekarda's (2012) baseline approach. In other words, the main innovation in Barnichon and Nekarda (2012) of using estimated flow rates from a VAR in the unemployment equation of motion, does not improve the accuracy *per se*. There is also evidence that a significant fraction of the forecast improvement is attributable to the use of leading indicators, especially the initial claims of unemployment benefits.

We find some modest support for Tasci (2012) as a forecasting model over longer, increasingly more policy-relevant, time horizons. Importantly, conditioned on emerging from a recession, it outperforms other flows-based methods. This is due to the ability of a trends-based approach to more realistically capture the dynamics of the labor market as it begins to normalize. This also implies that there is a drawback to relying on a reduced-form VAR to yield forecasts of unemployment flows, as forecasted paths around turning points can wildly persist away from their longer-run steady-state levels. We also found strong support for combining alternative forecasts to improve forecast accuracy. Forecasts that combine data on worker flows allow for trend dynamics in the unemployment rate, and those that leverage professional judgment are particularly accurate. Also, while forecast combinations that leveraged regression weights, in general, tended to have a slightly lower RMSFEs than simple averages, the difference between the two approaches wasn't economically meaningful. Finally, we show that the results are robust to several practical data issues.

Overall, our results point to a potentially successful forecasting strategy for the unemployment rate: Employ leading labor market indicators in addition to the flow rates in the near term, and allow for trend dynamics in job-finding and separations rates to influence the forecast over the longer term.

References

[1] Abraham, Katharine and Robert Shimer. 2001. "Changes in Unemployment Duration and Labor-Force Attachment," in *The Roaring '90s: Can Full Employment be Sustained?* ed. Alan B. Krueger and Robert Solow, 3–60. New York: Russell Sage Foundation and Century Foundation.

[2] Andrews, Donald W.K. and J. Christopher Monahan. 1992. "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator." *Econometrica* 60: 953-966.

[3] Barnichon, Regis. 2010. "Building a composite Help-Wanted Index" *Economics Letters*, 109(2010): 175-178.

[4] Barnichon, Regis and Christopher J. Nekarda. 2012. "The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market." *Brookings Papers on Economic Activity*, Fall: 83:117.

[5] Bates, J.M and C.W.J. Granger. 1969. "The Combination of Forecasts." *Operations Research Quarterly,* 20: 451-468.

[6] Bernanke, Ben S. 2012. "Recent Developments in the Labor Market." Speech delivered at NABE Annual Meetings Washington D.C., March 26. Board of Governors of the Federal Reserve System.

[7] Clark, Peter K. 1987. "The Cyclical Component of U.S. Economic Activity." *Quarterly Journal of Economics*, 102(4): 797–814.

[8] Clark, Peter K. 1989. "Trend Reversal in Real Output and Unemployment." *Journal of Econometrics*, 40: 15–32.

[9] Clark, Todd E. and McCraken, Michael W. 2011. "Advances in Forecast Evaluation." *Federal Reserve Bank of Cleveland Working Paper, No. 11-20.*

[10] Diebold, Francis X., and Roberto S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics*, 13: 253-263

[11] Elsby, Michael, Ryan Michaels, and Gary Solon. 2009. "The Ins and Outs of Cyclical Unemployment." *American Economic Journal: Ma*

[12] Elsby, Michael W.L., Bart Hobijn and Aysegul Sahin. 2011. "Unemployment Dynamics in the OECD." *Review of Economics and Statistics*, Forthcoming.

[13] Fujita, Shigeru and Gary Ramey. 2009. "The Cyclicality of Separation and Job Finding Rates." *International Economic Review*, 50(2): 415–430.

[14] Giacomini, R., and White, H. 2006. "Tests of Conditional Predictive Ability." *Econometrica*, 74: 1545-1578.

[15] Harvey, D. I., Leybourne, S. J., & Newbold, P. 1998. "Tests for Forecast Encompassing." *Journal of Business and Economic Statistics*,16(2): 254-259.

[16] Kim, Chang-Jin and Charles R. Nelson. 1999. *State-Space Models with Regime Switching*. Cambridge, MA: MIT Press.

[17] Kocherlakota, Narayana. 2010. "Inside the FOMC." Speech delivered in Marquette, Michigan., August 17. Federal Reserve Bank of Minneapolis.

[18] Montgomery, Alan L, Victo Zarnowitz, Ruey S. Tsay and George C. Tiao. 1998. "Forecasting the U.S. Unemployment Rate." *Journal of the American Statistical Association*, 93(442): 478:493.

[19] Mortensen, Dale and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies*, 61(3): 397–415.

[20] Polivka, Anne E. and Stephen M. Miller. 1998. "The CPS after the Redesign: Refocusing the Economic Lens," in *Labor Statistics Measurement Issues*, ed. John Haltiwanger, Marilyn E. Manser, and Robert Topel, 249–286. Chicago, IL: University of Chicago Press.

[21] Rothman, Philip. 1998. "Forecasting Asymmetric Unemployment Rates." *Review of Economics and Statistics*, 80(1): 164-168.

[22] Shimer, Robert. 2005. "The Cyclical Behavior of Unemployment and Vacancies: Evidence and Theory." *American Economic Review*, 95(1): 25–49.

[23] Shimer, Robert. 2012. "Reassessing the Ins and Outs of Unemployment." *Review of Economic Dynamics*, 15(2): 127–148.

[24] Tasci, Murat. 2012. "Ins and Outs of Unemployment in the Long-Run: Unemployment Flows and the Natural Rate." *Working Paper #12-24*, Federal Reserve Bank of Cleveland.

[25] Wright, Jonathan H. 2003. "Bayesian Model Averaging and Exchange Rate Forecasts." *International Finance Discussion Papers No. 779,* Board of Governors of the Federal Reserve System.

Tables

The tables below report RMSFEs either relative to a benchmark forecast or in absolute terms. In every table except for tables 3, 6, and 7, the benchmark forecast for the relative forecasting performance comparison and for the equality of prediction tests is the AR (simple autoregressive) model. The lowest RMSFE in each horizon is reported in bold font. "***", "**", and "*" denote statistically significant differences in forecasting accuracy at the 1 percent, 5 percent, and 10 percent levels, respectively, using the HLN equality of prediction test. The benchmark forecast for Table 3 is the VAR model (on the left-hand side) and FLOW-VAR model (on the right-hand side). For table 6, the benchmark forecast is the BEST specification (which is the lowest RMSFE for each forecasting horizon from the set of model-based forecasts in Table 1). For table 7, the benchmark forecast is the professional forecasts that correspond to the appropriate panel.

Table 1: Relative RMSFEs – Model Forecasts, 1976:M1-2010:M12

Horizon	FLOW-VAR	VAR	FLOW-UC	FLOW-UC|VAR	GAR	SETAR
t	0.809***	**0.791*****	0.975	0.809***	1.005	0.996
$t+1$	0.845**	**0.815*****	1.049	0.837***	1.036	0.994
$t+2$	0.866*	**0.851****	0.942	0.866**	1.041	0.992
$t+3$	0.891	0.873	0.917	**0.863****	1.026	0.997
$t+4$	0.919	0.900	0.909	**0.879***	1.002	1.022
$t+5$	0.943	0.920	0.912	**0.896**	0.982	1.059
$t+6$	0.958	0.934	0.923	**0.913**	0.970	1.103
$t+7$	0.975	0.948	0.935	**0.930**	0.961	1.163
$t+8$	0.993	0.966	0.949	**0.949**	0.959	1.223

Table 2: Absolute RMSFEs – Model Forecasts, 1976:M1-2010:M12

Horizon	FLOW-VAR	VAR	FLOW-UC	FLOW-UC|VAR	AR	GAR	SETAR
t	0.157***	**0.153*****	0.189	0.157***	0.194	0.195	0.193
$t+1$	0.370**	**0.357*****	0.459	0.366***	0.438	0.453	0.435
$t+2$	0.598*	**0.588****	0.650	0.598**	0.690	0.718	0.685
$t+3$	0.834	0.817	0.858	**0.807****	0.936	0.960	0.933
$t+4$	1.061	1.039	1.050	**1.016***	1.155	1.157	1.181
$t+5$	1.266	1.236	1.225	**1.204**	1.343	1.319	1.423
$t+6$	1.430	1.394	1.377	**1.363**	1.493	1.447	1.647
$t+7$	1.573	1.530	1.510	**1.501**	1.614	1.552	1.877
$t+8$	1.690	1.644	1.615	**1.615**	1.702	1.633	2.081

	VAR				FLOW-VAR			
Forecast Horizon	All	No HWI	No UIC	Only Flows	All	No HWI	No UIC	Only Flows
t	**0.153**	0.161*	0.177***	0.201***	**0.157**	0.163**	0.177***	0.201***
$t+1$	**0.357**	0.370	0.414**	0.460***	**0.370**	0.382	0.419**	0.461***
$t+2$	**0.588**	0.597	0.660**	0.710***	**0.598**	0.606	0.661*	0.705**
$t+3$	**0.817**	0.824	0.894	0.946**	**0.834**	0.841	0.898	0.942*
$t+4$	**1.039**	1.041	1.114	1.167*	**1.061**	1.064	1.118	1.162*
$t+5$	1.236	**1.231**	1.323	1.377*	1.266	**1.264**	1.327	1.369
$t+6$	1.394	**1.387**	1.494	1.559*	1.430	**1.427**	1.496	1.546
$t+7$	1.530	**1.519**	1.655	1.736	1.573	**1.568**	1656	1.716
$t+8$	1.644	**1.627**	1.802	1.902	1.690	**1.679**	1.797	1.868

Table 3: Leading Indicators - Absolute RMSFEs - 1976:M1-2010:M12

Table 4: Professional Forecasts - Relative RMSFEs - 1976:M1-2010:M12

	a) GB Sample					
Horizon	GB	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|GB	Size
t	0.876	**0.806***	0.953	0.818***	0.920	114
$t+1$	0.831	**0.793**	1.001	0.795**	0.966	114
$t+2$	**0.748**	0.791	0.860	0.812	0.772	114
$t+3$	**0.693**	0.788	0.840	0.795	0.805	114
$t+4$	**0.666**	0.787	0.823	0.798	0.815	114
$t+5$	**0.666**	0.816	0.840	0.826	0.826	98
$t+6$	**0.727**	0.877	0.912	0.879	0.920**	71
$t+7$	**0.757**	0.916	0.937	0.929	0.875	45
$t+8$	**1.075**	1.177	1.103	1.114	1.082	20
	b) SPF Sample					
Horizon	SPF	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|SPF	Size
t	0.909	**0.811***	1.002	0.853***	0.941	140
$t+1$	0.861	**0.828**	1.021	0.857*	0.923	140
$t+2$	**0.813***	0.852	0.919	0.876	0.820	140
$t+3$	**0.800***	0.865	0.901	0.877	0.812	140
$t+4$	**0.806***	0.885	0.898	0.887	0.817	140
	c) BC Sample					
Horizon	BC	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|BC	Size
t	1.135*	**0.809***	0.997	0.826***	0.937	371
$t+1$	0.960	**0.841**	1.094	0.871**	1.027	371
$t+2$	0.918	**0.896**	0.995	0.910	0.927	371
$t+3$	0.896*	0.922	0.964	0.904	**0.895**	371
$t+4$	0.886	0.948	0.950	0.920	**0.866**	367
$t+5$	0.889	0.950	0.947	0.930	**0.839**	273
$t+6$	0.897	1.001	0.972	0.953	**0.840**	180
$t+7$	0.881	0.986	0.966	0.962	**0.821**	90

Table 5: Professional Forecasts - Absolute RMSFEs - 1976:M1-2010:M12

	a) GB Sample					
Horizon	GB	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|GB	Size
t	0.164	**0.151*****	0.179	0.153***	0.172	114
$t+1$	0.333	**0.318****	0.401	0.319**	0.387	114
$t+2$	**0.478**	0.505	0.550	0.518	0.493	114
$t+3$	**0.596**	0.678	0.723	0.684	0.692	114
$t+4$	**0.712**	0.840	0.879	0.852	0.871	114
$t+5$	**0.810**	0.993	1.023	1.006	1.005	98
$t+6$	**0.957**	1.155	1.201	1.157	1.212**	71
$t+7$	**1.102**	1.333	1.364	1.352	1.273	45
$t+8$	**1.489**	1.631	1.528	1.543	1.499	20
	b) SPF Sample					
Horizon	SPF	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|SPF	Size
t	0.150	**0.133*****	0.165	0.140***	0.155	140
$t+1$	0.359	**0.345****	0.426	0.357*	0.385	140
$t+2$	**0.550***	0.577	0.622	0.593	0.555	140
$t+3$	**0.749***	0.810	0.844	0.822	0.761	140
$t+4$	**0.934***	1.025	1.041	1.027	0.947	140
	c) BC Sample					
Horizon	BC	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|BC	Size
t	0.216*	**0.154*****	0.190	0.158***	0.179	371
$t+1$	0.410	**0.359****	0.467	0.372**	0.439	371
$t+2$	0.607	**0.593**	0.658	0.602	0.614	371
$t+3$	0.810*	0.833	0.871	0.818	**0.809**	371
$t+4$	1.000	1.071	1.072	1.039	**0.977**	367
$t+5$	1.180	1.262	1.257	1.234	**1.114**	273
$t+6$	1.295	1.446	1.403	1.377	**1.213**	180
$t+7$	1.421	1.590	1.559	1.551	**1.323**	90

Table 6: RMSFEs –Full Sample (1976-2010) Model Forecast Combinations

Horizon	BEST	AR/VAR	AR/FLOW-UC	AR/VAR /FLOW-UC	REG_W	EQ_W
t	0.153	0.153	0.186***	0.153	**0.152**	0.166**
$t+1$	0.357	0.351	0.407**	0.350	**0.348**	0.368
$t+2$	0.588	0.571	0.625	0.568	**0.567**	0.585
$t+3$	0.807	0.787	0.835	0.782*	**0.778****	0.796
$t+4$	1.016	0.991	1.026	0.985	**0.976**	0.991
$t+5$	1.204	1.167	1.194	1.161	**1.143**	1.163*
$t+6$	1.363	1.306	1.336	1.302	**1.278**	1.302**
$t+7$	1.501	1.423	1.455	1.421	**1.395**	1.423**
$t+8$	1.615	1.516	1.547	1.516	**1.491**	1.523*

Table 7: RMSFEs for Forecast Combinations with Professional Forecasts

	a) GB Sample					
Horizon	GB	GB/AR	GB/VAR	GB/FLOW-UC	REG_W	EQ_W
t	0.164	0.151	0.140**	0.151	**0.136****	0.154
$t+1$	0.333	0.305	0.287*	0.314	**0.282***	0.308
$t+2$	0.478	0.451	0.439	0.454	**0.433**	0.484
$t+3$	0.596	0.577	0.572	0.578	**0.559**	0.646
$t+4$	0.712	0.696	0.695	0.696	**0.684**	0.803
$t+5$	0.810	0.807	0.807	0.807	**0.791**	0.946
$t+6$	0.957	0.955	0.956	0.957	**0.939**	1.079
$t+7$	1.102	1.095	1.098	1.099	**1.057**	1.258
$t+8$	1.489	1.355	1.467	1.472	**1.156**	1.469
	b) SPF Sample					
Horizon	SPF	SPF/AR	SPF/VAR	SPF/FLOW-UC	REG_W	EQ_W
t	0.150	0.139*	0.127***	0.142	**0.124*****	0.140
$t+1$	0.359	0.344	0.328**	0.354	**0.323****	0.348
$t+2$	0.550	0.542	0.532	0.549	**0.520**	0.564
$t+3$	0.749	0.743	0.735	0.746	**0.721**	0.784
$t+4$	0.934	0.930	0.926	0.933	**0.902**	0.979
	c) BC Sample					
Horizon	BC	BC/AR	BC/VAR	BC/FLOW-UC	REG_W	EQ_W
t	0.216	0.176***	0.151***	0.180***	**0.149*****	0.162***
$t+1$	0.410	0.381**	0.345***	0.407	**0.341*****	0.365**
$t+2$	0.607	0.585	0.558**	0.605	**0.549****	0.575
$t+3$	0.810	0.795	0.774	0.809	**0.760**	0.791
$t+4$	1.000	0.990	0.977	0.997	**0.949**	1.000
$t+5$	1.180	1.162	1.157	1.172	**1.121**	1.179
$t+6$	1.295	1.273	1.275	1.283	**1.227**	1.307
$t+7$	1.421	1.400	1.399	1.408	**1.348**	1.451

Table 8: Relative RMSFEs – Model Forecasts, Pre-2006

Horizon	FLOW-VAR	VAR	FLOW-UC	FLOW-UC|VAR	GAR	SETAR
t	0.781***	**0.771*****	0.949	0.781***	1.008	0.996
$t+1$	0.797**	**0.773*****	0.988	0.787***	1.047	0.982
$t+2$	0.790**	**0.773****	0.863**	0.804***	1.053	0.980
$t+3$	0.791*	**0.773****	0.830**	0.781**	1.033	0.993
$t+4$	0.801*	**0.784***	0.817**	0.787**	0.999	1.035
$t+5$	0.809	**0.792***	0.809**	0.793**	0.970	1.091
$t+6$	0.809	**0.792**	0.813*	0.801**	0.951	1.155
$t+7$	0.810	**0.795**	0.820*	0.811*	0.940	1.239
$t+8$	0.816	**0.805**	0.829*	0.826*	0.937	1.321*

Table 9: Absolute RMSFEs – Model Forecasts - Pre-2006							
Horizon	FLOW-VAR	VAR	FLOW-UC	FLOW-UC|VAR	AR	GAR	SETAR
t	0.150***	**0.148*****	0.183	0.150***	0.192	0.194	0.192
$t+1$	0.333**	**0.323*****	0.413	0.329***	0.418	0.438	0.410
$t+2$	0.513**	**0.502****	0.561**	0.522***	0.650	0.684	0.636
$t+3$	0.681*	**0.666****	0.716**	0.673**	0.862	0.890	0.856
$t+4$	0.840*	**0.822***	0.857**	0.825**	1.049	1.048	1.086
$t+5$	0.977	**0.955***	0.976**	0.956**	1.207	1.171	1.316
$t+6$	1.075	**1.053**	1.081*	0.065**	1.330	1.265	1.536
$t+7$	1.155	**1.134**	1.170*	1.157*	1.426	1.340	1.768
$t+8$	1.220	**1.202**	1.238*	1.234*	1.494	1.400	1.973*

Table 10: Professional Forecasts - Relative RMSFEs - Pre-2006						
	a) GB Sample					
Horizon	GB	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|GB	Size
t	0.861*	**0.804*****	0.951	0.817***	0.919	110
$t+1$	0.824	**0.789****	0.998	0.792**	0.965	110
$t+2$	**0.739***	0.783*	0.855	0.806*	0.768	110
$t+3$	**0.679**	0.773	0.830	0.784	0.801	110
$t+4$	**0.647**	0.764	0.808	0.781	0.808	110
$t+5$	**0.616**	0.770	0.808	0.791	0.808	94
$t+6$	**0.638**	0.799	0.861	0.816	0.899**	67
$t+7$	**0.599**	0.784	0.847	0.829	0.816	41
$t+8$	**0.957**	1.054	0.963	0.977	1.142	18
	b) SPF Sample					
Horizon	SPF	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|SPF	Size
t	0.906	**0.783*****	0.970	0.809***	0.930	120
$t+1$	0.848	**0.789****	0.991	0.805**	0.948	120
$t+2$	**0.785**	0.785	0.850	0.816	0.776	120
$t+3$	**0.756**	0.781	0.823	0.803	0.782	120
$t+4$	**0.750**	0.780	0.810	0.799	0.768	120
	c) BC Sample					
Horizon	BC	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|BC	Size
t	1.120	**0.790*****	0.971	0.798***	0.927	311
$t+1$	0.960	**0.799****	1.038	0.823**	1.028	311
$t+2$	0.911	**0.817***	0.919	0.850*	0.887	311
$t+3$	0.883	**0.822**	0.880	0.823*	0.872	311
$t+4$	0.857	0.832	0.857	0.828	**0.822**	307
$t+5$	0.839	0.793	0.849	0.827	**0.756**	228
$t+6$	0.852	0.833	0.870	0.853	**0.747**	150
$t+7$	0.831	0.838	0.857	0.861	**0.742**	75

Table 11: Professional Forecasts - Absolute RMSFEs - Pre-2006

	a) GB Sample					
Horizon	GB	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|GB	Size
t	0.163*	**0.152*** **	0.180	0.154***	0.174	110
$t+1$	0.335	**0.321** **	0.406	0.322**	0.392	110
$t+2$	**0.479***	0.508*	0.554	0.523*	0.498	110
$t+3$	**0.591**	0.672	0.722	0.682	0.696	110
$t+4$	**0.695**	0.821	0.869	0.839	0.868	110
$t+5$	**0.741**	0.926	0.972	0.951	0.972	94
$t+6$	**0.806**	1.010	1.088	1.031	1.136**	67
$t+7$	**0.808**	1.058	1.143	1.119	1.101	41
$t+8$	**0.914**	1.007	0.920	0.933	1.091	18
	b) SPF Sample					
Horizon	SPF	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|SPF	Size
t	0.144	**0.125*** **	0.154	0.129***	0.148	120
$t+1$	0.331	**0.308** **	0.386	0.314**	0.370	120
$t+2$	0.492	0.493	0.533	0.512	**0.487**	120
$t+3$	**0.645**	0.666	0.702	0.685	0.667	120
$t+4$	**0.784**	0.816	0.847	0.836	0.803	120
	c) BC Sample					
Horizon	BC	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|BC	Size
t	0.211	**0.149*** **	0.183	0.150***	0.174	311
$t+1$	0.385	**0.321** **	0.417	0.330**	0.413	311
$t+2$	0.552	**0.495***	0.557	0.515*	0.537	311
$t+3$	0.713	**0.664**	0.710	0.665*	0.704	311
$t+4$	0.851	0.827	0.852	0.823	**0.818**	307
$t+5$	0.970	0.917	0.981	0.955	**0.874**	228
$t+6$	1.051	1.027	1.073	1.052	**0.921**	150
$t+7$	1.143	1.153	1.179	1.184	**1.020**	75

Table 12: Absolute RMSFEs for GB Sample - Role of the Information Set

Horizon	GB	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|GB	Size
	a) GB Sample - Same Month Set					
Horizon	GB	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|GB	Size
t	0.165	**0.148*****	0.174	0.152***	0.168	127
$t+1$	0.332	**0.317***	0.391	0.321*	0.379	127
$t+2$	**0.480**	0.497	0.542	0.513	0.491	127
$t+3$	**0.608**	0.668	0.715	0.675	0.694	127
$t+4$	**0.735**	0.838	0.874	0.846	0.872	127
$t+5$	**0.851**	1.008	1.028	1.010	1.015	111
$t+6$	**0.977**	1.164	1.188	1.154	1.211	80
$t+7$	**1.103**	1.309	1.320	1.314	1.266	51
$t+8$	**1.443**	1.578	1.492	1.506	1.453	22
	b) GB Sample - Same Information Set					
Horizon	GB	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|GB	Size
t	0.164	**0.151*****	0.179	0.153***	0.172	114
$t+1$	0.333	**0.318****	0.401	0.319**	0.387	114
$t+2$	**0.478**	0.505	0.549	0.518	0.493	114
$t+3$	**0.596**	0.678	0.723	0.684	0.692	114
$t+4$	**0.712**	0.840	0.879	0.853	0.871	114
$t+5$	**0.810**	0.993	1.023	1.006	1.005	98
$t+6$	**0.957**	1.155	1.201	1.157	1.212**	71
$t+7$	**1.102**	1.333	1.364	1.352	1.273	45
$t+8$	**1.489**	1.631	1.528	1.543	1.499	20

Table 13: Absolute RMSFEs – Model Forecasts with Rounding - 1976:M1-2010:M12

Horizon	FLOW-VAR	VAR	FLOW-UC	FLOW-UC|VAR	AR	GAR	SETAR
t	0.164***	**0.160*****	0.195	0.164***	0.197	0.198	0.200
$t+1$	0.373**	**0.357*****	0.462	0.370***	0.440	0.455	0.438
$t+2$	0.599*	**0.588****	0.648	0.598**	0.692	0.718	0.686
$t+3$	0.836	0.819	0.856	**0.807****	0.936	0.960	0.932
$t+4$	1.061	1.040	1.049	**1.016***	1.155	1.158	1.177
$t+5$	1.269	1.235	1.225	**1.203**	1.343	1.320	1.426
$t+6$	1.431	1.395	1.378	**1.364**	1.492	1.445	1.646
$t+7$	1.574	1.526	1.509	**1.500**	1.614	1.550	1.874
$t+8$	1.689	1.645	1.614	**1.617**	1.703	1.633	2.079

Table 14: Professional Forecasts - RMSFEs - with Rounding

	a) GB Sample					
Horizon	GB	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|GB	Size
t	0.168**	**0.159*** **	0.187	0.161***	0.182	114
$t+1$	0.332	**0.315** **	0.398	0.320**	0.385	114
$t+2$	**0.477**	0.510	0.544	0.518	0.490	114
$t+3$	**0.593**	0.677	0.720	0.682	0.684	114
$t+4$	**0.711**	0.837	0.875	0.851	0.870	114
$t+5$	**0.810**	0.996	1.023	1.007	0.999	98
$t+6$	**0.957**	1.161	1.205	1.161	1.221	71
$t+7$	**1.100**	1.328	1.360	1.351	1.271	45
$t+8$	**1.484**	1.627	1.526	1.528	1.483	20
	b) SPF Sample					
Horizon	SPF	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|SPF	Size
t	0.155	**0.143*** **	0.174	0.153**	0.163	140
$t+1$	0.360	**0.343** **	0.430	0.358*	0.386	140
$t+2$	**0.550** *	0.583	0.620	0.592	0.554	140
$t+3$	**0.749** *	0.813	0.844	0.819	0.762	140
$t+4$	**0.933** *	1.027	1.038	1.028	0.947	140
	c) BC Sample					
Horizon	BC	VAR	FLOW-UC	FLOW-UC|VAR	FLOW-UC|BC	Size
t	0.218*	**0.160*** **	0.196	0.165***	0.183	371
$t+1$	0.410	**0.359** **	0.471*	0.373**	0.439	371
$t+2$	0.607	**0.593**	0.656	0.602	0.612	371
$t+3$	0.810*	0.835	0.869	0.816	**0.808**	371
$t+4$	1.000	1.072	1.071	1.039	**0.978**	367
$t+5$	1.178	1.262	1.256	1.232	**1.111**	273
$t+6$	1.294	1.448	1.404	1.378	**1.211**	180
$t+7$	1.417	1.589	1.561	1.543	**1.318**	90