The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market

Regis Barnichon
CREI, Universitat Pompeu Fabra, and Barcelona GSE

Christopher J. Nekarda
Board of Governors of the Federal Reserve System

First version: January 2012
This version: September 2012

Abstract

This paper presents a forecasting model of unemployment based on labor force flows data that, in real time, dramatically outperforms the Survey of Professional Forecasters, the Federal Reserve Board's Greenbook forecast, and basic time-series models. Our model reduces the root-mean-squared error of the best forecast by about 30 percent in the near term and performs especially well around recessions and turning points. Further, because our model uses information on labor force flows typically ignored by other approaches, a combined forecast including our model and the Greenbook forecast yields improvement of about 35 percent for current-quarter forecasts, and 15 percent for next quarter forecasts, as well as improvements at longer horizons.

The views in this paper are those of the authors and do not necessarily represent the views or policies of the Board of Governors of the Federal Reserve System or its staff. We would like to thank Wouter den Haan, Bart Hobijn, Òscar Jordà, Barbara Rossi, Tara Sinclair, Herman Stekler, and Paolo Surico. Regis Barnichon acknowledges financial support from the Spanish Ministerio de Economía y Competitividad (grant ECO2011-23188), the Generalitat de Catalunya (grant 2009SGR1157), and the Barcelona GSE Research Network.

1 Introduction

Forecasting the unemployment rate is an important and difficult task for policymakers, especially surrounding economic downturns. Despite decades of research on the topic, policymakers often rely on Okun's law—the empirical relationship between output growth and unemployment changes—or simple time-series models to forecast unemployment.

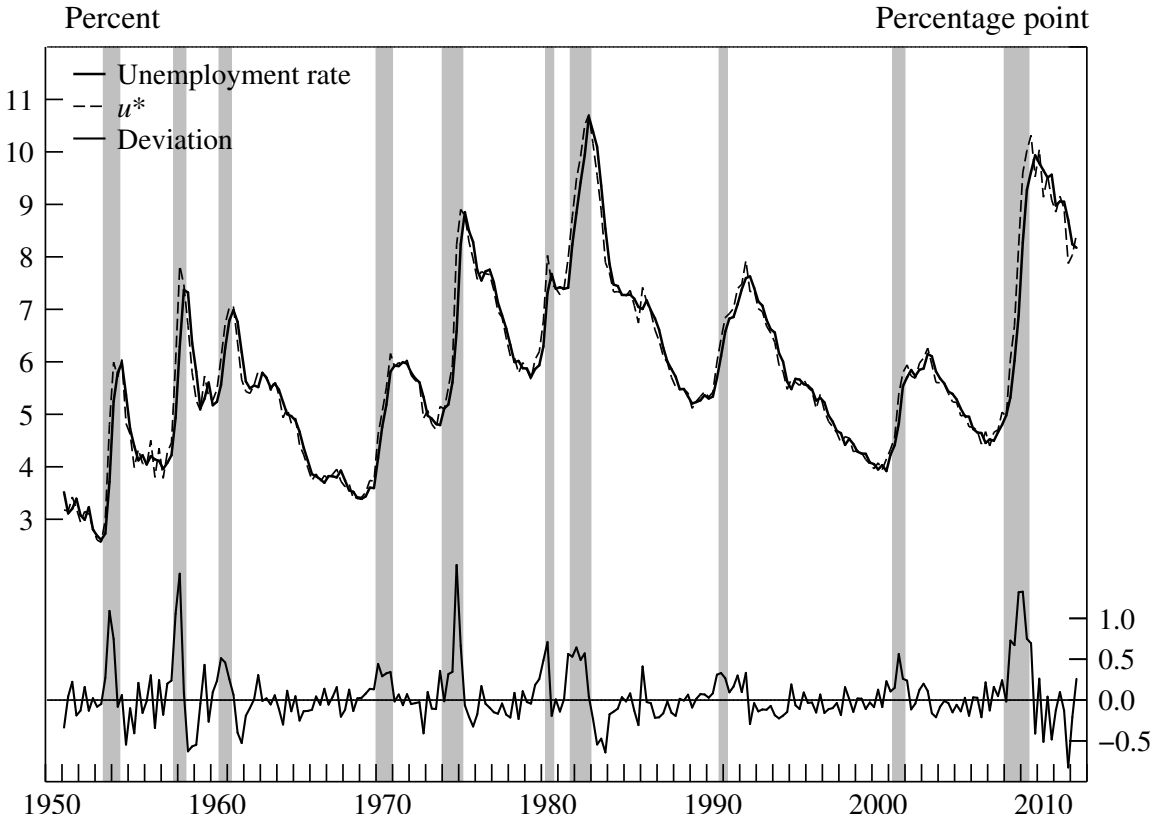
This paper presents a forecasting model of unemployment based on labor force flows. We exploit the tight relationship between the published unemployment rate, u, derived from the stocks of employed and unemployed persons, and the rate of unemployment implied by the underlying labor force flows, the *conditional steady-state unemployment rate*, u^*. Because the unemployment rate converges toward its steady-state rate, the flows provide information about the future unemployment rate that can improve forecasts of the unemployment rate. Figure 1 shows the tight relationship between the steady-state unemployment rate and the published unemployment rate. As shown by the deviation ($u^* - u$) plotted in the lower panel, in periods when u^* is above the actual rate, the unemployment rate tends to rise, and, conversely, when u^* lies below u, the unemployment rate tends to fall. This observation forms the motivation of our approach to forecasting unemployment with labor force flows.

Our model dramatically outperforms the Survey of Professional Forecasters (SPF), the Federal Reserve Board's Greenbook forecast, and standard time-series models for short-term forecasts, reducing the root-mean-squared error (RMSE) of the best model by about 30 percent. The model also does a good job at identifying turning points several quarters ahead of alternative models. Moreover, because our models' forecasts are based on information likely not incorporated by other forecasts, they can be combined with other models to further reduce RMSE.

Our forecasting model is built on two elements: a (nonlinear) law of motion describing how the unemployment rate converges to its conditional steady-state value—the rate of unemployment that would eventually prevail were the flows into and out of unemployment to remain at their current rates—and a forecast of these labor force flows. In turn, the model's improved performance stems from two principles: (1) Unemployment converges relatively quickly to its conditional steady state, and (2) the flows into and out of unemployment have different time-series properties than the stock.

First, our model exploits the relationship between the unemployment rate and its steady-state value implied by the flows. In steady state, the inflows into unemployment and outflows from unemployment are balanced. However, if the inflow rate were to jump, as tends to happen at the onset of a recession, then the conditional steady-state unemployment rate would also jump. With no additional shocks to the flows, the unemployment rate would progressively rise toward this new steady state. Since this convergence process occurs relatively quickly (within about three to five months), the conditional steady state provides information about the unemployment rate in the near future. Thus, by incorporating information from labor force flows, we can exploit this convergence

Figure 1. Unemployment Rate and Steady-State Unemployment Rate



Source: Bureau of Labor Statistics data and authors' calculations.

Notes: $u^* = s/(s + f)$; see equation 3. Quarterly average of monthly data. Shaded areas represent periods of business recession as determined by the National Bureau of Economic Research.

and improve near-term forecasts of the unemployment rate.

However, relying solely on current labor force flows constrains our approach to very near term forecasts, because the steady state to which the actual unemployment rate converges also changes over time as the underlying flows evolve. Thus, we forecast the underlying labor force flows using a time-series model and feed those forecasts into our law of motion to generate unemployment forecasts at longer horizons. Directly forecasting the flows into and out of unemployment rather than the unemployment stock itself, as is customary, is the second reason that our models outperform other approaches. Directly forecasting the flows allows our models to better capture the dynamics of unemployment—because the unemployment stock is driven by flows with different time-series properties, and because the contribution of the different flows changes throughout the cycle (Barnichon, 2012).

An additional advantage of focusing on labor force flows is that it allows us to capture the

asymmetric nature of unemployment movements—in particular, that increases are steeper than decreases.[1] Although our model is not explicitly asymmetric, it relies, in part, on the unemployment inflows, which are responsible for the asymmetry of unemployment (Barnichon, 2012). By using such information as inputs in the forecasts, our model can incorporate the impulses responsible for the asymmetry of unemployment.[2] Thanks to this property, we find that our model outperforms a baseline time-series model around turning points and large recessions. This property is particularly useful given that these are precisely the times when accurate unemployment forecasts are the most valuable.

One final benefit on focusing on the flows is that, as shown by Fujita and Ramey (2009), unemployment inflows lead outflows by about a quarter. As a result, our model does a good job at identifying turning points several quarters ahead of other forecasters and models. Indeed, because a turning point in the inflow rate typically signals a turning point in unemployment several quarters in advance, our model can better predict turning points. This is a significant improvement compared to the consensus forecast, the Greenbook, or time-series models—all of which typically miss turning points during contractionary periods.[3]

Our models are a useful addition to the set of forecasting models because our approach uses information on labor force flows typically ignored by standard approaches. A new forecast combining our models' forecasts and the Greenbook forecasts yields a reduction in RMSE of about 35 percent for current-quarter forecasts, 15 percent for next quarter forecasts, almost 10 percent for two-quarter-ahead forecasts, as well as slight improvements at longer horizons.

Finally, when incorporating flows into and out of the labor force, our model can be used to forecast the labor force participation rate. The forecast performance for labor force participation is nonetheless more modest than for the unemployment rate. Our model improves on the Greenbook for the current-quarter forecast, but performs worse thereafter. Surprisingly, however, a combined forecast of our model and the Greenbook yields dramatic reductions in RMSE.

This paper builds on the influential work of Montgomery et al. (1998) and extend the growing literature aimed at improving the performances of unemployment forecasting models.[4] We particularly draw on the recent literature on labor force flows, which has typically been overlooked by the forecasting literature, but has been the subject of numerous studies aimed at understanding the determinants of labor market fluctuations.[5]

1. For evidence on the asymmetry of unemployment, see Mitchell (1927), Neftçi (1984), DeLong and Summers (1986), and Sichel (2007).

2. And unlike standard time-series models used to capture asymmetries (such as threshold autoregressive models), our model does not rely on an arbitrary threshold to introduce asymmetry.

3. Montgomery et al. (1998).

4. See, for example, Rothman (1998), Golan and Perloff (2004), Brown and Moshiri (2004), and Milas and Rothman (2008).

5. See Shimer (2012); Petrongolo and Pissarides (2008); Elsby, Michaels and Solon (2009); Nekarda

2 The Steady-State Unemployment Rate

Our forecast is built on two elements: (1) A law of motion describing how the unemployment rate converges to its steady-state value, and (2) a forecast of the labor force flows determining steady-state unemployment and the speed at which actual unemployment converges to steady state. We first present a model with only two labor force states and then expand it to the more general case with three labor force states.

2.1 The Labor Market with Two States

We first develop a model with only two labor force states: employed and unemployed. That is, we explicitly assume that there are no movements into and out of the labor force. This approach is consistent with recent literature (for example, Shimer, 2012) showing that a two-state model does a good job of capturing unemployment fluctuations. In addition, it provides a simpler framework for understanding the basic flow-based accounting of the conditional steady-state unemployment rate, and it can be estimated over a long period using duration data. However, in section 2.2, we generalize our approach to three states and allow for movements into and out of the labor force.

2.1.1 The Law of Motion for Unemployment

Denote $u_{t+\tau}$ the unemployment rate at instant $t+\tau$ with t indexing months and $\tau \in [0,1]$ a continuous measure of time within a month. Assume that between month t and month $t+1$ all unemployed persons find a job according to a Poisson process with constant arrival rate f_{t+1}, and all employed workers lose their job according to a Poisson process with constant arrival rate s_{t+1}.[6] The unemployment rate then evolves according to

$$\frac{\mathrm{d}u_{t+\tau}}{\mathrm{d}\tau} = s_{t+1}\left(1-u_{t+\tau}\right) - f_{t+1}u_{t+\tau}, \tag{1}$$

as changes in unemployment are given by the difference between the inflows and the outflows. Solving equation 1 yields

$$u_{t+\tau} = \beta_{t+1}(\tau)u^*_{t+1} + \left[1-\beta_{t+1}(\tau)\right]u_t, \tag{2}$$

(2009); Barnichon (2012); and Elsby, Hobijn and Şahin (2011), among others.

6. We adopt this timing convention to reflect data availability, as the hazard rate is only observed in month $t+1$.

where

$$u^*_{t+1} \equiv \frac{s_{t+1}}{s_{t+1} + f_{t+1}} \tag{3}$$

denotes the conditional steady-state unemployment rate, and $\beta_{t+1}(\tau) \equiv 1 - \mathrm{e}^{-\tau(s_{t+1}+f_{t+1})}$ is the rate of convergence to that steady state.

Equation 2 relates variation in the unemployment stock $u_{t+\tau}$ over the course of a month to variation in the underlying flow hazards, f_{t+1} and s_{t+1}. A one-month-ahead forecast for the unemployment rate, $\hat{u}_{t+1|t}$, can thus be obtained from

$$\hat{u}_{t+1|t} = \hat{\beta}_{t+1} u^*_{t+1} + \left(1 - \hat{\beta}_{t+1}\right) u_t, \tag{4}$$

where $\hat{\beta}_{t+1}$ is the month t forecast of β_{t+1}, the convergence speed between t and $t + 1$.

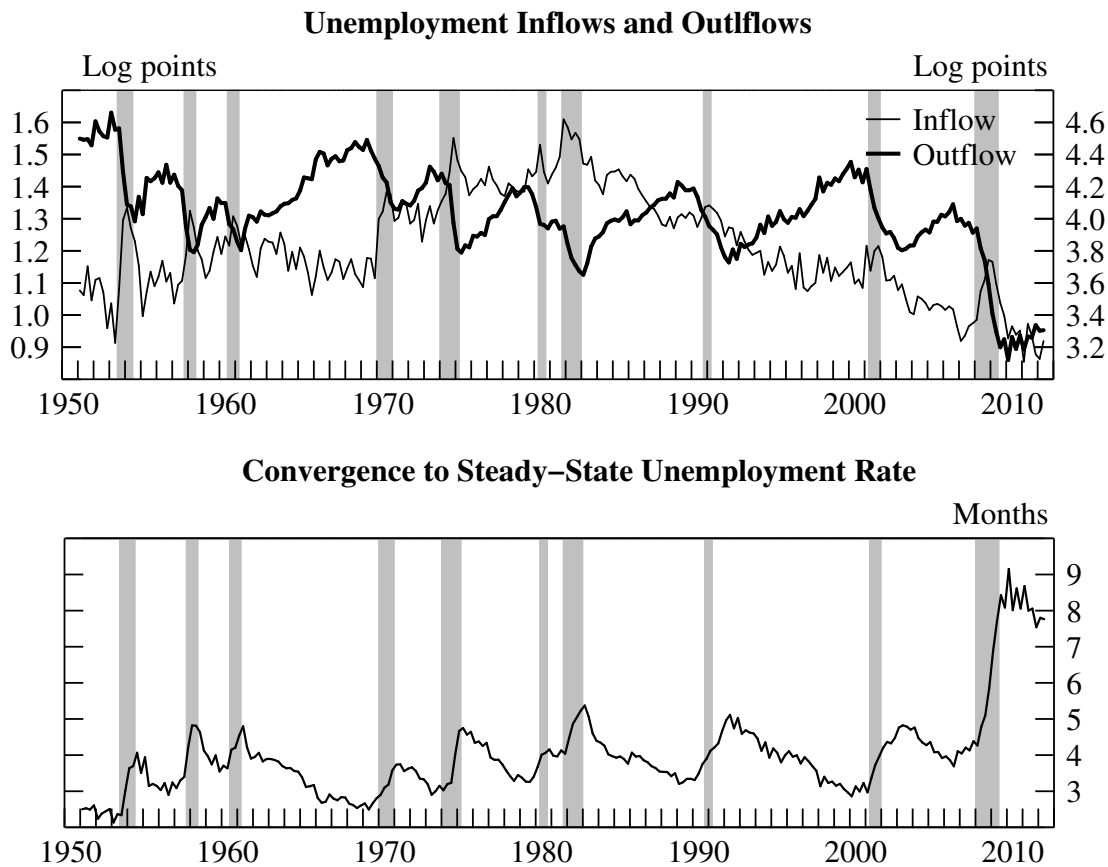
Over 1951–2011, the sum of monthly unemployment inflow and outflow rates averaged 0.62, implying that the half-life deviation of unemployment from its steady-state is slightly more than a month. As a result, unemployment gets 90 percent of the way to its conditional steady-state value in about four months, on average. However, as the lower panel of figure 2 shows, the convergence speed varies considerably over the business cycle, as inflow and outflow rates evolve. As a result, the time needed to close 90 percent of the gap with steady state unemployment ranges from about three months in tight labor markets to about five months in slack markets. In the 2007–08 recession, the drop in the unemployment exit rate was so dramatic, that the figure increased to an unprecedented nine months. It has since edged lower to just under eight months in the second quarter 2012.

Part of the exceptional increase owes to a dramatic decline in the job finding rate, specifically, exceptionally low job creation and low matching efficiency (Barnichon and Figura, 2010). Moreover, Elsby et al. (2011) showed that the exceptional decline was also an artifact of measurement error, because not all persons flowing into unemployment had a duration of fewer than five weeks. This phenomenon became much more prevalent in the recent recession, where a larger fraction of unemployment inflows reported a duration of more than five weeks, leading the duration-based measure of the unemployment exit rate to suffer from a larger downward bias. As we discuss in section 7, to the extent that the bias is stronger than in previous recessions, forecasting performances could deteriorate.

2.1.2 Forecasting Labor Force Flows

Because equation 4 only forecasts the unemployment rate one month ahead given current values of the hazard rates, forecasting the unemployment rate at longer horizons requires making forecasts of the hazard rates.

Figure 2. Unemployment Inflow and Outflow Hazard Rates and Convergence to Steady-State



Source: Authors' calculations based on Bureau of Labor Statistics data.

Notes: Time in months needed to close 90 percent of the gap with steady-state unemployment rate $u^* = s/(s + f)$. Quarterly average of monthly data. Shaded areas represent periods of business recession as determined by the National Bureau of Economic Research.

A simple approach is to assume that the hazard rates remain constant at their last observed value over the forecast horizon. However, in real time a forecaster does not observe s_{t+1} and f_{t+1}, but only s_t and f_t. This is because at month t one can only observe labor force flows from $t-1$ to t.[7] Thus, the j-period-ahead forecast of the unemployment rate can be formed from the month t values of s

7. A concrete example helps clarify this point. The August employment report (published on September 7, 2012) provided information on the stock of unemployment in August and the average unemployment inflow and outflow rates between July and August (s_t and f_t). Looking back at equation 4, this allows us to measure β_t, u_t^* and u_t. Thus, to forecast $\hat{u}_{t+1|t}$ (the unemployment rate in September), we need forecasts of $\hat{f}_{t+1|t}$ and $\hat{s}_{t+1|t}$ (that is, the flows from August to September).

and f by[8]

$$\hat{u}_{t+j|t} = \left[1 - \mathrm{e}^{-j(f_t+s_t)}\right] u_t^* + \mathrm{e}^{-j(f_t+s_t)} u_t. \tag{5}$$

If the hazard rates are persistent enough, equation 5 will provide reasonable forecasts.[9] However, as figure 2 shows, the hazard rates do evolve, and with them the conditional steady-state unemployment rate and the speed of convergence.

Because the hazard rates are not sufficiently persistent, we use a vector autoregression (VAR) to forecast the inflow and outflow rates. We also include two leading indicators of labor force flows: vacancy posting and initial claims for unemployment insurance. Specifically, let

$$\mathbf{y}_t = (\ln s_t, \ln f_t, \Delta \ln u_t, \ln uic_t, \Delta \ln hwi_t)' ,$$

where *uic* is the monthly average of weekly initial claims for unemployment insurance and Δhwi is the change in Barnichon's (2010) composite help-wanted index. Note that, given our timing convention for the flows, the hazard rates effectively enter the VAR lagged by one month. We estimate the VAR

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \varepsilon_t \tag{6}$$

over a fifteen-year rolling window.[10]

After generating forecasts of the hazard rates, we obtain j-period-ahead forecasts of unemployment by iterating on[11]

$$\hat{u}_{t+j|t} = \hat{\beta}_{t+j} \hat{u}^*_{t+j|t} + \left(1 - \hat{\beta}_{t+j}\right) \hat{u}_{t+j-1|t}, \tag{7}$$

8. This law of motion forms the basis of Elsby, Hobijn and Şahin's (2011) strategy to generalize Shimer's (2012) unemployment decomposition to incorporate out-of-steady-state dynamics.

9. In section 4, we consider a forecast based only on the convergence to the steady-state unemployment rate.

10. We found that, in real time, a rolling window (in which the model is estimated over the previous K months) yielded more accurate forecasts than a recursive window (in which the model is estimated over the entire observed history), likely because of the low-frequency patterns; fifteen years was superior to ten- and twenty-year windows. We also considered lag lengths between 1 and 12.

11. Because $\hat{u}_{t+j|t}$ is a nonlinear function of $\hat{f}_{t+j|t}$ and $\hat{s}_{t+j|t}$, Jensen's inequality, in theory, prohibits us from directly forecasting the unemployment rate from equation 7 and forecasts of $f_{t+j|t}$ and $s_{t+j|t}$. To avoid this problem, we use Monte Carlo simulation and sample with replacement from the VAR residuals $\{\epsilon^f, \epsilon^s\}$ and form forecasts (equations 8 and 9) using the sampling distribution of ϵ^f and ϵ^s. In practice, given the magnitude of the inflows and outflows, the unemployment rate forecasts obtained by Monte Carlo simulation are not different from those formed from equation 7. For simplicity, we use that approximation.

with

$$\hat{u}^*_{t+j} = \frac{\hat{s}_{t+j|t}}{\hat{s}_{t+j|t} + \hat{f}_{t+j|t}} \tag{8}$$

and

$$\hat{\beta}_{t+j} = 1 - \mathrm{e}^{-\left(\hat{s}_{t+j|t} + \hat{f}_{t+j|t}\right)}. \tag{9}$$

With month $t + j$ forecasts of the flow rates in hand, we can forecast the month $t + j$ values of u^* and β. The month $t + j$ unemployment forecast is then obtained by taking a weighted average of the previous-period (month $t + j - 1$) unemployment rate and the current-period (month $t + j$) conditional steady-state unemployment rate, with the weights determined by the speed of convergence to steady state.

2.2 The Labor Market with Three States

So far, we have only considered a labor market with two states: employed or unemployed. However, not all those without jobs are unemployed. Indeed, flows into and out of the labor force dwarf those into and out of unemployment.[12] This section considers a model that incorporates flows among all three labor force states.

An important advantage of the three-state model is the ability to capture more accurately the flows taking place in the labor market. For instance, the unemployment inflow rate comprises both those losing or leaving jobs as well as entrants to the labor force. Since these two flows (and in fact all six flows) display different time-series properties, a three-state model may produce better forecasts than a two-state model.[13] In addition, the three-state model can be used to forecast the labor force participation rate.

To generalize our two-state framework to three states, we need to specify and solve the system of differential equations governing the number of people in unemployment, U; in employment, E; or out of the labor force, N.

Between month t and month $t + 1$, individuals can transit from state $a \in \{E, U, N\}$ to state $b \in \{E, U, N\}$ according to a Poisson process with constant arrival rate λ^{ab}_{t+1}. The stocks of unemployed,

12. See Blanchard and Diamond (1990) for the seminal study of gross flows.
13. See Barnichon and Figura (2010) for more on the properties of the different flows.

employed, and persons not in the labor force satisfy the system[14]

$$\begin{cases} \dot{U}_{t+\tau} = \lambda_{t+1}^{EU} E_{t+\tau} + \lambda_{t+1}^{NU} N_{t+\tau} - (\lambda_{t+1}^{UE} + \lambda_{t+1}^{UN}) U_{t+\tau} \\ \dot{E}_{t+\tau} = \lambda_{t+1}^{UE} U_{t+\tau} + \lambda_{t+1}^{NE} N_{t+\tau} - (\lambda_{t+1}^{EU} + \lambda_{t+1}^{EN}) E_{t+\tau} \\ \dot{N}_{t+\tau} = \lambda_{t+1}^{EN} E_{t+\tau} + \lambda_{t+1}^{UN} U_{t+\tau} - (\lambda_{t+1}^{NE} + \lambda_{t+1}^{NU}) N_{t+\tau}. \end{cases} \tag{10}$$

For instance, as shown in the first line, changes in unemployment are given by the difference between the inflows to unemployment (workers losing or quitting their jobs and persons joining the labor force) and the outflows from unemployment (unemployed persons finding a job or dropping out of the labor force).

Then, using the initial and terminal conditions, the one-step ahead forecasts of the three stocks can be solved as functions of the transition probabilities (λ^{ab}s). The details of the solution are shown in the appendix. We then use the solution to these equations to generate one-period-ahead forecasts of the unemployment rate and labor force participation rate from

$$\hat{u}_{t+1|t} = \frac{\hat{U}_{t+1|t}}{\hat{U}_{t+1|t} + \hat{E}_{t+1|t}} \tag{11}$$

and

$$l\hat{f}pr_{t+1|t} = \frac{\hat{U}_{t+1|t} + \hat{E}_{t+1|t}}{\hat{E}_{t+1|t} + \hat{U}_{t+1|t} + \hat{N}_{t+1|t}}. \tag{12}$$

Note that, in effect, what we assume for population growth does not affect our forecasts, because we forecast population shares.

As with the two-state model, to construct forecasts beyond one period ahead, we used a VAR to forecast the six transition probabilities. Specifically, we estimate

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \mathbf{\Phi}_3 \mathbf{y}_{t-3} + \varepsilon_t, \tag{13}$$

over a ten-year rolling window, where

$$\mathbf{y}_t = \left(\ln \lambda_t^{EU}, \ln \lambda_t^{UE}, \ln \lambda_t^{EN}, \ln \lambda_t^{NE}, \ln \lambda_t^{NU}, \ln \lambda_t^{UN}, \ln u_t, \ln uic_t, \ln hwi_t \right)'.$$

Note that in this specification, the unemployment rate and help-wanted index enter in levels, rather than in changes because it yielded marginally better forecasts.

14. Equation 10 assumes that P_t is constant within a month and that inflows and outflows to the civilian noninstitutional population aged sixteen and older are negligible. This assumption is reasonable given that the working-age population increases by about 150,000 per month, an order of magnitude or two higher than the flows into and out of E, U or N.

3 Data

We constructed measures of the transition rates in the two-state and three-state models using different approaches: an indirect one (using information on the stocks of unemployment and short-term unemployment to infer the transition rates) for the two-state model and a direct one (using measures of labor force flows) for the three-state model.

For the two-state model, we follow Shimer (2012) and use information on the number of persons unemployed, U_t, and those unemployed for fewer than five weeks, U_t^s, to infer job finding and job separation hazard rates. Specifically, the unemployment outflow probability, F, was calculated from

$$F_{t+1} = 1 - \frac{U_{t+1} - U_{t+1}^s}{U_t},$$

with $f_{t+1} = -\ln(1 - F_{t+1})$ the hazard rate. The unemployment inflow rate, s, was then obtained by solving equation 1 forward over $[t, t+1]$ and finding the value of s_{t+1} that solves

$$U_{t+1} = \frac{\left[1 - \mathrm{e}^{-(f_{t+1}+s_{t+1})}\right] s_{t+1}}{f_{t+1} + s_{t+1}} (U_t + E_t) + \mathrm{e}^{-(f_{t+1}+s_{t+1})} U_t.$$

Note that in this accounting, given a value for the unemployment outflow rate (which also captures movements out of the labor force) and the stock of unemployed persons, the inflow rate is the rate that explains the observed stock of unemployed persons in the next month. As a result, the inflow rate incorporates all movements in unemployment not accounted for by the unemployment outflow rate.

For the three-state model, we used aggregate labor market transition probabilities among employment, unemployment, and nonparticipation calculated from longitudinally-matched Current Population Survey (CPS) microdata. We constructed the transition rates from labor market flows as $\lambda_t^{ab} \equiv ab_t / a_{t-1}$, where ab_t is the number of persons who were observed having state a in month $t-1$ and subsequently having state b in month t. (The time series of λ^{ab}s are collectively referred to as "gross flows.") The Bureau of Labor Statistics publishes a research series of gross flows that begins in February 1990. For contemporary forecasts, the published data have a sufficiently long history to estimate the model. However, to evaluate historical model forecasts prior to February 2000, we needed data with a longer history. Thus, we constructed measures of gross flows that cover June 1967 to January 1990, allowing us to begin our historical forecasts in 1976. From 1976 to 1990, we constructed gross flows from Nekarda's (2012) Longitudinal Population Database. Before 1976, we used gross flows tabulated by Joe Ritter.[15]

Finally, weekly initial claims for unemployment insurance are published by the Department of Labor, Employment and Training Administration. Our measure of vacancy posting is the composite

15. See Shimer (2012).

help-wanted index presented in Barnichon (2010).

4 Forecasting Performance

We evaluated the performance of our flows models by comparing their unemployment rate forecasts with alternative forecasts along several dimensions. First, we assessed the RMSE of out-of-sample forecasts. Next, because it is harder to forecast the unemployment rate around recessions, we assessed our model's performance relative to a baseline time-series model over the business cycle. Finally, we examined the conditions under which the model can forecast business cycle turning points. In what follows, we refer to the forecasts from the two-state model as "SSUR-2" and the forecasts from the three-state model as "SSUR-3."

4.1 Real-Time Forecasts

Our objective in this section is to evaluate our models' forecasts against the best alternative forecasts, by both professional forecasters and other time-series models. We consider five alternative forecasts of the unemployment rate. The first two alternatives are professional forecasters. We consider the Greenbook forecast, which the literature has generally shown to be the benchmark forecast, and the median forecast from the SPF.[16]

The other three alternative forecasts are from time-series models, and are intended to disentangle the mechanisms behind the performance of our model. We consider a basic univariate time-series model, intended as a "naive" forecast that takes into account no other information but the unemployment rate. Following Montgomery et al., we use an ARIMA(2,0,1) model for the unemployment rate. We also consider the unemployment rate forecast derived from the law of motion for unemployment rate (equation 5) holding the inflow and outflow rates constant at their last known value. We refer to this model as the u^* model. Shutting down the evolution of the hazard rates isolates the contribution of the current conditional steady-state unemployment rate. Our last alternative is the unemployment rate forecast from a VAR that includes the labor force flows and the two leading indicators. By comparing our SSUR models against the VAR, we can directly evaluate the nonlinear relationship implied by the theory compared to an atheoretical linear time-series model using the same information set. The three alternative time-series models are estimated over a fifteen-year rolling window.

Historical forecasts were necessarily made with the data in hand at the time the projection was made. Some economic data, such as real GDP and payroll employment, are subject to substantial revision over time. For these variables, the current-vintage data may differ substantially from

16. See, for example, Romer and Romer (2000); Sims (2002); Faust and Wright (2007); and Tulip (2009).

the historical data used to make a historical forecast. In the case of the unemployment rate, however, revisions are relatively minor. The labor force data obtained from the CPS are revised only to reflect updated estimates of seasonal fluctuations. Indeed, the not-seasonally adjusted stocks of employment and unemployment are never revised, reflecting their origin from a point-in-time survey of households. Nevertheless, even the small revisions to seasonal factors may have important consequences for our models' performance.

To construct real-time estimates of the hazard rates, we begin with monthly vintages of the published seasonally-adjusted stock of employed, unemployed, and short-term unemployed.[17] For each month, we estimate the time series of the inflow and outflow hazard rates from the real-time stocks as described in section 3. These series are then used in the VAR to forecast the evolution of the hazard rates. Real-time data for initial claims and the help-wanted index are not available, so current-vintage data are used in the VAR. As with the unemployment rate, revisions to these series are relatively small.[18] Nonetheless, in our evaluation section, we assess the implications of this limitation.

To reflect the environment within which forecasters must operate, we estimated our models and generated forecasts using real-time data, except for initial claims and the help wanted index. For each alternative forecast, we estimated all models taking as inputs only the data that were available at the time of the forecast.

The SPF sends out the survey questionnaire sometime in the first month of a quarter, and the survey participants are asked to mail back the completed questionnaire by the middle of the second month of the quarter. Thus, the forecasts included in the SPF incorporate labor market data from the first month of each quarter. To make forecasts comparable, the model forecasts are made treating the first month's unemployment rate as data.

The information set for the Greenbook forecast is more irregular, and we are careful to mimic the information set for each Greenbook date. Because the timing of the forecast is dictated by the date of the FOMC meeting, Greenbook forecasts are made at different points in a quarter—some forecasts have no monthly data for the current quarter, while others have two months of labor market data. For example, at the time the March 2004 Greenbook was published, the unemployment rate was known through February 2004, and thus the $t + 0$ forecast was made with data for the first two months of the quarter. However, when the April 2004 Greenbook was published, the unemployment rate was known only through March 2003, and thus the $t+0$ forecast (for the second quarter of 2004)

17. An alternative approach that sidesteps the issue of seasonal revisions altogether is to forecast the not-seasonally adjusted unemployment rate from not-seasonally adjusted CPS data. That model performed similarly to the two-state model we presented here.

18. There are no revisions to the print help-wanted index. Real-time data for initial claims are available beginning in June 2009. Over the 39 months for which real-time data are available, the maximum absolute variation in the monthly average level of weekly initial claims over this period was tiny at about 3 percent.

was made without any published data for the quarter. Finally, because the Greenbook forecasts are made public with a five-year lag, our comparison using the Greenbook forecast ends in 2006.

4.2 Forecast Errors

Table 1 reports the RMSE of real-time forecasts for quarterly unemployment rates over a one-year horizon (including a forecast of the current quarter, $t + 0$). To evaluate the statistical significance of our results, we report the p values of the unconditional Giacomini and White (2006) predictive ability test statistic of equal predictive ability between our SSUR-2 forecast and the comparison forecast.[19]

The SSUR-2 model outperforms the Greenbook forecast and the SPF dramatically for short-term forecasts. As shown in the first two rows of table 1, the SSUR-2 model has a RMSE that is more than 30 percent lower than the other two forecasts in the current quarter and 10 percent lower for a one-quarter-ahead forecast. This corresponds to a reduction in RMSE for current-quarter forecasts of roughly 0.05 percentage point. Moreover, the improvement in the current-quarter forecast is statistically significant at the 1-percent level against both forecasts. Although the improvement in next-quarter forecast is of similar magnitude, it is not statistically significant (or, in the case of the Greenbook forecast, only marginally significant). At longer horizons, the improvement over the SPF and Greenbook diminishes. This is not surprising given that the Greenbook and SPF forecasts are based on an array of economic data and models of the broader economy, while SSUR-2 is a statistical model that incorporates only near-term information about the labor market.

The lower panel of table 1 reports the performances of SSUR-2 against the time-series models. The univariate ARIMA model performs worse than SSUR-2 at all horizons. The unemployment rate forecast from the VAR performs worse than SSUR-2 at all horizons, showing that the nonlinearity is an important feature of our model. Finally, the contribution of forecasting the flows is evident from the last row, which reports the performance of a forecast based only on convergence to the conditional steady-state unemployment rate. This model performs worse than SSUR-2 at all horizons, indicating that time variation in the flow rates is, indeed, an important element of our model. It is remarkable that a forecast from the theoretical law of motion (5) that relies on only the last known value of u^* performs as well or better than both estimated time-series models. Foreshadowing section 5 on forecast combination, this result suggests that combining a model based on the steady-state unemployment rate with an estimated time-series model may yield further improvements.

19. We use the Giacomini and White (2006) predictive ability test, because it is robust to both non-nested and nested models (as are the VAR, U^* and SSUR-2 models), unlike the Diebold and Mariano (1995) test.

Table 1. Unemployment Rate Forecasts

Root-mean-squared error (percentage point)

	Forecast horizon (quarters)				
Model	$t+0$	$t+1$	$t+2$	$t+3$	$t+4$
Professional forecasters, common sample, 1976–2006					
SSUR-2	0.12	0.35	0.54	0.70	0.86
Greenbook	0.17***	0.39*	0.54	0.65	0.78
	(0.00)	(0.10)	(0.83)	(0.41)	(0.26)
SPF	0.18***	0.38	0.53	0.66	0.82
	(0.00)	(0.51)	(0.86)	(0.49)	(0.46)
No. obs.	89	89	89	89	89
Models, monthly, 1976–2011					
SSUR-2	0.15	0.38	0.60	0.83	1.06
ARIMA	0.22***	0.52**	0.84*	1.14	1.41*
	(0.00)	(0.05)	(0.08)	(0.11)	(0.08)
VAR	0.19***	0.47***	0.73*	1.03*	1.30
	(0.00)	(0.00)	(0.09)	(0.07)	(0.14)
u^*	0.20***	0.48***	0.70**	0.92	1.11
	(0.00)	(0.00)	(0.01)	(0.16)	(0.41)
No. obs.	432	432	432	429	426

Source: Authors' calculations using data from Bureau of Labor Statistics, Department of Labor, Federal Reserve Bank of Philadelphia, Board of Governors of the Federal Reserve System, and Barnichon (2010).

Notes: The evaluation of professional forecasts is calculated from 89 forecasts over 1976–2006, and the evaluation of the models' forecasts is calculated from 493 forecasts over 1976–2011. $t+0$ denotes current-quarter forecast. p values of Giacomini–White test statistic are reported in parentheses. ***/**/* indicates statistically different from SSUR-2 at 1/5/10 percent.

4.3 Quasi-Real-Time Forecasts and SSUR-3

As we noted earlier, our preferred VAR specifications—that is, including initial claims and the help-wanted index—cannot be estimated in true real time because vintages of the two leading indicators are not available. We assessed whether this gave our model an unfair advantage over the historical professional forecasters by estimating the VAR for the SSUR-2 model without *uic* and Δhwi—a true real-time exercise. Over the sample used in the upper panel of table 1, this model still had an RMSE almost 20 percent lower than the Greenbook at $t+0$ and essentially the same at $t+1$.

We also compared the performance of the true real-time forecasts of SSUR-2 with the model's "quasi-real-time" forecasts, where we used the same rolling estimation and forecasting procedure as in the real-time exercise, but used the current-vintage data at all points; that is, we omitted all variation associated with revisions to the seasonal factors. Over the same sample, the truly real-time forecasts were actually slightly *better* than the quasi-real-time forecasts at all but the current-quarter horizon (where they were equal). This suggests that evaluating the models in quasi-real time likely does not alter significantly the spirit of the real-time exercise.

With this in mind, we assessed the quasi-real-time forecasts of the SSUR-3 model. (Because historical records of seasonal revisions to gross flows are not available, we could not evaluate the performance of the SSUR-3 model in true real time.) Table 2 evaluates the performances of SSUR-3 against SSUR-2 in quasi-real time. Over the period from 1976 to 2006, the three-state model performs a bit worse than SSUR-2 in the current-, next-, and two-quarter-ahead forecasts, while at longer horizons it performs appreciably worse. However, the gross flows we calculate from the CPS microdata (prior to 1990) are noisier than duration-based hazard rates, in part owing to spurious transitions between unemployment and nonparticipation. If, however, we restrict the time period to use only forecasts that were estimated using the published gross flows data, the differences are small at horizons of up to two-quarters-ahead. As with the longer sample, SSUR-3 performs appreciably worse than SSUR-2 at forecast horizons of $t + 3$ and beyond.

4.4 Forecasting Performance over the Business Cycle

The unemployment stock is driven by flows with different time-series properties, and the contribution of the different flows changes throughout the cycle.[20] For instance, inflows are responsible for the sharp increase in unemployment at the onset of recessions, but outflows are the main driving force of unemployment in normal times.

This property suggests that the performance of our flows models relative to other models may vary over the business cycle. For instance, because the SSUR-2 model incorporates the unemployment inflow rate, which is responsible for the asymmetry of unemployment, it may better capture the asymmetric nature of unemployment than standard models. Thus, it may perform better during recessions, especially compared to standard models, which do not include labor force flows.

To test this idea and evaluate whether SSUR-2 performs differently over the course of the business cycle, we use the Giacomini and Rossi (2010) predictive ability test in unstable environments. The test develops a measure of the relative local forecasting performance of two models and is ideal

20. At a quarterly frequency, the autocorrelation of the outflow rate is 0.91, but the inflow rate is only 0.73 (Shimer, 2012). Further, while the distribution of the (detrended) inflow rate is positively skewed and highly kurtotic, the distribution of the (detrended) outflow rate exhibits no skewness and low kurtosis (Barnichon, 2012).

Table 2. Quasi-Real-Time Unemployment Rate Forecasts, 1976–2006

Root Mean Squared Forecast Error (percentage point)

	Forecast horizon (quarters)				
Model	$t+0$	$t+1$	$t+2$	$t+3$	$t+4$
GB/SPF sample, 1976–2006					
SSUR-2	0.13	0.32	0.50	0.67	0.84
SSUR-3	0.15***	0.37*	0.61**	0.85**	1.09**
	(0.01)	(0.08)	(0.02)	(0.01)	(0.02)
GB/SPF sample, 2000–2006					
SSUR-2	0.10	0.23	0.30	0.44	0.54
SSUR-3	0.11*	0.23	0.37**	0.57**	0.77**
	(0.07)	(0.39)	(0.04)	(0.06)	(0.04)

Source: Authors' calculations using data from Bureau of Labor Statistics, Department of Labor, and Barnichon (2010).

Notes: Upper panel is calculated from 89 forecasts over 1976–2006 that share a common information set; lower panel uses 83 monthly forecasts over February 2000 to December 2006. $t+0$ denotes current-quarter forecast. p values of Giacomini–White test statistic are reported in parentheses. ***/**/* indicates statistically different from SSUR-2 at 1/5/10 percent.

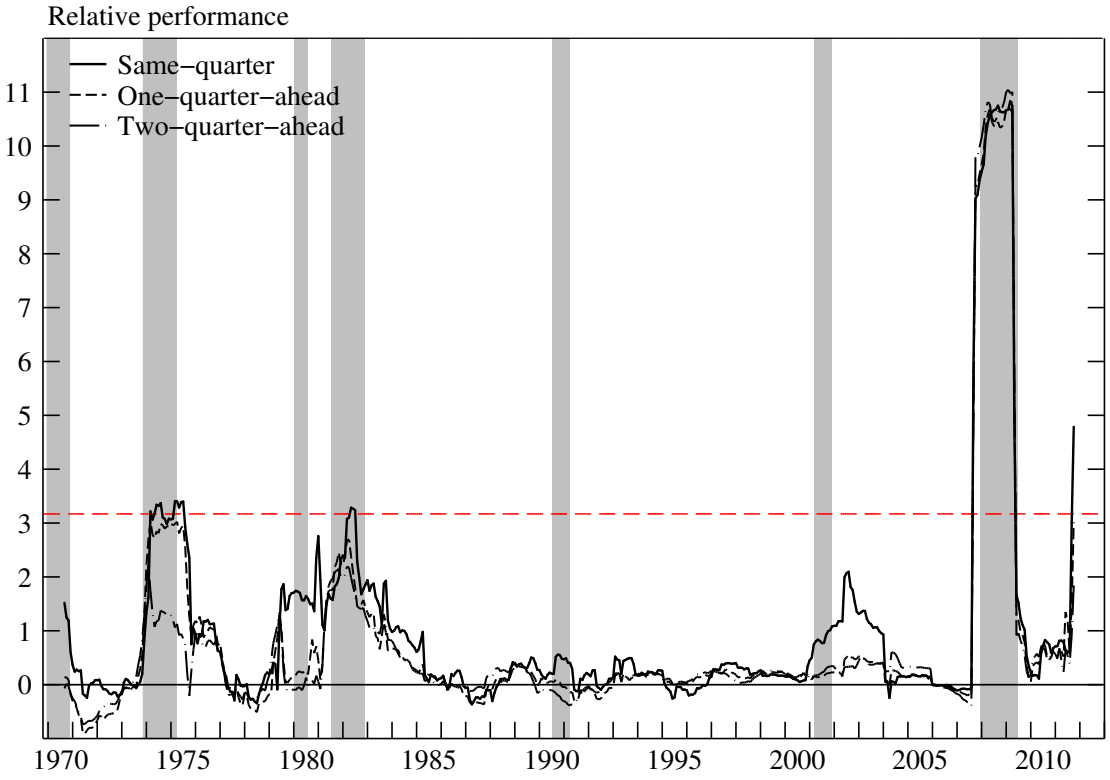
for testing whether the performance of our model varies over the cycle (compared to a benchmark model). We use as a benchmark the ARIMA model presented in table 1. We evaluate the local forecasting performance over a five-year window from monthly forecasts spanning November 1968 to February 2012.[21]

Figure 3 plots the Giacomini and Rossi fluctuation test for current-quarter, one-quarter-ahead, and two-quarter-ahead forecasts, along with the corresponding 5 percent critical value. The unit of the y-axis is the (standardized) rolling difference in mean-squared-error between the two models. This is measure of the relative performance; a positive value indicates a superior performance of SSUR-2.

While SSUR-2 forecasts are almost always more accurate than those of the benchmark model, SSUR-2 performs especially well around recessions—and particularly during the deep recessions in 1973, the early 1980s, and 2007–08—and during times of large and swift movements in the inflow rate. In other words, SSUR-2 yields the greatest improvement over a naive baseline around turning

21. Although the Greenbook forecast would be an interesting benchmark to compare against, we use the ARIMA model as our benchmark comparison because the Giacomini and Rossi (2010) test is only valid for models estimated over rolling-windows. (Both models are estimated over a fifteen-year rolling window.)

Figure 3. Giacomini–Rossi Fluctuation Test Statistic



Authors' calculations using data from Bureau of Labor Statistics, Department of Labor, Federal Reserve Bank of Philadelphia, and Barnichon (2010).

Notes: Relative performance is the fifteen-year rolling difference in MSE between forecasts from SSUR-2 and ARIMA(2,0,1) models. Both models are estimated over a ten-year rolling window. Dashed horizontal line indicates 5 percent critical value. Shaded areas represent periods of business recession as determined by the National Bureau of Economic Research.

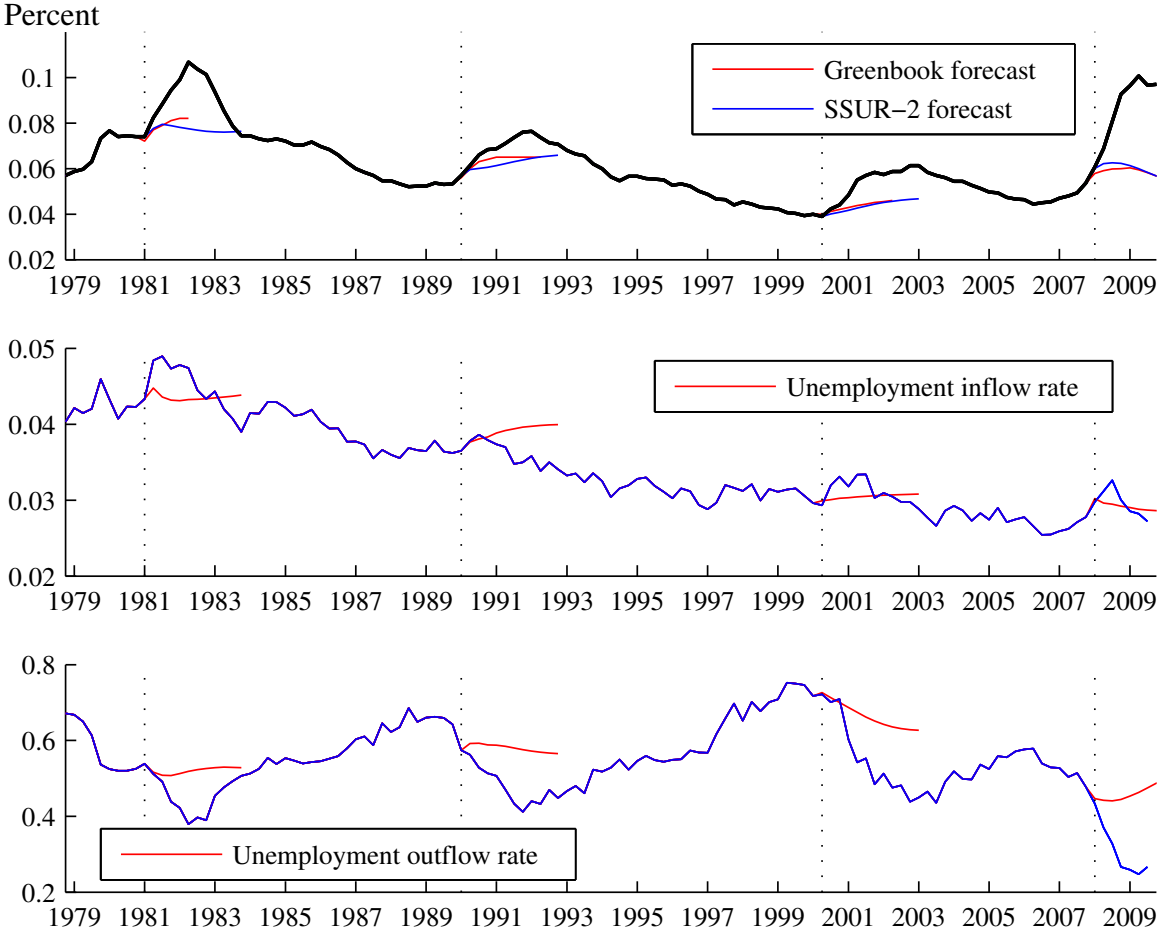
points, precisely when accurate unemployment forecasts are the most valuable.

4.5 Business Cycle Turning Points

To study the performances of the flows model during contractionary periods and before turning points, figures 4 and 5 plot the forecasts from SSUR-2 and the Greenbook over the last four recessions.[22] The figures illustrate the evolution of the forecasts as the recessions gain in momentum. Specifically, for each recession, we consider two different jumping-off points. The first point, shown

22. As discussed previously, the Greenbook forecast is generally viewed as the benchmark forecast, and we thus use it as the benchmark comparison. Because the Greenbook forecasts for the 2007–08 recession are not yet public, we use the SPF forecast.

Figure 4. Unemployment Rate Forecasts at the Onset of Recessions



Authors' calculations using data from Bureau of Labor Statistics, Department of Labor, Federal Reserve Bank of Philadelphia, and Barnichon (2010).

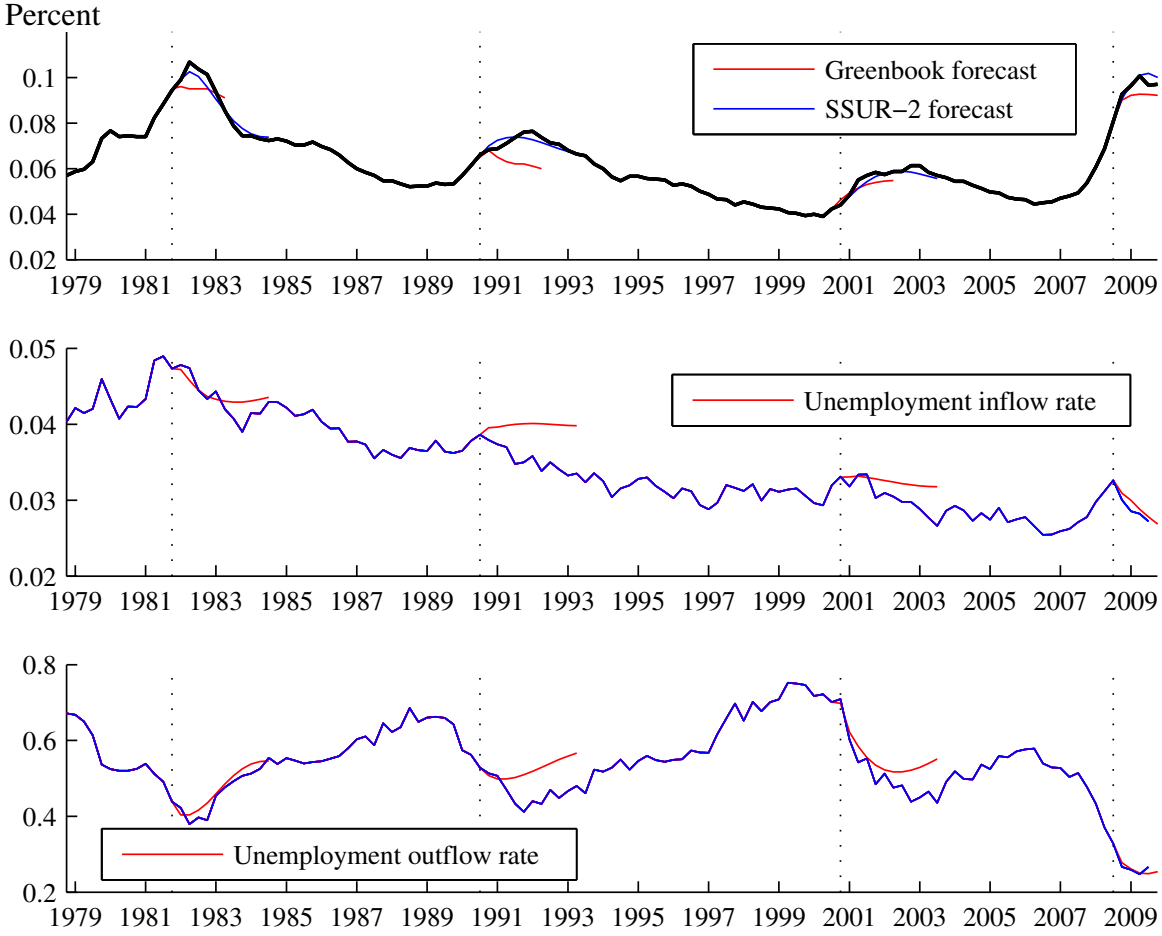
Notes: Dashed vertical lines indicate forecast jump-off dates: 1981m9, 1990m9, 2000m11, and 2008m8. Because Greenbook forecasts after 2006 are not yet public, the 2008m8 forecast is from the SPF.

in figure 4, is at the nascent onset of a recession—roughly before any significant increase in unemployment. The second point, shown in figure 5, corresponds to the peak of the unemployment inflow rate, which is about half-way through the increase in the unemployment rate.[23]

Jumping off from the very early stages of a recession, the SSUR-2 model does not perform significantly better than the Greenbook at horizons of a year or more. Except for the early 1980s recession, the impetus from the inflow rate is too small, and the SSUR-2 model understates the increases in unemployment.

23. While the peak in the inflow rate is unknown in real time, this exercise illustrates how the model can be helpful in identifying turning points ahead of other models.

Figure 5. Unemployment Rate Forecasts Mid-Way through Recessions



Source: BLS and Federal Reserve Bank of Philadelphia data and authors' calculations.

Notes: Dashed vertical lines indicate forecast jump-off dates: 1982m6, 1991m3, 2001m5, 2009m1. Because Greenbook forecasts after 2006 are not yet public, the 2009m1 forecast is from the SPF.

However, once the unemployment rate has risen, the SSUR-2 model is better at identifying turning points. When jumping off roughly mid-way through the increase in unemployment, the model clearly outperforms the Greenbook forecast and is able to predict the turning point in unemployment for the last four recessions as far a year in advance. Indeed, at each of the dates picked (and especially in 1982 and 1990), the Greenbook had projected that the unemployment rate was near its peak for the cycle, when in fact it would continue to rise for some time. In contrast, the SSUR-2 model predicted that the turning point would occur much later—sometimes up to a year later. And for all four recessions, the model's predicted turning point was close to the actual turning point. The model forecast was also much closer to the actual path of unemployment than any of the Greenbook

forecasts.

To get some intuition for the model's performance, the middle and lower panels of figures 4 and 5 plot the behavior of the unemployment inflow and outflow rates around the jumping-off points. The inflow rate is most responsible for the model's superior performance. Because the inflow rate leads the unemployment rate and because bursts of separation are responsible for the sharp increases in unemployment at the onset of recessions, incorporating information from the inflow rate allows the model to capture the fast increase in unemployment during recessions.[24] Thus, SSUR-2 correctly predicts a period of increasing unemployment following all four jumping-off points. In contrast, the Greenbook missed or understated the increase in unemployment in all four recessions.[25] In addition, because the turning point in the inflow rate signals the turning point in the unemployment rate several quarters (and sometimes as far as a year) in advance, the model is able to predict the turning point in the unemployment rate with relatively high confidence several quarters in advance.

5 Combining Forecasts

The array of forecasting models we considered reflect different information sets. The SSUR models' forecasts rely mainly on labor force flows data and other labor market indicators but not on information outside of the labor market. In contrast, the professional forecasts (in particular, the SPF and Greenbook forecasts) are based on an array of economic data and models beyond the labor market, but they may ignore information on labor force flows. The ARIMA model forecasts unemployment from its past behavior.

Given these different information sets, a natural question is whether any additional improvements can be made to the forecasts by combining our flows models with a professional forecast such as the SPF and a simple time-series model such as the ARIMA. To wit, we constructed a new combined forecast, which exploits the differences in correlation among the forecast errors.[26]

This combined forecast was constructed by taking a weighted average of the forecasts from SSUR-2, the SPF and the ARIMA model. The weights were determined by ordinary least-squares regression, with a constant included to account for any systematic biases in the estimate. We estimated weights separately for each forecast horizon. These weights allowed us to evaluate the marginal contributions of each model over the SPF forecast. If the SSUR-2 model forecast had no incremental benefit over the SPF forecast, the weight on the SSUR-2 forecast would be zero.

24. Fujita and Ramey (2009) and Barnichon (2012).

25. As discussed in Montgomery et al. (1998) and Baghestani (2008), this tendency to understate increases in unemployment during recessions is not only a property of the Greenbook forecast, but is an undesirable feature of the SPF and most other models.

26. Granger and Newbold (1986).

Table 3. Optimal Combined Unemployment Rate Forecasts, 1976–2006

	Forecast horizon (quarters)				
Forecast	$t+0$	$t+1$	$t+2$	$t+3$	$t+4$
Root mean squared forecast error (percentage point)					
SSUR-2	0.11	0.33	0.50	0.68	0.82
SPF	0.16	0.35	0.49	0.63	0.77
ARIMA	0.14	0.37	0.59	0.80	0.98
Combined	0.11***	0.30**	0.45*	0.61	0.74
	(0.00)	(0.02)	(0.08)	(0.33)	(0.35)
Optimal weights					
SSUR-2	0.80	0.48	0.42	0.35	0.40
	(0.10)	(0.14)	(0.14)	(0.15)	(0.16)
SPF	0.15	0.42	0.55	0.65	0.63
	(0.08)	(0.11)	(0.12)	(0.13)	(0.14)
ARIMA	0.06	0.11	0.04	0.01	−0.04
	(0.11)	(0.12)	(0.12)	(0.12)	(0.13)
Constant	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Source: Authors' calculations using data from Bureau of Labor Statistics, Department of Labor, Federal Reserve Bank of Philadelphia, and Barnichon (2010).

Notes: Calculated from 124 forecasts. $t+0$ denotes the current-quarter forecast. Upper panel: Numbers in parentheses are p values of Giacomini–White test of equal predictive ability; ***/**/* indicates significantly different from SPF at 1/5/10 percent. Lower panel: regression of $u_t = \beta_0 + \beta_1 \hat{u}_t^{\text{SSUR}} + \beta_2 \hat{u}_t^{\text{SPF}} + \beta_3 \hat{u}_t^{\text{ARIMA}} + \nu_t$; standard errors are reported in parentheses.

As shown in table 3, this was not the case, and combining the SSUR-2 model with the SPF and the ARIMA improved forecasting performance significantly at horizons up to two quarters ahead. The evaluation is a real-time exercise where, for each forecast at a given time, the weights are determined using available history only. Compared to the baseline SPF forecast, the reduction in RMSE achieved by the combined forecast amounted to about 35 percent for current-quarter forecasts, 15 percent for one-quarter-ahead forecasts, almost 10 percent for two-quarter-ahead forecasts, as well as small improvement at longer horizons. The improvements for current-quarter, one-quarter-ahead, and two-quarter-ahead forecasts are statistically better than the SPF forecast alone.

The optimal weights reflect the contribution of the SSUR-2 model for short-term forecasting, and the combined forecast puts a lot more weight on SSUR-2 at short-term horizons.[27] Importantly, the fact that the combined forecast performs significantly better than any single forecast indicates

27. The reported optimal weights are the weights estimated over the full sample.

that the flows models bring relevant information not contained in the SPF or ARIMA forecasts. In other words, because the forecast errors of the models are not strongly correlated, the combined forecast performs substantially better.

6 Forecasting Labor Force Participation

Unlike with the unemployment rate, there is less of a systematic aggregate relationship between labor force participation and output growth. In fact, aggregate participation was largely thought to be acyclical over 1960–2006, where changes in the labor force participation rate were only weakly related to output growth.[28] As a result, forecasting the labor force participation rate was often seen as subordinate to forecasting the unemployment rate.

The large and unexpected decline in labor force participation during and after the 2007–08 recession challenged that conventional wisdom and highlighted the importance of forecasting the labor force participation rate. However, given the historical absence of a strong relation between output and labor force participation, forecasters have few models to turn to.

Thus, one an advantage of the three-state model over the two-state model is that it also generates forecasts of the labor force participation rate (and, by extension, the employment-to-population ratio). Table 4 evaluates the performance of the SSUR-3 model compared to the Greenbook forecasts.[29] SSUR-3 improves on the Greenbook forecast for the current-quarter forecast, although the reduction in RMSE is not statistically significant. At longer forecast horizons, SSUR-3 performs markedly less well than the Greenbook.

However, SSUR-3 forecast errors need not be correlated with Greenbook forecast errors, and a combined forecast may generate significant improvements. The third row of table 4 confirms this intuition. The optimal combined labor force participation rate forecast between the Greenbook and SSUR-3 performs significantly better at all horizons considered, and especially at longer horizons. The reduction in RMSE is trivial for current-quarter forecasts, but grows to between 0.1 to 0.3 percentage point at longer horizons.

The high weight on SSUR-3 at all horizons reflects the superior performances of SSUR-3 compared to the Greenbook. This was initially not apparent in the direct comparison between SSUR-3 and the Greenbook, because SSUR-3 presents a larger bias than the Greenbook. The constant in the

28. Nonetheless several papers identified the importance of demographics for the aggregate participation rate. Aaronson et al. (2006) and Fallick and Pingle (2007) use cohort-based models to help isolate demographic and other structural factors from cyclical variation in the participation rate. They find that the apparent acyclicality of aggregate participation is the result of moderately cyclical participation among certain demographic groups that roughly offsets when aggregated.

29. The historical Greenbook forecasts contain quarterly forecasts for the participation rate beginning only in 2000.

Table 4. Labor Force Participation Rate Forecasts and Optimal Combination, 2000–2006

	Forecast horizon (quarters)				
Forecast	$t+0$	$t+1$	$t+2$	$t+3$	$t+4$
Root mean squared forecast error (percentage point)					
SSUR-3	0.13	0.26	0.38	0.46	0.56
Greenbook	0.14	0.23**	0.31***	0.39*	0.47
	(0.21)	(0.03)	(0.00)	(0.06)	(0.14)
Combined	0.10***	0.15***	0.17***	0.18***	0.19***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Optimal weights					
SSUR-3	0.40	0.36	0.67	0.85	0.69
	(0.16)	(0.16)	(0.14)	(0.12)	(0.10)
Greenbook	0.51	0.35	0.44	0.53	0.50
	(0.14)	(0.14)	(0.11)	(0.09)	(0.07)
Constant	0.06	0.10	0.07	0.00	0.09
	(0.02)	(0.04)	(0.05)	(0.05)	(0.06)

Source: Authors' calculations using data from Bureau of Labor Statistics, Department of Labor, Federal Reserve Bank of Philadelphia, Board of Governors of the Federal Reserve System, and Barnichon (2010).

Notes: Calculated from 56 forecasts over January 2000 to December 2006. $t+0$ denotes the current quarter forecast. Upper panel: Numbers in parentheses are p values of Giacomini–White test of equal predictive ability; ***/**/* indicates significantly different at 1/5/10 percent. Lower panel: regression of $lfpr_t = \beta_0 + \beta_1 \hat{lfpr}_t^{\text{SSUR}} + \beta_2 \hat{lfpr}_t^{\text{SPF}} + \beta_3 \hat{lfpr}_t^{\text{ARIMA}} + \nu_t$; standard errors are reported in parentheses.

optimal forecast accounts for this systematic bias.

7 Recent Performance and Near-Term Prospects

Thus far our evaluation against professional forecasts ended in 2006, shortly before the "Great Recession." A crucial question, however, is how the model performed during the recent recession and the ongoing recovery. In particular, can the model capture the steep increase in unemployment and the lack of rapid decline following this recession compared to other deep recessions?

Beyond 2006, we can no longer compare to the Greenbook forecast and thus use the SPF as a benchmark. As table 1 showed, the SPF and Greenbook forecasts have roughly similar RMSEs over a one-year forecast horizon.

Table 5 reports the RMSE for forecasts starting in February 2007 and ending in February 2012. Although the SPF's current-quarter forecast error is roughly comparable to the earlier period, fore-

Table 5. Performance of Recent Unemployment Rate Forecasts, 2007–12

Root-mean-squared error (percentage point)

	Forecast horizon (quarters)				
Model	$t+0$	$t+1$	$t+2$	$t+3$	$t+4$
SPF	0.18	0.48	0.81	1.20	1.63
SSUR-2	0.17	0.58	0.97	1.50	1.98
SSUR-3	0.13	0.46	0.95	1.53	2.16
No. obs.	21	21	20	19	18

Source: Authors' calculations using data from Bureau of Labor Statistics, Department of Labor, Federal Reserve Bank of Philadelphia, and Barnichon (2010).

Notes: $t+0$ denotes current-quarter forecast.

cast errors at longer horizons are 0.1 to 0.8 percentage point higher during this period than over 1976–2006. The flows models forecasts' are similarly higher. In particular the two-state model's forecast for the current quarter—by far its comparative advantage—worsens appreciably. Whereas it had outperformed the SPF by 25 percent in the earlier period, the SSUR-2 model now performs only slightly better than the SPF.

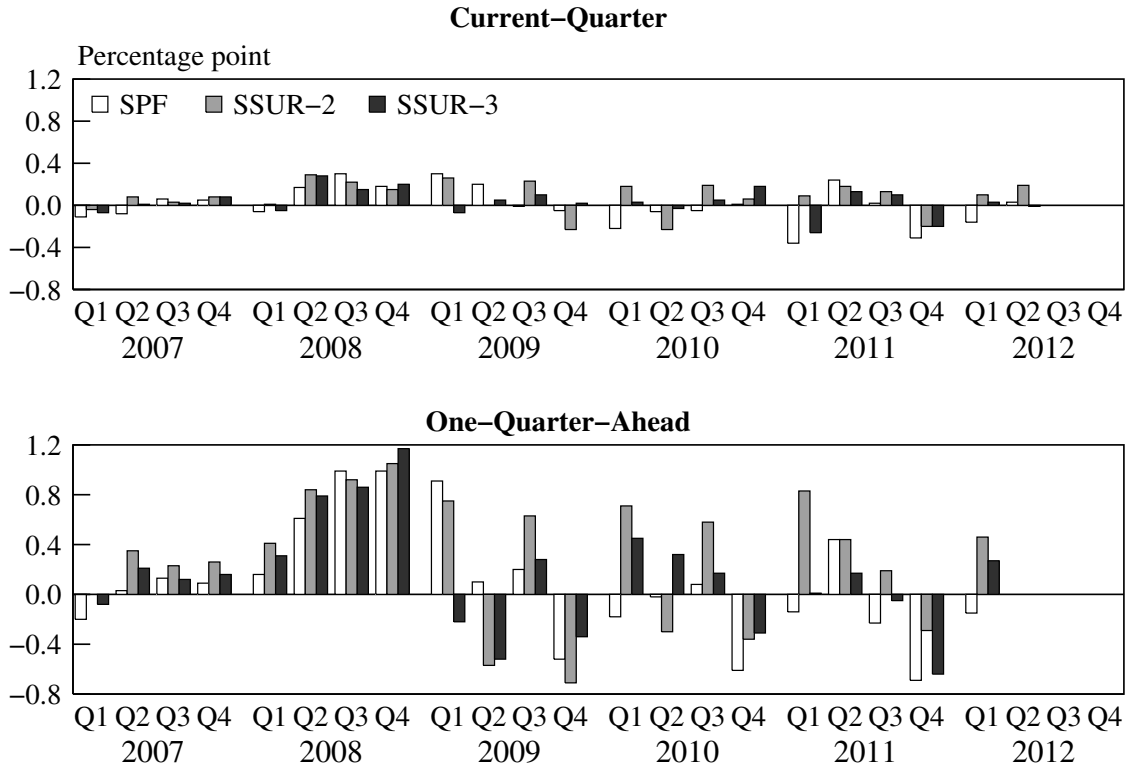
In contrast, the three-state model, which had performed worse than the two-state model and the SPF over 1976–2006 at all horizons, now beats the SPF by almost 30 percent in the current quarter and is even a bit lower one quarter ahead. We point to three factors to explain this striking difference.

First, the gross flows are much better measured in the published data than in the historical tabulations. Because for much of the 1976–2006 sample the model was estimated using the transitions we calculated from the microdata (rather than from the published data), the three-state model performs worse than the two-state model. Indeed, table 2 showed that SSUR-3 performed about the same as the two-state model when estimated using only the published gross flows data.

Second, the two-state model uses cross-sectional data on unemployment to infer the outflow rate and backs out the inflow rate using an unemployment accounting identity. As Elsby et al. (2011) note, a key assumption needed to derive the hazards appears to have broken down starting in 2009. They show that, historically, the two measures of unemployment outflows moved closely together over the business cycle, but that since 2009 Shimer's (2012) measure has exhibited a much larger decline than the unemployment outflow. They show that the discrepancy is being driven by the large increase in the unemployment duration of persons flowing into unemployment, whereas Shimer's calculation assumes that all unemployment inflows have a duration of five weeks or less.

Third, and most important, the two-state model by design abstracts from movements into and out of the labor force. Historically, the labor force participation rate had not exhibited much cyclicality.

Figure 6. Recent Forecast Misses



Source: Authors' calculations using data from Bureau of Labor Statistics, Department of Labor, and Barnichon (2010).

However, in the recent recession and recovery, the participation rate has fallen 2½ percentage points. The Congressional Budget Office's estimate of the trend labor force suggests that only about ½ percentage point of this decline can be accounted for by declining trend participation (primarily due to aging of the population). The remaining 2 percentage points reflects an unusual cyclical decline. The two-state model cannot account for the cyclical decline, and thus projects an employment-population ratio that is systematically too high. In contrast, the flows in the three-state model reflect the declining participation rate, in particular the sizable increase in flows from unemployment to out of the labor force and modest decline in flows into unemployment from out of the labor force.

Figure 6 shows the time pattern of recent forecast misses by the SPF and the two models. The upper panel shows the miss, in percentage points, on the unemployment rate in the current quarter of forecast (two months of forecast). The lower panel shows the miss on the unemployment rate in the next quarter (five months of forecast). A positive value indicates that the unemployment rate was higher than the model predicted. As the unemployment rate started to increase from 4.5

Table 6. Model Forecasts for 2012
Percent

Model	*July*	*Aug.*	*Sept.*	*Oct.*	*Nov.*	*Dec.*	*Q3*	*Q4*
Unemployment rate								
SSUR-2	**8.25**	8.12	8.00	7.87	7.73	7.59	8.13	7.73
SSUR-3	**8.25**	8.20	8.17	8.16	8.16	8.15	8.21	8.16
Labor force participation rate								
SSUR-3	**63.70**	63.83	63.89	63.95	63.99	64.04	63.81	64.00

Source: Authors' calculations using data from Bureau of Labor Statistics, Department of Labor, and Barnichon (2010).
Notes: Bold cells indicate published data.

percent in the first quarter of 2007 to 4.8 percent at the end of 2007, the SPF and the models had relatively small upside surprises at both the $t + 0$ and $t + 1$ horizons (the unemployed was higher than projected). The unemployment rate accelerated over 2008 and the first half of 2009, rising from 5 percent to 9½ percent. In 2008, the models and SPF alike were modestly surprised to the upside in their current-quarter forecasts, with misses of about ¼ percentage point. However, from 2009 forward, the three-state model consistently outperforms the SPF and two-state model in the current quarter, with misses both to the upside and downside. At the one-quarter-ahead horizon, the two-state model does noticeably worse than the SPF or three-state model from 2009 on.

Turning to the models' outlook for the coming months, table 6 presents forecasts for the unemployment rate and labor force participation rate. The two-state model projects the unemployment rate to decline rapidly over the rest of the year, from 8.3 percent in July to 7.6 percent in December. This decline reflects a projected increase in the unemployment outflow hazard and little change in the inflow hazard (not shown). The three-state model projects little improvement in the unemployment rate in 2012, as declines in unemployment are partially offset by an increase in the labor force participation rate.

8 Conclusion

Although the unemployment rate is typically not considered a leading or coincident indicator, increases in the unemployment rate have preceded the last three recessions. Recent research by Fleischman and Roberts (2011) finds that the unemployment rate provides the best single signal about the state of the business cycle in real time. Nevertheless, despite extensive research on the topic, forecasters and policymakers often rely on Okun's law or basic time-series models to forecast the

unemployment rate.

This paper presents a nonlinear model for forecasting the unemployment rate based on labor force flows that, in real time, dramatically outperforms basic times-series models, the SPF, and the Federal Reserve Board's Greenbook forecast at short horizons. The model is based on two principles: (1) The unemployment rate converges to its conditional steady-state value in three to five months according to a nonlinear law of motion, and (2) the labor force flows have different time-series properties.

Empirically, the two-state model has a root-mean-squared-forecast error about 30 percent lower than the next-best forecast for current-quarter, and 10 percent lower for next-quarter forecast. Our model also does a good job at identifying turning points several quarters ahead of other models and forecasters. In addition, because the model brings new information to the forecast, a combined forecast including our model and the SPF forecast yields improvement of about to 35 percent for current-quarter forecast, 25 percent for next quarter forecast, almost 10 percent for two-quarter-ahead forecasts, as well as slight improvements at longer horizons. Additionally, our model has the highest predictive ability surrounding business cycle turning points and large recessions.

The two new models that we propose have both advantages and disadvantages. The two-state model is easier to understand conceptually and to implement. The duration-based unemployment inflow and outflow hazard rates have a longer history and are somewhat less noisy. However, the hazard rates are not directly measured but rather inferred from a theoretical model. More important, a key assumption for deriving the hazards appears to have broken down starting in 2009.

The three-state model is a more realistic characterization of the labor market and the model produces internally consistent forecasts for the unemployment rate, labor force participation rate, and employment-population ratio. In addition, since 2007, the three-state model outperforms the two-state model, in part because it accounts for the unprecedented large decline in labor force participation during this cycle.

References

Aaronson, Stephanie, Bruce Fallick, Andrew Figura, Jonathan Pingle, and William Wascher. 2006. "The Recent Decline in the Labor Force Participation Rate and Its Implications for Potential Labor Supply." *Brookings Papers on Economic Activity*, 37(1): 169–54.

Baghestani, Hamid. 2008. "Federal Reserve versus Private Information: Who is the Best Unemployment Rate Predictor?" *Journal of Policy Modeling*, 30(1): 101–10.

Barnichon, Regis. 2010. "Building a Composite Help-Wanted Index." *Economics Letters*, 109(3): 175–78.

Barnichon, Regis. 2012. "Vacancy Posting, Job Separation, and Unemployment Fluctuations." *Journal of Economic Dynamics and Control*, 36(3): 315–30.

Barnichon, Regis, and Andrew Figura. 2010. "What Drives Movements in the Unemployment Rate?" Federal Reserve Board of Governors Finance and Economics Discussion Series paper 2010-48.

Blanchard, Olivier Jean, and Peter Diamond. 1990. "The Cyclical Behavior of the Gross Flows of U.S. Workers." *Brookings Papers on Economic Activity*, 1990(2): 85–155.

Brown, Laura, and Saeed Moshiri. 2004. "Unemployment Variation over the Business Cycles: A Comparison of Forecasting Models." *Journal of Forecasting*, 23(7): 497–511.

DeLong, J. Bradford, and Lawrence H. Summers. 1986. "Are Business Cycles Symmetric?" In *The American Business Cycle, Continuity and Changes*. , ed. Robert J. Gordon. University of Chicago Press.

Diebold, Francis X, and Roberto S Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics*, 13(3): 253–63.

Elsby, Michael W. L., Bart Hobijn, and Ayşegül Şahin. 2011. "Unemployment Dynamics in the OECD." Federal Reserve Bank of San Francisco Working paper.

Elsby, Michael W. L., Bart Hobijn, Ayşegül Şahin, and Robert G. Valletta. 2011. "The Labor Market in the Great Recession—An Update to September 2011." *Brookings Papers on Economic Activity*, 42(2): 353–71.

Elsby, Michael W. L., Ryan Michaels, and Gary Solon. 2009. "The Ins and Outs of Cyclical Unemployment." *American Economic Journal: Macroeconomics*, 1(1): 84–110.

Fallick, Bruce, and Jonathan Pingle. 2007. "A Cohort-Based Model of Labor Force Participation." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series paper 2007-09.

Faust, Jon, and Jonathan H. Wright. 2007. "Comparing Greenbook and Reduced Form Forecasts using a Large Realtime Dataset." National Bureau of Economic Research, Inc NBER Working Papers 13397.

Fleischman, Charles A., and John M. Roberts. 2011. "From Many Series, One Cycle: Improved Estimates of the Business Cycle from a Multivariate Unobserved Components Model." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series paper 2011-46.

Fujita, Shigeru, and Garey Ramey. 2009. "The Cyclicality Of Separation And Job Finding Rates." *International Economic Review*, 50(2): 415–30.

Giacomini, Raffaella, and Barbara Rossi. 2010. "Forecast Comparisons in Unstable Environments." *Journal of Applied Econometrics*, 25(4): 595–620.

Giacomini, Raffaella, and Halbert White. 2006. "Tests of Conditional Predictive Ability." *Econometrica*, 74(6): 1545–78.

Golan, Amos, and Jeffrey M. Perloff. 2004. "Superior Forecasts of the U.S. Unemployment Rate Using a Nonparametric Method." *The Review of Economics and Statistics*, 86(1): 433–38.

Granger, Clive W., and Paul Newbold. 1986. *Forecasting Economic Time Series.* New York:Academic Press.

Milas, Costas, and Philip Rothman. 2008. "Out-of-Sample Forecasting of Unemployment Rates with Pooled STVECM Forecasts." *International Journal of Forecasting*, 24(1): 101–21.

Mitchell, Wesley C. 1927. *Business Cycles: The Problem of Its Setting.* Cambridge, MA:National Bureau of Economic Research.

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

Neftçi, Salih N. 1984. "Are Economic Time Series Asymmetric Over the Business Cycle?" *Journal of Political Economy*, 92(2): 307–28.

Nekarda, Christopher J. 2009. "Understanding Unemployment Dynamics: The Role of Time Aggregation." Board of Governors of the Federal Reserve System Unpublished paper.

Nekarda, Christopher J. 2012. "The Longitudinal Population Database." Board of Governors of the Federal Reserve System Unpublished paper.

Petrongolo, Barbara, and Christopher A. Pissarides. 2008. "The Ins and Outs of European Unemployment." *American Economic Review*, 98(2): 256–62.

Romer, Christina D., and David H. Romer. 2000. "Federal Reserve Information and the Behavior of Interest Rates." *American Economic Review*, 90(3): 429–57.

Rothman, Philip. 1998. "Forecasting Asymmetric Unemployment Rates." *The Review of Economics and Statistics*, 80(1): 164–168.

Shimer, Robert. 2012. "Reassessing the Ins and Outs of Unemployment." *Review of Economic Dynamics*, 15(2): 127–48.

Sichel, Daniel E. 2007. "Business Cycle Asymmetry: A Deeper Look." *Economic Inquiry*, 31(2): 224–36.

Sims, Christopher A. 2002. "The Role of Models and Probabilities in the Monetary Policy Process." *Brookings Papers on Economic Activity*, 33(2): 1–62.

Tulip, Peter. 2009. "Has the Economy Become More Predictable? Changes in Greenbook Forecast Accuracy." *Journal of Money, Credit and Banking*, 41(6): 1217–231.

Appendix

Solution to Three-State Model

Denoting $Y_{t+\tau} = (U_{t+\tau}, E_{t+\tau}, N_{t+\tau})'$, we can rewrite equation 10 as

$$\dot{Y}_{t+\tau} = A_t Y_{t+\tau}, \tag{A.1}$$

with

$$A_t = \begin{pmatrix} -\lambda_t^{UE} - \lambda_t^{UN} & \lambda_t^{EU} & \lambda_t^{NU} \\ \lambda_t^{UE} & -\lambda_t^{EU} - \lambda_t^{EN} & \lambda_t^{NE} \\ \lambda_t^{UN} & \lambda_t^{EN} & -\lambda_t^{NE} - \lambda_t^{NU} \end{pmatrix}.$$

Since the columns of A_t sum to zero, A_t has one eigenvalue equal to zero. Denoting R_t the matrix of eigenvectors of A_t corresponding to the eigenvalues $\{r_{1t}, r_{2t}, 0\}$, a solution to equation A.1 is

$$Y_{t+\tau} = R_t \begin{pmatrix} c_1 e^{\tau(r_{1t})} \\ c_2 e^{\tau(r_{2t})} \\ c_3 \end{pmatrix} \tag{A.2}$$

with c_1, c_2 and c_3 the constants of integration. The two nonzero eigenvalues are negative and are functions of the hazard rates:

$$\begin{aligned} r_{1t} \approx -\beta_{1t} &\equiv \lambda_t^{UE} + \lambda_t^{UN} \\ r_{2t} \approx -\beta_{2t} &\equiv \lambda_t^{EU} + \lambda_t^{EN} + \lambda_t^{NE} + \lambda_t^{NU}. \end{aligned} \tag{A.3}$$

To find the values of c_1, c_2 and c_3, we use initial conditions $Y_t = (U_t, E_t, N_t)'$ and terminal conditions $Y_t \underset{t\to\infty}{\longrightarrow} (U_t^*, E_t^*, N_t^*)'$, the vector of the steady-state numbers of unemployed (U_t^*), employed (E_t^*), and nonparticipants (N_t^*). The steady-state stocks are given by

$$\begin{aligned} U_t^* &= k\frac{s_{t+1}}{s_{t+1} + f_{t+1} + o_{t+1}} \\ E_t^* &= k\frac{f_{t+1}}{s_{t+1} + f_{t+1} + o_{t+1}} \\ N_t^* &= k\frac{o_{t+1}}{s_{t+1} + f_{t+1} + o_{t+1}}, \end{aligned}$$

where k is a constant set so that U_t^*, E_t^*, and N_t^* sum to P_t, the working age population in month t; and s_{t+1}, f_{t+1}, and o_{t+1} defined by

$$\begin{cases} s_{t+1} = \lambda_{t+1}^{EN}\lambda_{t+1}^{NU} + \lambda_{t+1}^{NE}\lambda_{t+1}^{EU} + \lambda_{t+1}^{NU}\lambda_{t+1}^{EU} \\ f_{t+1} = \lambda_{t+1}^{UN}\lambda_{t+1}^{NE} + \lambda_{t+1}^{NU}\lambda_{t+1}^{UE} + \lambda_{t+1}^{NE}\lambda_{t+1}^{UE} \\ o_{t+1} = \lambda_{t+1}^{EU}\lambda_{t+1}^{UN} + \lambda_{t+1}^{UE}\lambda_{t+1}^{EN} + \lambda_{t+1}^{UN}\lambda_{t+1}^{EN}. \end{cases}$$

Some algebra yields the one-month-ahead forecasts of unemployment, employment, and non-participation:

$$\begin{aligned} U_{t+1} &= p_{11}c_1\mathrm{e}^{-\beta_{1t}} + p_{12}c_2\mathrm{e}^{-\beta_{2t}} + U_t^* \\ E_{t+1} &= p_{21}c_1\mathrm{e}^{-\beta_{1t}} + p_{22}c_2\mathrm{e}^{-\beta_{2t}} + E_t^* \\ N_{t+1} &= p_{31}c_1\mathrm{e}^{-\beta_{1t}} + p_{32}c_2\mathrm{e}^{-\beta_{2t}} + N_t^* \end{aligned} \tag{A.4}$$

with c_1 and c_2 given by

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}^{-1} \times \begin{pmatrix} U_t \\ E_t \end{pmatrix}.$$