w o r k i n g

p a p e r

15 19R

Forecasting Inflation: Phillips Curve Effects on Services Price Measures

Ellis W. Tallman and Saeed Zaman

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 15-19R September 2016*

Forecasting Inflation:
Phillips Curve Effects on Services Price Measures

Ellis W. Tallman and Saeed Zaman

We estimate an empirical model of inflation that exploits a Phillips Curve relationship between a measure of unemployment and a subaggregate measure of inflation (services). We generate an aggregate inflation forecast from forecasts of the goods subcomponent separate from the services subcomponent, and compare the aggregated forecast to the leading time-series univariate and standard Phillips curve forecasting models. Our results indicate marked improvements in point and density forecasting accuracy statistics for models that exploit relationships between services inflation and the unemployment rate.

Keywords: Inflation forecasting, Phillips curve, disaggregated inflation forecasting models, trend-cycle model, density combinations.

JEL Codes: C22, C53, E31, E37.

Suggested citation: Tallman, Ellis W., and Saeed Zaman, 2015. "Forecasting Inflation: Phillips Curve Effects on Services Price Measures," Federal Reserve Bank of Cleveland, working paper no. 15-19R.

Ellis Tallman and Saeed Zaman are at the Federal Reserve Bank of Cleveland. Correspondence about the paper may be directed to saeed.zaman@clev.frb.org.
The authors are very grateful for valuable comments and suggestions from the editor, associate editor, two anonymous referees, Todd Clark, Edward Knotek, James Morley, Andrea Stella, and the seminar and conference participants at the 24th Symposium of the Society for Nonlinear Dynamics and Econometrics in Tuscaloosa, the 36th International Symposium on Forecasting in Santander, the third annual conference of the International Association of Applied Econometrics in Milan, and the Federal Reserve System Committee on Macroeconomics meeting Day-Ahead Inflation workshop. They are grateful to Andrea Stella for letting them use his computer code for estimating their state-space model as described in his work with Jim Stock.

*Original version posted October 2015.

1 Introduction

Forecasting models of aggregate inflation including those that employ a Phillips Curve have been unable to outperform consistently univariate statistical models of aggregate inflation (Faust and Wright 2013).[1] Peach, Rich, and Linder (2013) suggest that the empirical estimates of a Phillips Curve may be diluted when applied to aggregate inflation because the influence of resource gap factors (such as the difference between measured unemployment and its "natural" rate) may affect the costs of services (services as non-tradable) more directly and substantially than on the costs of goods. Similarly, using data from New Zealand Hargreaves, Kite, and Hodgetts (2006) demonstrate how a Phillips Curve relationship is important for modeling inflation for non-tradable prices and therefore model tradeables and nontradeables prices separately. In each of these cases, the supporting intuition centers on how the key factors that influence tradeables (or goods) can differ materially from factors that affect prices in non-tradeables (services). Following from these ideas, we hypothesize that a resource gap measure has an important effect on services price determination, and not on goods price inflation.

In this paper, we build a composite model for inflation that consists of bi-variate state space model (unobserved components) of services inflation and the unemployment rate combined with a parsimonious univariate model for goods inflation. The services inflation model adapts the bivariate state space model as in Stella and Stock (2015) and exploits an empirical relationship between services inflation and the unemployment gap.[2] The forecasting model for aggregate inflation in this paper captures the apparent relationship in that unemployment rate deviations from trend (a latent variable estimate of the 'natural rate') appear useful for predicting services inflation.

We estimate an inflation in parts model in which we separately measure services inflation and goods inflation. Using these two inflation series separately, the model isolates a durable statistical relationship between services inflation and the unemployment rate. The bivariate state space model of services inflation exploits the empirical Phillips Curve correlation suggested in Peach, Rich, and Linder (2013). From the estimated model, we generate forecasts of services inflation and we combine it with the goods inflation forecast from an estimated trend in goods inflation (i.e. trend cycle decomposition with stochastic volatility along the lines of Stock and Watson, 2007)[3] to compute a composite forecast of the aggregate inflation.[4] We then evaluate the forecasts of aggregate inflation, services in-

[1]Here we interpret the Phillips Curve as being the correlation between deviations of unemployment from its natural rate and deviations of inflation from its trend or expected rate.

[2]See Peach, Rich, and Linder (2013). They employ an empirical approach to the data that differs from our methods.

[3]Goods inflation trend estimated using 5-year average or exponential smoothing method with alpha=0.05 leads to much more accurate point forecasts of goods inflation and resulting aggregate inflation forecasts. The results are presented in the Online Appendix A.4.

[4]Among alternative models of goods inflation, we estimated one that specifically extends the bivariate UC state space model to three variables. The resulting tri-variate unobserved components state space model consisted of: unemployment rate, services inflation, and goods inflation. Separate models for services inflation and goods inflation produces substantially more accurate forecasts than a single model

flation (from the bi-variate state space model), and goods inflation (parsimonious model).

We find that modeling inflation sub-components separately (goods as a univariate time-series model and services inflation in a state space form with unemployment) produces significant improvements in both point and density forecast accuracy for aggregate inflation relative to a forecasting application of Stella and Stock's bivariate model for total inflation.[5] The model provides a good benchmark because it modestly outperforms relatively accurate univariate approaches such as Atkeson and Ohanian (2001) and Stock and Watson (2007).[6] Similarly, our "inflation in parts" framework performs well in terms of forecast accuracy of aggregate inflation relative to common univariate benchmarks that are usually the most accurate in terms of root mean squared error and other point-based accuracy criteria. We also demonstrate forecast accuracy improvements for density forecasts using our inflation in parts framework, consistent with Ravazzolo and Vahey (2014) who generate forecasts of inflation that are aggregated from sub-components of inflation.[7]

Our empirical model estimates a Phillips curve relationship between services inflation and unemployment rate as evidenced by a relatively stable and negative estimate of the slope, a result consistent with Lee and Nelson (2007), Stella and Stock(2015), and King and Watson (1994).[8] Our results contrast with Atkeson and Ohanian (2001) and Stock and Watson (2007), for example, those papers document the failure of standard Phillips curve models in forecasting aggregate inflation compared to simple univariate benchmarks in the post 1990 period.

The results employ the (headline) Personal Consumption Expenditures price deflator (PCE) as the measure for inflation, and unemployment rate as measure from which we infer economic slack. We focus on headline PCE inflation because Federal Reserve's long-run objective for inflation is defined in terms of this measure. We also generate results for four additional combinations: (i) PCE inflation and short-term unemployment rate (results provided in online Appendix A.1)[9] (ii) (headline) Consumer Price Index (CPI) inflation and unemployment rate (provided in the online Appendix A.2), (iii) CPI inflation and short-term unemployment rate (available on request from authors). (iv) CPI inflation excluding food and energy (core CPI) and unemployment rate (provided in the

of the aggregate. These results are consistent with Peach, Rich, and Linder (2013).

[5]The specification in Stella and Stock (2015) exploits the Phillips curve relationship between total unemployment rate and total inflation. We apply their specification to services inflation with the unemployment rate.

[6]We note that the forecast accuracy improvement for the Stella and Stock model forecasts are mostly limited to short-horizons.

[7]Ravazzolo and Vahey (2014) disaggregates PCE into 16 sub-components, each of which is modeled as a univariate autoregressive process. They combine the forecasts of these sub-components to form a forecasts of aggregate inflation and show that the density forecasts of that model out-perform density forecasts from a univariate integrated moving average specification of aggregate inflation (their focus was on headline PCE inflation).

[8]Each of the cited works estimates a negative correlation between aggregate inflation and unemployment rate at business cycle frequencies.

[9]Gordon(2013) and Ball and Mazumder (2014) suggest that short-term unemployment rate is a more appropriate measure from which to infer economic slack.

online Appendix A.3). Our finding of forecast accuracy improvements extend to these combinations (i.e. headline CPI inflation, using overall unemployment rate or short-term unemployment rate) with one exception. For "core" CPI inflation results are mixed. The metrics for forecast accuracy (for both point and density) support the notion that the model remains among the more accurate forecasting models under consideration. However, an alternative model that exploits the Phillips curve relationship between aggregate core inflation and unemployment rate can forecast about as accurately as the model in parts.

The paper is organized as follows: Section 2 describes the relevant literature, Section 3 outlines the data and empirical strategy. Section 4 discusses the estimation and forecasting results, and Section 5 concludes.

2 Review of Relevant Literature

Existing research indicates the potential for forecasting accuracy improvement from distinguishing between inflation of goods and of services in models of inflation. Peach, Rich, and Antoniades (2004) use a vector error correction model to estimate goods inflation and services inflation separately while also imposing a long-run relationship between the two measures. The paper demonstrates that the short-run to medium-run dynamics of the two inflation series depend on the deviation of the long-run equilibrium between the two inflation rates. Their empirical model consists only of the lags of good and services inflation as explanators, and therefore they do not investigate the Phillips Curve.

Clark (2004) provides a qualitative analysis of the behavior of core goods and core services inflation as measured by PCE price index. He identifies 1994 as the year in which the dynamics of the two series began to diverge.[10] Given differences in the dynamics of the two series, there maybe benefits modeling each series separately, and then combine the disaggregate forecasts as an alternative to modeling the aggregate directly.

More recently, Peach, Rich, and Linder (2013) stress the importance of separately modeling goods and services inflation and show that in the case of core goods inflation, the unemployment gap does not play any material role. In contrast, the unemployment gap has significant influence on core services inflation.[11] In their study, they emphasize that the behavior and determinants of services and goods inflation are distinct, and they limit their forecast comparison to their composite model versus the standard Phillips curve model from 2005 to 2012.

[10] Clark(2004) attributes this shift in dynamics to both the exchange rate and the increase in global competition.

[11] Peach, Rich, and Linder also suggest that long-run inflation expectations play an important role for the behavior of core services inflation but not for core goods inflation. In the case of core goods inflation they find that short-term inflation expectations (one-year out) plays a material role in explaining its time-series behavior.

Existing research is divided on whether there are forecasting accuracy improvements from modeling aggregate inflation using sub-components of the aggregate inflation measure. In general, exising evidence suggests there may be gains from modeling disaggregates of the index using data from the United States. Bermingham and DAgostino (2011) empirically test the forecast performance of time series models with multiple-level of dis-aggregation both for US PCE inflation and Euro Area inflation. They employ Bayesian Vector Auto Regression (BVAR) and simple Auto Regressive (AR) models. They find in the USA case, dis-aggregated BVAR (especially the 15-component) generally performs well relative to random walk benchmark. Whereas, in the case of Euro Area, disaggregation through the AR models works better. In the USA case, the authors suggest that strong common co-movement among the dis-aggregated series are captured by BVARs (multivariate) (consistent with Reis and Watson (2010)). Measures for the Euro Area display less commonality among the dis-aggregated components and more individual series dynamics, which is consistent with superior performance of aggregation based on AR models.

Hubrich (2005), and Hendry and Hubrich (2006,2011) find that forecasting aggregate inflation through disaggregation does not help in forecasting Euro Area inflation but it helps for US inflation. The degree of improvement in the forecast accuracy of US inflation importantly depends on the length of the sample period and the level of the disaggregation. They use set of VARs and include disaggregates directly in the model that has aggregate inflation. Following similar approach, Luetkepohl (2010) estimates system of VARs that include both aggregate and disaggregate information for Euro Area inflation only and finds that including too many disaggregates could lead to estimation error and specification error.

Espasa et al (2002) found favorable results in forecasting aggregate monthly U.S. CPI inflation by separately modeling the disaggregates using univariate ARIMA processes. The disaggregates they employ included food, energy, core goods, and core services. Aron and Muelbauer (2008) show evidence of gains in forecasting monthly PCE price index using disaggregates price indices: durable goods, non-durable goods, and services. They use single equation error correction model (ECM) for each of the component with same predictors and the combined forecast outperforms the forecast of the aggregate PCE generated from the ECM with the same predictors as used for the disaggregates over the evaluation sample 2000 to 2007. Faust and Wright (2013) found no material differences between the aggregate and disaggregate approaches unless parameter restrictions were imposed on the disaggregate equations.

Furthermore, country specific studies of inflation such as Duarte and Rua (2007), Bruneau et al (2007) and Moser et al (2007) use similar time series models and find that generating aggregate inflation forecasts using disaggregates helps for Portugal, France, and Austria respectively.

Stock and Watson (2015) documents evidence of estimating a superior trend inflation by using dis-aggregated 17-components of aggregate PCE inflation relative to trend estimated from the univariate model of aggregate PCE inflation. They use a multivariate extension of the univariate unobserved components stochastic volatility model of trend inflation in Stock and Watson (2007).

The studies cited above focus on point forecast accuracy metrics; Ravazzolo and Vahey(2014) focus on density forecasting. They show that for headline PCE inflation density forecasts computed by combining 16 dis-aggregate components (with weights based on linear opinion pool) are more accurate than those from the integrated moving average specification and AR2 model for aggregate inflation. They model each of the disaggregate series as a univariate autoregressive AR2 process with constant volatility. To our knowledge their study is the only one that looked at the density forecasts for U.S. inflation based on disaggregated component data.

The evidence of forecasting improvements by modeling inflation using its components motivates our paper. We build on the results from the aforementioned studies, especially Peach, Rich and Linder (2013). In light of ample evidence documenting an important role of stochastic volatility in the inflation process and forecast accuracy (e.g. Stock and Watson 2007; Clark 2011), we allow for time variation in the variance of the innovations to various components, and that in turn implies time-varying relationship between changes in services inflation and unemployment rate.[12] We perform extensive forecast evaluation exercises, specifically comparing our model's forecast performance against a number of popular alternative models as well as across sub-samples. We use only two major components of aggregate inflation (services and goods inflation). The final product is a composite model of services inflation and goods inflation to forecast aggregate inflation.

3 Data and the Model

We employ the following quarterly data series in this research: the overall unemployment rate (16 years and over), the short-term unemployment rate (share of labor force unemployed for 26 weeks or less), Personal Consumption Expenditures deflator (PCE), two components of the PCE deflator – the services component as well as the goods component, and these two components' relative shares of Personal Consumption Expenditures deflator.[13] The unemployment rate(s) series are quarterly average of the monthly series available from the Bureau of Labor Statistics, and PCE inflation rate(s) including their shares are available from Bureau of Economic Analysis. The sample starts in 1960:Q1, and goes through 2014:Q4. We are interested in forecasting the quarterly annualized

[12]Moreover, our framework accounts for a slowly varying local mean for services inflation, which is an important feature of accurate inflation forecasts (Faust and Wright 2013).

[13]For robustness, we use Consumer Price Index (CPI), its two components goods and services components along with these two components' relative shares of CPI. Also used in one of the robustness exercises is core CPI, core CPI goods (i.e.commodities less food and energy), core CPI services (i.e. services less energy services), along with these two components' relative share of core CPI. CPI data is available from the Bureau of Labor Statistics.

Personal Consumption Expenditures deflator in the aggregate, so we combine our quarterly forecasts of services and goods components using the actual weights to produce the aggregate inflation forecast. The weights represent the relative share of services and goods in the personal consumption expenditures.

Let U_t represent the unemployment rate for the aggregate data series and let U_t^N represent the natural rate of unemployment, which may vary over time. The natural rate of unemployment is a latent variable in this approach and will be estimated as part of the model estimation. We use P_t^s as the price level measure for services, P_t^g as the price level measure for goods, and P_t^T as the aggregate price level measure. For inflation as measured by these series, we represent the quarterly annualized rate (computed as 400 times the natural log difference in price levels) in the services component by π_t^s and in the goods component by π_t^g.

We assume that both services inflation and unemployment rate data series can be modeled as an unobserved components, that is, the sum of a long-term trend component that is a random-walk, a stationary transitory (cyclical) component, and measurement error component. We assume a common cyclical component. The model is estimated with Bayesian Gibbs sampler.

The model specifies the latent variables (the unobserved trend and cyclical components) within a state-space form of the time series model. We outline the specification below, let:

U_t represent the aggregate unemployment rate
U_t^N represent the natural rate of unemployment
U_t^C represent the cyclical component of the unemployment rate
π_t represent the aggregate inflation
π_t^s represent the services inflation component
$\pi_t^{s,*}$ represent the trend services inflation
$\pi_t^{s,C}$ represent the cyclical services inflation component
π_t^g represent the goods inflation component

The variables $U_t,\ \pi_t^s,\ and\ \pi_t^g$ are observable, and the other four measures are unobservable.

3.1 Modeling Services Inflation

Our setup follows closely that of Lee and Nelson (2007) and Stella and Stock (2015). We view the multivariate unobserved components model in Stella and Stock (2015) as state of the art so we use that approach as applied to modeling services inflation. Specifically, the services inflation rate is modeled as a sum of three unobserved stochastic processes: random walk trend component, a stationary cyclical component (common also to cyclical unemployment), and a serially uncorrelated measurement error component.[14] Similarly, the unemployment rate, is decomposed into the random walk trend, a stationary cyclical unemployment, and a measurement error component. The stochastic trend unemployment rate can be interpreted as the natural rate of unemployment (or alternatively as the non-accelerating inflation rate of unemployment or NAIRU), and is usually assumed to evolve independently of monetary policy actions. Other studies that have incorporated time varying random walk trends in both unemployment rate and inflation include Lee and Nelson (2007), Harvey (2011), and Stella and Stock (2015).

$$U_t = U_t^N + U_t^C + \eta_t \tag{1}$$
$$\pi_t^s = \pi_t^{s,*} + \pi_t^{s,C} + \tilde{\eta}_t \tag{2}$$

The trend or "permanent" components are modeled as:

$$U_t^N = U_{t-1}^N + \epsilon_t, \qquad \epsilon_t \quad i.i.d. \;\; N(0, \omega I_1) \tag{3}$$
$$\pi_t^{s,*} = \pi_{t-1}^{s,*} + \tilde{\epsilon}_t, \qquad \tilde{\epsilon}_t = \sigma_{\tilde{\epsilon},t}\xi_{\tilde{\epsilon},t} \qquad \xi_{\tilde{\epsilon},t} \quad i.i.d. \;\; N(0, I_1) \tag{4}$$
$$ln(\sigma^2_{\tilde{\epsilon},t}) = ln(\sigma^2_{\tilde{\epsilon},t-1}) + \nu_{\tilde{\epsilon},t} \qquad \nu_{\tilde{\epsilon},t} \quad i.i.d. \;\; N(0, \gamma I_1) \tag{5}$$

Our specification allows the variance of the innovations to the services inflation trend to vary over time. However, as in Stella and Stock (2015) we don't allow stochastic volatility in the innovations to trend unemployment ($\omega = 0.01$).[15]

The cyclical components are modeled as follows:

$$U_t^C = \alpha_1 U_{t-1}^C + \alpha_2 U_{t-2}^C + \zeta_t \qquad \zeta_t = \sigma_{\zeta,t}\xi_{\zeta,t} \qquad \xi_{\zeta,t} \quad i.i.d. \;\; N(0, I_1) \tag{6}$$
$$ln(\sigma^2_{\zeta,t}) = ln(\sigma^2_{\zeta,t-1}) + \nu_{\zeta,t} \qquad \nu_{\zeta,t} \quad i.i.d. \;\; N(0, \gamma I_1) \tag{7}$$
$$\pi_t^{s,C} = \lambda U_t^C \tag{8}$$

The cyclical services inflation component is a function of unemployment cycle. The parameter λ relates the unemployment gap and the services inflation gap across business cycle frequencies. The estimated parameters determine the importance of a Phillips curve relationship. Specifically, there is time-variation in the innovation variances to the various latent components of the multivariate UC model described above. This implies a corresponding time-varying bivariate vector autoregressive (VAR) in the changes in unemployment gap and services inflation gap. At each point in time, the particular values

[14]The assumption that the services inflation trend follows a random walk is a reasonable one because the factors that drive it are unknown but are quite persistent, a point emphasized in Lee and Nelson (2007) for the case of aggregate inflation trend.

[15]Our results are not sensitive to this restriction.

of the innovation variances characterize the implied coefficients of the bivariate VAR.[16] The model generates an implied estimate of the slope of the Phillips curve which is the sum of the coefficients on the current and lagged changes in unemployment gap of the services inflation gap equation of this implied bivariate VAR.[17]

Finally, the measurement error components are as follows[18]:

$$\eta_t \qquad i.i.d.\ N(0,\gamma I) \tag{9}$$

$$\tilde{\eta}_t = \sigma_{\tilde{\eta},t}\xi_{\tilde{\eta},t} \qquad \xi_{\tilde{\eta},t}\quad i.i.d.\ N(0,I) \tag{10}$$

$$ln(\sigma^2_{\tilde{\eta},t}) = ln(\sigma^2_{\tilde{\eta},t-1}) + \nu_{\tilde{\eta},t} \qquad \nu_{\tilde{\eta},t}\quad i.i.d.\ N(0,\gamma I) \tag{11}$$

As discussed in Stella and Stock (2015), allowing for stochastic volatility in the measurement error of unemployment equation poses a challenging empirical problem of separately estimating the cyclical unemployment and the related measurement error. To simplify this challenge, we restrict the variance of the measurement error to be constant. Stella and Stock (2015) highlight that the forecasting performance of their model was not sensitive to whether one has constant or time-varying variance in the measurement error, but they show that time-varying variance in the measurement error introduced considerable high-frequency noise to the estimated trend unemployment.

[16] A multivariate unobserved component (UC) model with common trend such as the one laid out above has a companion implied vector autoregressive (VAR) representation of an infinite lag length. This VAR representation results from the Kalman filter's (recursive) predictive filtering algorithm, specifically the estimated predictive filtering weights constitute the coefficients of the corresponding implied VAR model. The predictive weights computed through the Kalman filter depend on the joint autocovariances of the variables of the UC model, which in turn depend on the innovation variances. A UC model that allows for stochastic volatility in the innovation variances would lead to a time-varying joint autocovariance of the variables in the UC model, implying time-varying predictive filtering weights (resulting in time-varying coefficients of the corresponding implied VAR).

[17] This feature of time-variation in the Phillips curve slope is an innovative characteristic of the Stella and Stock model that differentiates their application from most other specifications of Phillips curve models based on unobserved components. Koopman and Harvey (2003), and Harvey (2006) provide generic algorithms for the unobserved components class of models to derive the implied weights of the corresponding VAR model representation.

[18] The scalar parameter γ is fixed at 0.2. With an aim towards forecasting we focused on tighter parametrized models and therefore fixed the γ to a constant value (same as in Stock and Watson 2007 and Stella and Stock 2015). Given that these two models constitutes our benchmark models we fixed γ to the same value used in these two studies to facilitate their comparison. Furthermore, Clark and Doh (2014) show that forecasting results were very similar whether using a fixed γ of 0.2 or estimating it (even though estimated value is slightly higher than 0.2). The Appendix displays evidence on the sensitivity of our results to using different values of the parameter γ that governs the smoothness of stochastic volatility process. Our sensitivity results (reported in the online Appendix A.5) show that on average point forecast evaluation results both quantitatively and qualitatively are very similar for values of γ from 0.2 to 0.6. For density forecasts, as expected, results are little different quantitatively when using the forecast metric log score. However, the results are not so different for most horizons (except for the horizon h=12) when using the forecast metric CRPS. On average the quantitative differences are small enough (except for the longer horizons) that qualitatively the results hold.

The unobserved components models of Stella and Stock (2015) and Lee and Nelson (2007) share common features although there are three major differences between them. Firstly, Stella and Stock (2015) incorporate a measurement error component in addition to random walk trend and stationary cyclical components; secondly, they introduce stochastic volatility in all the unobserved components (total of five); and they estimate the model using Bayesian methods which together with multiple stochastic volatility processes makes it computationally quite intensive to generate recursive forecasts.[19]

The model is estimated using Bayesian Gibbs sampler, and the algorithm is described briefly below:[20]

The prior density for the parameters (α_1, α_2, and λ) is assumed to be normal conjugate with mean of zero, and variance of 100.
Steps:

1. Conditional on observed data (services inflation, and unemployment rate), parameters (α_1, α_2, and λ), and stochastic volatilities draw the unobserved state variables $U^N, \pi^{S,*}$, and U^C
2. Conditional on the unobserved state variables, and stochastic volatilities draw the parameters (α_1, α_2, and λ) from the posterior distribution. The draws of autoregressive coefficients on the lags of the unemployment cyclical component, α_1, and α_2 are subject to linear constraints so to ensure stationary cyclical component.[21]
3. Conditional on the unobserved state variables, and the parameters (α_1, α_2, and λ) we draw stochastic volatilities. This step is based on work by Kim, Shephard, and Chib (1998).

The above steps are repeated for N+"burn in" draws.

At each forecast origin, the model is simulated with N+ "burn in" draws discarding the first "burn in" draws and storing every k^{th} draw.[22] For each draw, the forecasts of

[19] Lee and Nelson (2007) estimate two specifications of the bivariate model, unrestricted, and restricted. The restricted version imposes two sets of restrictions: (1) the cyclical unemployment rate was treated as exogenous with respect to cyclical inflation by setting the coefficients on lags of cyclical inflation to zero in the cyclical unemployment rate equation. (2) they also restrict the autoregressive coefficients of cyclical inflation in its own equation to zero on the premise that by removing the random walk inflation trend, any persistence that still remains is due to the cyclical unemployment rate, which displays considerable inertia. The restricted specification in Lee and Nelson (2007) closely resembles the specification in Stella and Stock (2015) except that the latter introduces stochastic volatility to all the latent components. Using Lee and Nelson framework, our forecasting results remain similar to the baseline results using Stella and Stock techniques, although the baseline is ordinally more accurate. That said, the estimates of the latent components are different, especially the estimate of the trend unemployment rate. The Lee and Nelson (2007) model is, as noted, simpler to implement and interested readers can contact the authors for those results.

[20] We employ the same procedures as found in Stella and Stock (2015)

[21] Specifically, they are constrained to satisfy the following conditions: $\alpha_2 <= 1 - abs(\alpha_1)$, and $\alpha_2 >= -1$, see Morley (1999).

[22] We simulate the model with N=10,000, "burn in"=5,000, and k=5 (i.e. skip interval to reduce

the services PCE inflation are generated by iterating forward the above model equations h periods forward (recursive substitution). The mean forecast of these N forecasts forms our posterior forecast of services PCE inflation.

The h-step ahead forecast of services inflation for draw i, is

$$\hat{\pi}^{s}_{t+h,i} = \hat{\pi}^{s,*}_{t+h,i} + \hat{\pi}^{s,C}_{t+h,i} + \hat{\hat{\eta}}_{t+h,i} \tag{12}$$

$$where \ \ h = 1, ..., 8$$

Accordingly, the mean forecast for the PCE services inflation is:

$$\hat{\pi}^{s}_{t+h} = 1/N \sum_{i=1}^{N} \hat{\pi}^{s}_{t+h,i} \tag{13}$$

3.2 Modeling Goods Inflation

In modeling goods inflation, we adopt a parsimonious approach[23]. That is, we assume that our best forecast for goods inflation next period and beyond is the current estimated trend of goods inflation. To estimate trend inflation, we model goods inflation as a univariate unobserved components process with stochastic volatility (along the lines of Stock and Watson (2007) for aggregate inflation).[24]

Specifically, goods inflation is decomposed into a random walk trend component, and a serially uncorrelated transitory component whose variance is allowed to vary over time.

correlation among draws), resulting in 1,000 retained draws that are used for estimation and forecasting. The model results are very similar if we instead run our exercise with N=30,000, burnin=5,000, k=5 resulting in 5,000 retained draws.

[23]A tri-variate state space model that jointly estimates the relationship between goods inflation, service inflation, and unemployment rate produced inferior forecasts for goods inflation. The model estimates since 1985 suggested weak empirical relationship between the cyclical unemployment rate and cyclical goods inflation. The estimation results of this tri-variate model specification are available on request from the authors.

[24]To estimate the trend in goods inflation, we have also explored various other univariate specifications and assessed their out-of-sample forecasting accuracy for goods inflation over the sample 1985-1993, a time-period that predates the formal forecast evaluation sample. We include various exponential smoothing models, moving average models ranging from one year to six years, HP filter. The exponential smoothing method with alpha=0.05, or alpha=0.15, or 5-year average, or 6-year average all produced forecasts that were more accurate than our choice of UC-SV for goods inflation. We choose UC-SV for its convenience in generating density forecasts. In the online Appendix A.4 we report results if we would have instead selected 5-year average as our model choice for goods inflation. The results are qualitatively similar but quantitatively we get a slightly better payoff in terms of accuracy of point forecasts. Movements in exchange rates and in energy prices influence goods inflation, so we also evaluate the forecasts of goods inflation from a multivariate model (BVAR) that consists of goods inflation, energy inflation, and exchange rates. With the exception of one period ahead, forecasts from this model were generally inferior to the best smoothing procedures

The random walk trend component has a shock term, the variance of which is also allowed to change over time.

$$\pi_t^g = \tau_t^g + \eta_t^g, \qquad where\ \eta_t^g = \sigma_{\eta^g,t}\zeta_{n^g,t} \qquad \zeta_{n^g,t}\ \ is\ \ i.i.d.\ N(0, I_1)$$

$$\tau_t^g = \tau_{t-1}^g + \varepsilon_t^g, \qquad where\ \varepsilon_t^g = \sigma_{\varepsilon^g,t}\zeta_{\varepsilon^g,t} \qquad \zeta_{\varepsilon^g,t}\ \ is\ \ i.i.d.\ N(0, I_1)$$

$$ln(\sigma^2_{\eta^g,t}) = ln(\sigma^2_{\eta^g,t-1}) + \nu_{\eta^g,t}, \qquad where\ \nu_{\eta^g,t}\ \ is\ \ i.i.d.\ N(0, \gamma I_1)$$

$$ln(\sigma^2_{\varepsilon^g,t}) = ln(\sigma^2_{\varepsilon^g,t-1}) + \nu_{\varepsilon^g,t}, \qquad where\ \nu_{\varepsilon^g,t}\ \ is\ \ i.i.d.\ N(0, \gamma I_1)$$

The point forecast of the future goods inflation for next h horizons is the current estimated trend inflation (i.e. filtered estimate of τ_t^g):

$$\hat{\pi}_{t+h}^g = \hat{\tau}_t^g$$

The model is estimated with three-step Metropolis within Gibbs MCMC algorithm:

Step 1. Conditional on the histories of $\sigma^2_{\varepsilon^g,t}$, $\sigma^2_{\eta^g,t}$, and $\gamma = 0.2$ draw the time series of the latent components: goods inflation trend τ_t^g, and transitory component η_t^g. This step utilizes the Kalman forward filter and Durbin and Koopman (2002) algorithm for backward smoother.

Step 2. Conditional on the draws of the latent components in Step 1, and $\gamma = 0.2$ draw the time series of the shock variance to the trend component, $\sigma^2_{\varepsilon^g,t}$.

Step 3. Conditional on the draws of the latent components in Step 1, and $\gamma = 0.2$ draw the time series of the innovation variance corresponding to the transitory component, $\sigma^2_{\eta^g,t}$.

Note, the parameter γ that governs the standard deviation of the shocks to log volatilities can either be estimated or fixed. We have fixed it to 0.2 (please refer to footnote 18). Also, standard deviation of the shocks for two log volatilities does not have to be the same, but here we set it to be equal. If we instead estimate both of them, then the algorithm would become a five-step Metropolis within Gibbs MCMC algorithm.

To compute density forecasts (i.e. predictive density $h(\pi^g_{t+h,t})$ for t+h horizon at forecast origin t) we follow the approach laid out in Clark and Doh (2014):

Simulate the above model specification with 10,000 draws, discard the first 5,000 (burin), and retain the every fifth draw of the remaining 5,000 draws resulting in 1,000 retained draws. These retained draws correspond to draws of τ_t^g, $\sigma^2_{\eta^g,t}$, and $\sigma^2_{\varepsilon^g,t}$ for periods 1 to t, where t is the forecast origin.

For each retained draw i, do the following:

1. Draw innovation to volatility of the transitory component for period t+1 to t+h using a normal distribution with variance γ^2(=0.04). Feed these innovation draws to the random walk process of $\log(\sigma^2_{\eta^g,t})$ to generate time series $\log(\sigma^2_{\eta^g,t+1})$,....,$\log(\sigma^2_{\eta^g,t+h})$.
2. Draw innovation to volatility of the shocks to trend component for period t+1 to t+h using a normal distribution with variance γ^2(=0.04). Feed these innovation draws to the random walk process of $\log(\sigma^2_{\varepsilon^g,t})$ to generate time series $\log(\sigma^2_{\varepsilon^g,t+1})$,....,$\log(\sigma^2_{\varepsilon^g,t+h})$.
3. Draw innovations to τ^g_{t+h} using normal distribution with variance $\sigma^2_{\varepsilon^g,t+h}$. Using these innovations compute τ^g_{t+1},...τ^g_{t+h}.
4. Draw η^g_{t+1},...η^g_{t+h} using normal distribution with corresponding variance $\sigma^2_{\eta^g,t+1}$,...$\sigma^2_{\eta^g,t+h}$.
5. Finally using τ^g_{t+1},...τ^g_{t+h} obtained in step 3 above, and η^g_{t+1},...η^g_{t+h} from step 4 compute the sum of these two series to obtain an estimate of π^g_{t+1},....π^g_{t+h}.

3.3 Forecasting Aggregate Inflation

The forecast of the aggregate inflation (quarterly annualized) at time t for h quarters ahead is simply the composite forecast of the services inflation forecast and the goods inflation forecast (both quarterly annualized) combined using the share weights available as of time t. The weights reflect the relative share of services inflation, and goods inflation in overall headline inflation. Specifically, the weight for services inflation is computed as nominal share of personal consumption expenditures of services over nominal PCE, similarly weight for goods inflation is computed as the nominal share of goods consumption expenditures over nominal PCE. Over our forecast evaluation sample (1994.Q1 to 2014.Q4), the shares have been fairly stable at roughly 65 percent going to services expenditure and the remaining 35 percent to goods expenditures.

The point forecast is computed as follows:

$$\hat{\pi}_{t+h} = w^s_t \hat{\pi}^s_{t+h} + w^g_t \hat{\pi}^g_{t+h} \tag{14}$$

where $\hat{\pi}^s_{t+h}$ is a posterior mean of N draws (i.e. 1000 draws) at forecast horizon t+h corresponding to the bivariate model of services inflation and unemployment, and $\hat{\pi}^g_{t+h}$ is a posterior mean of N draws (i.e. 1000 draws) at forecast horizon t+h corresponding to the univariate unobserved components model for goods inflation.

Density forecast is computed similarly, that is for each draw i:

$$\hat{\pi}_{t+h,i} = w^s_t \hat{\pi}^s_{t+h,i} + w^g_t \hat{\pi}^g_{t+h,i} \tag{15}$$

$$p(\pi_{t+h,t}) = w^s_t g(\pi^s_{t+h,t}) + w^g_t h(\pi^g_{t+h,t}) \quad t = 1993.Q4 \;\; to \;\; 2014.Q3 \tag{16}$$

where $p(\pi_{t+h,t})$ is the h step-ahead predictive density for the aggregate inflation, $g(\pi^s_{t+h,t})$ is the h step-ahead predictive density formed at the forecast origin t for the services inflation, and $h(\pi^g_{t+h,t})$ is the h step-ahead predictive density formed at the forecast origin t for the goods inflation.

We compare the forecasting results of our framework against the following seven benchmark models:

Random Walk model (Atkeson and Ohanian (2001)). According to this model, the forecasts of aggregate inflation for h-quarters into the future is simply the average of the most recent four available quarterly readings.

$$\hat{\pi}_{t+h} = 1/4 \sum_{i=t}^{t-3} \pi_i$$

Univariate AR4 model. An (unrestricted) autoregressive model of aggregate inflation with four lags. The forecasts of aggregate inflation for h-quarters into the future are computed iteratively.

We estimate the model using OLS, and then compute the one step ahead forecast for aggregate inflation as:

$$\hat{\pi}_{t+1} = \hat{\beta}_0 + \hat{\beta}_1 \pi_t + \hat{\beta}_2 \pi_{t-1} + \hat{\beta}_3 \pi_{t-2} + \hat{\beta}_4 \pi_{t-3}$$

Similarly, the remainder of the forecasts for h-1 horizons are generated by recursive substitution as follows:

$$\hat{\pi}_{t+h} = \hat{\beta}_0 + \hat{\beta}_1 \hat{\pi}_{t+h-1} + \hat{\beta}_2 \hat{\pi}_{t+h-2} + \hat{\beta}_3 \hat{\pi}_{t+h-3} + \hat{\beta}_4 \hat{\pi}_{t+h-4}$$

Univariate AR1 inflation in gap model. This model is similar to that used in Faust and Wright (2013). They show that this simple model was the most accurate, only judgment forecasts (e.g. Survey of Professional Forecasts) were able to outperform it. Specifically inflation is modeled in a gap form[25],

$$\pi^{gap}_t = \beta_0 + \beta_1 \pi^{gap}_{t-1} + \epsilon_t, \quad \epsilon_t \ i.i.d \ N(0,1)$$

where $\pi^{gap}_t = \pi_t - \pi^{LR}_t$ (π^{LR}_t refers to the long-run inflation expectations. There are various sources that can be utilized to obtain the expectations series see Clark and Doh(2014). We use the series that comes from Federal Reserve Board of Governors FRB/US econometric model, and is referred as PTR (known as survey-based long-run 5- to 10- year-ahead PCE

[25]We modify the Faust and Wright (2013) 'fixed-ρ' specification by adding the intercept term, and estimated coefficients.

inflation expectations series).[26]

The forecasts of inflation gap for h-quarters into the future are computed iteratively. Then the last available value of the trend (as of time t) is added to the forecast of the gap to compute the implied forecast of the aggregate inflation.

Stock and Watson (2007) univariate unobserved component model with stochastic volatility (UC-SV) model. This univariate UC-SV model has been among the most accurate models over the forecasting period of our focus. The model forecasts for aggregate inflation for h-quarters into the future are simply the model's current estimated trend inflation rate. Specifically it decomposes aggregate inflation into a stochastic trend component and a transitory component, assuming time varying variances of the respective shocks to these two components. The specification (consisting of four equations for estimation and one equation for forecasting) is as follows (retaining the notation in Stock and Watson (2007)):

$$\pi_t = \tau_t + \eta_t, \qquad where\ \eta_t = \sigma_{\eta,t}\zeta_{n,t} \qquad \zeta_{n,t}\ \ is\ \ i.i.d.\ N(0, I_1)$$

$$\tau_t = \tau_{t-1} + \varepsilon_t, \qquad where\ \varepsilon_t = \sigma_{\varepsilon,t}\zeta_{\varepsilon,t} \qquad \zeta_{\varepsilon,t}\ \ is\ \ i.i.d.\ N(0, I_1)$$

$$ln(\sigma^2_{\eta,t}) = ln(\sigma^2_{\eta,t-1}) + \nu_{\eta,t}, \qquad where\ \nu_{\eta,t}\ \ is\ \ i.i.d.\ N(0, \gamma I_1)$$

$$ln(\sigma^2_{\varepsilon,t}) = ln(\sigma^2_{\varepsilon,t-1}) + \nu_{\varepsilon,t}, \qquad where\ \nu_{\varepsilon,t}\ \ is\ \ i.i.d.\ N(0, \gamma I_1)$$

γ is a scalar parameter that helps characterize the smoothness of the stochastic volatility process, and so it can either be estimated or fixed. We follow Stock and Watson (2007) and fix it at 0.2.

Point forecast of the future aggregate inflation for the next h periods is the current estimated trend inflation (i.e. filtered estimate of τ_t)

$$\hat{\pi}_{t+h} = \hat{\tau}_t$$

[26]This series for inflation expectations has also been used in several studies such as Peach, Rich, and Linder (2013), Clark and Doh (2014), and Chan, Clark, and Koop (2015). The FRB/US model splices the following three series (i.e. sources) to construct the PTR: econometric estimates of inflation expectations from Kozicki and Tinsley (2001), 5- to 10-year ahead survey measures compiled by Richard Hoey, and 10-year ahead forecasts from the Survey of Professional Forecasts (SPF). These three sources are used because SPF does not go back to 1960. Beginning 2012:Q1 (when the Federal Reserve announced its inflation goal of 2 percent) we set inflation expectations to that value (SPF long run inflation forecast is also fairly stable at 2 percent since then). That said, as shown by Clark and Doh (2014) using either PTR, or Blue Chip (constructed as in Faust and Wright, 2013) gives very similar results.

Since this model has stochastic volatility we use it as a benchmark for density forecast comparison. Research on density forecasting (see Clark and Doh 2014, Chan, Clark, and Koop, 2015) has shown this is a tough benchmark just like the case for point forecast. The density forecasts are computed as discussed in the section 3.2 Modeling Goods Inflation.

Stella and Stock (2015) Phillips Curve (PC). As discussed earlier, our services inflation Phillips curve specification closely follows Stella and Stock (2015). They jointly model aggregate inflation and overall unemployment and find empirical support of the existence of the Phillips curve. We use their model to produce point and density forecasts of aggregate inflation.

Three variable BVAR. Given quarterly data on services inflation, goods inflation, and unemployment rate, one could simply estimate a Bayesian Vector Auto-Regression of these three variables. The quarterly forecasts of the services and goods inflation from the BVAR can be combined using the expenditure weights to form a composite forecast of aggregate inflation. We estimate a small BVAR in growth rates (four lags) estimated with the Minnesota and Sum of coefficient (SOC) priors as one of the benchmark models. We set to one the values of the both the hyper parameters that control for the tightness for Minnesota and SOC priors. Results in which values for the hyper parameters at each forecast origin are determined by maximizing the marginal likelihood leads to substantially less accurate forecasts of inflation medium to long-run.

Let us denote vector $Y_t = [\pi_t^g, \pi_t^s, UR_t]$, then the point forecasts are computed as follows

$$\begin{aligned}\hat{Y}_{t+h} &= B_0 + B_1\hat{Y}_{t+h-1} + B_2\hat{Y}_{t+h-2} + B_3\hat{Y}_{t+h-3} + B_4\hat{Y}_{t+h-4}\\ \hat{\pi}_{t+h} &= w_t^s\hat{\pi}_{t+h}^s + w_t^g\hat{\pi}_{t+h}^g\end{aligned}$$

Steady-State Stochastic Volatility BVAR (SS-SV-BVAR).[27]. Clark 2011 shows that this model generates superior point and density forecasts for inflation and other variables, so we use it as one of the benchmark forecasting models. As the name suggests, the model is a Bayesian VAR with steady state priors on the variables of the system augmented with stochastic volatility. For our exercise we use four variables: unemployment rate, federal funds rate, headline PCE inflation, and core PCE inflation.[28]. Headline PCE inflation is modeled as deviation from core PCE inflation. Following the setup in Clark (2011) and using mostly the same notation, we can write steady state

[27]We thank Todd Clark for providing his RATS code for this particular model. We translated it line-by-line to Matlab

[28]We choose these variables as they are used in the Stochastic Volatility Time Varying BVAR (TVP-BVAR) of Agostino, Gambetti, and Giannone (2013), and we examined this model as well. From the general perspective of accurate forecasts, both SS-SV-BVAR and TVP-BVAR were effectively similar, although in an absolute sense, the SS-SV-BVAR model slightly outperformed TVP-BVAR. For reporting simplicity, we report only results for SS-SV-BVAR in the paper.

BVAR with k lags as:

$$
\begin{aligned}
B(L)(y_t - \Psi d_t) &= v_t \\
v_t &= A^{-1}\Lambda_t^{0.5}\epsilon_t,\ \ \epsilon_t\ is\ \ i.i.d\ \ N(0, I_p) \\
\Lambda_t &= diag(\lambda_{1,t}, \lambda_{2,t}, ..., \lambda_{p,t}) \\
log(\lambda_{i,t} &= log(\lambda_{i,t-1} + \nu_{i,t}),\ \ \nu_t\ \ is\ \ i.i.d\ \ N(0, \phi_i)
\end{aligned}
$$

where i=1,...p, and p denotes the size of vector of variables y_t. In our case, p=4 (i.e. number of variables). The deterministic variables are represented by d_t but for our case we have only one which is the constant, $B(L) = I_p - B_1 L - B_2 L^2 - .... - B_k L^k (where\ \ k = 4)$, Ψ is p x 1 vector of coefficients of the deterministic variables (we have one deterministic variable per equation),[29] and A is a lower triangular matrix that has ones on the diagonal.

This model is estimated using Bayesian methods with five-step Metropolis within Gibbs MCMC algorithm. We refer you to Clark (2011) for the detailed algorithm including the prior specification.

To compute the point and density forecasts we simulate and estimate the model using 10,000 draws, discard the first 5,000 and retain every fifth of the remaining 5,000 draws resulting in 1,000 draws. The posterior means of the retained MCMC distributions (i.e. predictive density) are denoted as the point forecasts.

4 Results

4.1 In Sample Results

Figure 1 plot the posterior estimate of the natural rate of unemployment along with the 90 percent probability band from the inflation in parts model. The estimated trend unemployment (i.e. model's implied natural rate) averages around 5.5 percent, and the point estimate does not move above 6 or below 5 through the entire estimation and forecast sample. This suggests that the model attributes most of the time variation in the overall unemployment rate to the cyclical component. The degree of uncertainty around the posterior estimate of the natural rate is quite wide as the 90 percent probability band on average ranges from -1 to +1 percentage points.

Figure 2 plots the posterior estimate of the (common) cyclical component of the unemployment rate along with the 90 percent probability band. The shaded area represents the recession bars (as dated by the National Bureau of Economic Research dating committee). The visual inspection of the figure suggests movements in the cyclical component unemployment rate are in line with the business cycle. That is, cyclical unemployment increases during recessions, and similarly the cyclical component of unemployment falls

[29] If there were more than one deterministic variables per system say for example q of them then Ψ will be p x q matrix of coefficients.

with an expanding economy.

Figure 3 plots the posterior estimate of the trend in services inflation (including the uncertainty around it) alongside the actual services inflation. The trend estimate for services inflation generally moves in tandem with actual services inflation up until about 1990. Since then, trend inflation has ranged between 2 and 4 percent and movements in actual services inflation are largely attributed to cyclical factors. Over the last three years, estimated trend services inflation has been declining and was 2.1 percent at the end of 2014.[30]

Figure 4 plots the posterior estimate of the trend in goods inflation along with the uncertainty estimate. Just like for the case of services inflation, beginning 1980 the trend in goods inflation fell sharply (in line with the actual movements in goods inflation) and it continued to trend lower albeit at a much slower pace until 1998. From thereon it gradually trended up for two years or so and has generally been flat for several years averaging around 1.3 percent until 2011.Q2. Since 2011:Q3 it has been gradually trending lower, estimated to be around -0.7 percent at the end of 2014.

Figure 5 plots the time-varying posterior estimate of the slope of the Phillips curve displaying a negative (inverse) relationship between services inflation gap and the unemployment rate gap. Over our entire sample, the smoothed posterior estimate varies very little between -0.17 and -0.2 percent with 90 percent band ranging between -0.1 and -0.35 percent.[31] To get a sense of how different the estimated slope will be if we stopped the estimation at various points in time (i.e. recursive estimates from 1993:Q4 to 2014:Q4 pretty much coinciding with our forecasting exercise) table 6 shows the comparison between the recursive estimate and the smoothed estimate over the sub-sample. Surely, there is more time variation, for example, during the Great Recession estimated Phillips Curve was more negative (i.e. steepened) but since then has been flattening.

4.2 Pseudo Out-of-Sample Forecasting Results: Point forecasts

In this section, we compare the point forecast evaluation statistics of our inflation in parts framework with the models described above. Relative root mean squared error (relative RMSE) is the accuracy metric for comparing point forecast accuracy for horizons from 1 to 12 quarters ahead, some of which are reported in Table 1. The target variable is the

[30]The overall unemployment rate declined slowly so that the gap between the estimated trend and actual services inflation has been elevated. In the past few quarters the gap has diminished driven mainly by sharp fall in the overall unemployment rate. Overall, the contours of the trend services inflation estimated by the model are quite similar to the estimated trend in aggregate inflation estimated in various other studies such as Cogley and Sargent (2002), Ireland (2007), Lee and Nelson (2007), Cogley and Sbordone (2008), and Kim, Manopimoke, and Nelson (2014).

[31]In the Appendix we also report results for CPI inflation. For the model using CPI inflation, the slope estimate varies more than in the model using PCE inflation (the smoothed estimate for the slope varies between .3 and .4). In addition, the time-varying slope of the Phillips curve between GDP deflator and unemployment rate reported in Stella and Stock (2015) displays time variation quite similar to ours.

quarterly annualized PCE inflation rate. Given limited availability of real time data for services and goods PCE inflation separately, we resort to pseudo out of sample forecast evaluation. Hence, we use the data vintage as of first quarter of 2015. To gauge whether the forecast gains are statistically significant we report the significance statistics from the Diebold-Mariano test (with the Newey-West correction) for equal forecast accuracy between the inflation in parts model and the alternative benchmark models.

We evaluate forecast accuracy using an expanding window of data. That is, we increase the estimation sample by one quarterly observation for each forecast. Specifically, the initial estimation sample runs from 1960:Q1 to 1993:Q4 (generating forecasts from 1994:Q1 to 1996:Q4). We then estimate the model using data from 1960:Q1 to 1994:Q1 (forecasts from 1994:Q2 to 1997:Q1), .., and the final sample runs from 1960:Q1 to 2014:Q3. Accordingly, the forecast evaluation sample spans 1994:Q1 to 2014:Q4, giving us about 84 one-step ahead forecast errors, 83 two-step ahead errors, 82 three-step ahead errors and so on. We denote this as the full-sample, but also report results of the forecast sample that end in 2007:Q3 (denote it as 'pre-crisis' sample), meaning that we use no data beyond 2007:Q3 to evaluate the forecasts. The limited sample results provide a check on the robustness of our results and help evaluate the impact of the financial crisis on the forecast accuracy of inflation from these models.

Table 1 reports forecast evaluation results for the forecast of annualized headline PCE inflation for one, four, five, eight, ten, and twelve steps ahead. The first panel of the table reports results for the full sample (1994:Q1 - 2014:Q4) and the second panel for the pre-crisis sample (i.e. 1994:Q1 to 2007:Q3). The first row reports the root mean squared errors (RMSE) of the inflation in parts model. The remaining rows report relative root mean squared error (relative RMSE) in which each row reports the ratio of RMSE of benchmark model listed in that row (numerator) relative to inflation in parts model. So a **ratio (relative RMSE) of greater than one indicates that inflation in parts model is on average more accurate in forecasting aggregate inflation** than the corresponding benchmark alternative.

On average over the full forecast evaluation sample, the inflation forecasts from the inflation in parts model that exploits the Phillips curve relationship between services inflation and unemployment rate are at least as accurate than any of the alternative benchmark models. The statistically significant forecasting improvement occurs mainly from the 4 quarter or longer forecast horizon.

Among the specific alternative benchmarks, forecasts generated from the model that exploits the Phillips curve relationship between aggregate inflation and unemployment (Stella and Stock 2015) are about 3 to 8 percent less accurate (and in many cases statistically significantly) compared to forecasts from inflation in parts model. The inflation in parts model retains an accuracy improvement over the pre-crisis sample. Over the full sample, forecasts from the the univariate autoregressive process of aggregate inflation (AR4) are on average 5 to 23 percent less accurate than forecasts from the inflation in parts model. Again the forecast improvements of the parts model are statistically significant. The forecasting gains extend to the pre-crisis sample, though they are statistically not significant in most cases.

The inflation forecasts from AR1 gap model over the full-sample are on average marginally less accurate (and about 2 to 15 percent less accurate over the pre-crisis sample) compared to inflation in parts model but are statistically not significant. It is worth noting that once inflation is modeled in gap form (i.e. deviation from its long-run trend) inflation forecasts in the medium to longer horizons are substantially more accurate compared to say AR4 model in which inflation is modeled without accounting for its underlying slowly varying long-run trend (results consistent with Faust and Wright 2013, and Zaman 2013).

Over the full-sample the inflation in parts model outperforms the univariate benchmarks: Atkeson and Ohanian (2001) (RW) and Stock and Watson (2007). We see forecasting accuracy improvements relative to Atkeson and Ohanian (2001) that are statistically significant in many of the horizons. The Stock and Watson (2007) forecasts are statistically not significantly different from those of the inflation in parts model. Over the pre-crisis forecast evaluation sample, the inflation in parts model displays RMSE statistics that are generally lower than both Atkeson and Ohanian (2001) and Stock and Watson (2007), and are statistically significant for the latter case starting at 5 quarter forecast horizon and beyond.

The inflation in parts model outperforms the three variable BVAR both in full-sample and pre-crisis sample. The RMSE from the BVAR over the full sample is on average 20 percent higher and differences are statistically significant; and in the pre-crisis sample the RMSE is 13 percent higher in the short-term increasing to 37 percent further out, and differences are mostly statistically significant.

Compared to our final benchmark model (SS-SV-BVAR), the inflation in parts model marginally out-performs it for most horizons both in the full and pre-crisis sample periods though the gains for the most part are statistically not significant. In the very near-term (h=1) and very long-term (h=12) the SS-SV-BVAR marginally outperform inflation in parts framework. This suggests that the forecasts from this model are on average competitive to those of the inflation in parts model. It is worth noting that by modeling headline inflation as deviation from its underlying trend (in this case core PCE inflation), imposing steady-state priors, and allowing stochastic volatility importantly helps the BVAR to produce accurate inflation forecasts that on average are at par with those of inflation in parts framework.

Table 2 reports forecast evaluation results for the forecast of **average headline PCE inflation** for one to twelve quarters ahead for the full-sample[32]. By construction the one step ahead forecast for average inflation will coincide with the forecast of quarterly annualized inflation, and so the RMSE and relative RMSE for one-step ahead are going to be

[32] For the sake of brevity we report results for full-sample only.

exactly the same as reported in Table 1. The two step ahead average forecast of inflation is the arithmetic average of the annualized quarterly inflation one step ahead and two step ahead. Similarly, twelve step ahead forecast of average inflation is the arithmetic average of one step to twelve steps ahead annualized quarterly inflation forecast. Consistent with results reported in Table 1, both qualitatively and statistically the results corresponding to forecast of average inflation are very similar. Quantitatively, the forecasting gains of inflation in parts framework over the benchmark models are higher with the magnitude of superiority increasing with the horizon. For example, forecasts of average inflation from AO benchmark are on average 16 percent (h=4) to 27 percent (h=12) less accurate compared to inflation in parts model. Similarly, forecasts from the SS-SV-BVAR starting from h=4 and beyond are on average 10 percent less accurate than inflation in parts model, though the differences are statistically not significant.

4.3 Pseudo Out-of-Sample Forecasting Results: Density forecasts

Recently, policy makers and central bankers in particular have shown increasing interest in probabilistic information for policy (see Aastveit et al. (2014), Clark (2011)). As such we evaluate the density forecast performance of our model against models viewed as accurate benchmarks in the inflation forecasting literature. Density forecasting performance is typically measured by the log predictive score [the log of the predictive density corresponding to the variable of interest evaluated at the actual realization]. That metric provides the broadest measure of the accuracy and calibration of the entire predictive density (see Geweke and Amisano (2010)). However, it is sensitive to outlier observations. Continuous Rank Probability Score (CRPS) is an alternative metric of density forecast performance and it measures the closeness of the predictive distribution and the actual realization. This metric penalizes big forecast misses, and thereby favors predictive densities that have higher probability near or at the actual outcome (see Ravazzolo and Vahey (2014), Gneiting and Raftery (2007), Panagioteles and Smith (2008), and Groen, Paap, and Ravazzolo(2013)).[33] We assess the accuracy of the density forecasts using the sum of log predictive scores and sum of CRPS computed over our forecast evaluation sample. Higher the log score more accurate the density forecasts, conversely lower the sum of CRPS more accurate the density forecasts.

Tables 3 and 4 report the density forecasting evaluation results based on sum of log scores, and sum of CRPS respectively. To have a fair horse-race, our inflation in parts framework accuracy is compared against three benchmarks that all have stochastic volatility. The three benchmark models are Stella and Stock (2015), Stock and Watson

[33]A detailed rationale for CRPS is provided in Gneiting and Raftery (2007) who show that CRPS favors forecast densities that have high sharpness and smaller distance. They define sharpness of the predictive density as the mass of the density clustered around its center. Higher the fraction of the mass concentrated around its center, higher will be the sharpness. The "distance" is defined as the distance between the center of the predictive density and the actual outcome.

(2007), and SS-SV-BVAR.

The numbers reported in table 3 are **relative log predictive scores**. Specifically, log of the predictive density from the inflation in parts model - log of the predictive density from the benchmark model. So a **positive number** suggests that density forecasts from the inflation in parts model are on average more accurate compared to the benchmark. Comparing the density forecast accuracy of inflation in parts model with the Stella and Stock (2015) clearly inflation in parts model outperforms it for all horizons and by significant magnitude. The superiority and the magnitude of outperformance is confirmed by alternative metric **relative CRPS** reported in table 4. For this metric, (relative) **negative numbers** correspond to better accuracy of the inflation in parts model relative to the benchmark as lower values of CRPS are preferable to higher values. This suggests that combined density forecast from goods inflation model and services inflation model generates superior density forecasts relative to a model that exploits the Phillips curve relationship between aggregate inflation and unemployment rate. Overall, relative to Stella and Stock (2015) the inflation in parts model generates more accurate both point and density forecasts.

Compared to Stock and Watson (2007) the density forecast accuracy of the two models are quite competitive. Based on relative log predictive score, the Stock and Watson (2007) model marginally outperform the inflation in parts framework. On the other hand, if we look at the relative CRPS, with the exception of horizon h=1 and h=8, inflation in parts model marginally outperform Stock and Watson. So overall, it is safe to say that just like in the case of point forecast accuracy, these two models' density forecast accuracy is on average about the same.

Finally, compared to the SS-SV-BVAR, density forecast accuracy of the inflation in parts model is more accurate for all horizons except the first horizon (h=1). This is true whether we look at the relative log score or relative CRPS. It is worth noting that even though point forecast accuracy of SS-SV-BVAR is competitive to the inflation in parts model, the density forecast accuracy is generally inferior to the parts model.

Overall, the inflation in parts model generates density forecasts that are at least as good or better than models that inflation forecasting literature has shown to be among the best.

5 Conclusions

In this paper, we model aggregate inflation by estimating sub-components of inflation (services and goods separately). We employ the unobserved components model with stochastic volatility as in Stella and Stock (2015) to exploit the correlation between unemployment and services inflation as emphasized in Peach, Rich and Linder (2013) and use a univariate unobserved components model with stochastic volatility for the goods component of inflation. We estimate the models from 1960:Q1 to 1993:Q4, forecast from 1 to 12 quarters ahead and iterate this process in an increasing data window until the end of 2014. The combination of point forecasts from the subcomponents produces aggregate inflation forecast that displays smaller RMSE than a set of the leading inflation forecasting (benchmark) models. The superior forecasting performance of the inflation in parts framework extends to density forecasts as well. We note that the contribution to forecast accuracy arising from the empirical Phillips Curve relationship is notable, and our results are robust to using the short-term unemployment rate in the model for services inflation.

References

Aastveit, K, Gerdrup, K.R., and Jore, A.S. (2014) “Nowcasting GDP in Real Time: A Density Combination Approach,“ *Journal of Business and Economic Statistics* Volume 32, Issue 1, pp 48-68.

Aron, J., and Muellbauer, J. (2008) ”New Methods For Forecasting Inflation and Its Subcomponents: Application to the USA,” *University of Oxford Dept. of Economics Discussion Paper Series* Number 406.

Atkeson, Andrew, and Lee E. Ohanian (2001). “Are Phillips Curves Useful for Forecasting Inflation?,” *Federal Reserve Bank of Minneapolis Quarterly Review* 25(1), 2-11.

Ball, Laurence, and Sandeep Mazumder (2014). “A Phillips Curve with Anchored Expectations and Short-Term Unemployment,” *NBER Working Paper* Number 20715.

Benalal, N., Diaz del Hoyo, J. L., Landau, B., Roma, M. and Frauke Skudelny (2004). “To aggregate or not to aggregate? euro area inflation forecasting,“ *ECB Working Paper* 374

Bermingham, C., and Antonello D’Agostino (2014) “Understanding and Forecasting Aggregate And Disaggregate Price Dynamics,“ *Empirical Economics* Volume 46, Issue 2, pp 765-788

Brayton, Flint, Roberts, John M., and John C. Williams (1999). “What Happened to the Phillips Curve?,” *FEDS Working Paper* Number 1999-49

Bruneau, C., De Bandt, O., Flageollet, A., and Emmanuel Michaux (2007) “Forecasting inflation using economic indicators: The case of France,“ *Journal of Forecasting* Vol. 26, pp. 1 22

Chan, Joshua, Clark, Todd E., and Gary Koop (2015) “A New Model of Inflation, Trend Inflation, and Long-Run Inflation Expectations,“ *Cleveland Fed Working Paper,* 15-20

Clark, Todd E. (2004). “An Evaluation of the Decline in Goods Inflation,” *Federal Reserve Bank of Kansas City Economic Review* Second Quarter 2004

Clark, Todd E. (2011). “Real-Time Density Forecasts from Bayesian Vector Autoregressions with Stochastic Volatility,” *Journal of Business and Economic Statistics* 29(3): 32741

Clark, Todd E. (2014). “The Importance of Trend Inflation in the Search for Missing Deflation,” *Federal Reserve Bank of Cleveland Economic Commentary* 2014-16

Clark, Todd E., and Taeyoung Doh (2014). “Evaluating Alternative Models of Trend Inflation,” *International Journal of Forecasting* 30, 426-448

Cogley, Timothy, and Argia M. Sbordone (2008) “Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve,“ *American Economic Review,* 98, 2101-26

Cogley, Timothy, and Thomas J. Sargent (2002) “Evolving Post-World War II U.S. Inflation Dynamics,“ *In NBER Macroeconomics Annual 2001,* edited by Mark Gertler and Kenneth Rogoff, pp. 331-73. Cambridge, MA: The MIT Press

D’Agostino, Antonello, Gambetti, Luca and Domenico Giannone (2013) “Macroeconomic Forecasting and Structural Change,“ *Journal of Applied Econometrics,* 28: 82-101

Diebold, Francis X. and Roberto S. Mariano (1995) “Comparing Predictive Accuracy,“ *Journal of Business and Economic Statistics,* 13(3): 25363.

Duarte, C. and Antonio Rua (2007) “Forecasting inflation through a bottom-up approach: How bottom is bottom?,“ *Economic Modelling* Vol. 24, No. 6, pp. 941-953

Espasa, A., Poncela,P., and Senra,E. (2002) ”Forecasting Monthly US Consumer Price Indexes Through A Disaggregated I(2) Analysis,” *Statistics and Econometric Series 09, Universidad Carlos III, Departamento de Estadstica y Econometra* Working Paper 02-0301.

Faust, Jon, and Jonathan H. Wright (2013) Forecasting Inflation, in *Handbook of Economic Forecasting*, volume 2 (North Holland)

Geweke, John, and Gianni Amisano (2010) “Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns,“ *International Journal of Forecasting* 26, 216-230

Gneiting, Tilmann, and Adrian Raftery (2007) “Strictly Proper Scoring Rules, Prediction and Estimation,“ *Journal of the American Statistical Society* 102(477): 359-378

Gordon, Robert J. (2013). “The Phillips Curve Is Alive and Well: Inflation and the NAIRU during the Slow Recovery,” *NBER Working Paper* No. 19390

Groen, Jan, Paap, Richard, and Francesco Ravazzolo (2013) “Real-Time Inflation Forecasting in a Changing World,“ *Journal of Business and Economic Statistics* 31:1,29-44

Hargreaves, D., Hannah Kite, and Bernard Hodgetts (2006) “Modeling New Zealand inflation in a Phillips Curve“ *Reserve Bank of New Zealand: Bulletin* Vol. 69, No.3

Harvey, Andrew (2006). “Forecasting with Unobserved Components Time Series Models,” *Handbook of Economic Forecasting* Volume 1 2006, 327-412

Harvey, Andrew (2011). "Modeling the Phillips Curve with Unobserved Components," *Applied Financial Economics,* 21, 7-17

Hendry, D. F. and Kirstin Hubrich (2006) "Forecasting aggregates by disaggregates" *European Central Bank Working Paper* No. 589

Hendry, D. F. and Kirstin Hubrich (2011) "Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate," *Journal of Business and Economic Statistics* Volume 29, Issue 2

Hubrich, Kirstin (2005) "Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy¿' *International Journal of Forecasting* Vol. 21, pp. 119-136

Ireland, Peter N. (2007) "Changes in the Federal Reserve's Inflation Target: Causes and Consequences." *Journal of Money, Credit, and Banking* 39, 1851-82

Kim, Chang-Jin, Manopimoke, Pym, and Charles R. Nelson (2014). "Trend Inflation and the Nature of Structural Breaks in the New Keynesian Phillips Curve," *Journal of Money, Credit, and Banking* Volume 46, No. 2-3

King, Robert G., and Mark Watson (1994). "The post-war U.S. Phillips curve: a revisionist econometric history," *Carnegie-Rochester Conference Series on Public Policy* 41 (1994):157-219

Koopman, Siem Jan., and Andrew Harvey (2003). "Computing observation weights for signal extraction and filtering," *Journal of Economic Dynamics and Control* 27 (2003):1317-1333

Kozicki, Sharon, and P.A. Tinsley (2001) "Shifting Endpoints in the Term Structure of Interest Rates," *Journal of Monetary Economics* 47, 613-652

Lee, Jaejoon, and Charles R. Nelson (2007). "Expectation Horizon and the Phillips Curve: The Solution to an Empirical Puzzle," *Journal of Applied Econometrics* 22(1), 161-178.

Luetkepohl, Helmet (2010) "Forecasting Non-linear Aggregates and Aggregates with Time-Varying Weights" *EUI Working Paper* No. 2010/11

Morley, James (1999). "A Note on Constraining AR(2) Parameters in Estimation," Washington University manuscript

Moser, G., Rumler, F., and Johann Scharler (2007) "Forecasting Austrian inflation," *Economic Modelling* Volume 24, Issue 3, 470-480

Nelson, Charles R., and Williams G. Schwert (1977). "Short-Term Interest Rates as Predictors of Inflation: On Testing the Hypothesis that the Real Rate of Interest Is Constant," *American Economic Review* 67, 478-86

Panagiotelis, Anastasios, and Michael Smith (2008). "Bayesian Density Forecasting Intraday Electricity Prices Using Multivariate Skew t Distribution," *International Journal of Forecasting* 24: 710-727

Peach, Richard, Rich, Robert, and Alexis Antoniades (2004). "The Historical and Recent Behavior of Goods and Services Inflation," *Federal Reserve Bank of New York Economic Policy Review* Volume 10, No. 3 (December):18-31

Peach, Richard, Rich, Robert, and M. Henry Linder (2013). "The Parts Are More Than the Whole: Separating Goods and Services to Predict Core Inflation," *Federal Reserve Bank of New York Current Issues in Economics and Finance* Volume 19, Issue Number 7, 2013

Ravazzolo, Francesco, and Shaun Vahey (2014). "Forecast Densities for Economic Aggregates from Disaggregate Ensembles," *Studies in Nonlinear Dynamics and Econometrics* 18(4): 367-381

Reis, Ricardo, and Mark W. Watson (2010). "Relative Goods' Prices, Pure Inflation, and The Phillips Correlation," *American Economic Journal: Macroeconomics 2* 128-157

Stella, Andrea, and James H. Stock (2015). "A State-Dependent Model for Inflation Forecasting," *in Unobserved Components and Time Series Econometrics*, Siem Jan Koopman and Neil Shephard eds., Oxford University Press, pp. 14-29.

Stock, James H., and Mark Watson (2007). "Why Has U.S. Inflation Become Harder to Forecast?," *Journal of Money, Credit, and Banking* 39(1), 3-33

Stock, James H., and Mark Watson (2015). "Core Inflation and Trend Inflation," *NBER Working Paper* 21282

Zaman, Saeed (2013). "Improving Inflation Forecasts in the Medium to Long Term," *Federal Reserve Bank of Cleveland Economic Commentary* 2013-16

6 Tables and Figures

Figure 1: Unemployment Rate and the Estimated Trend

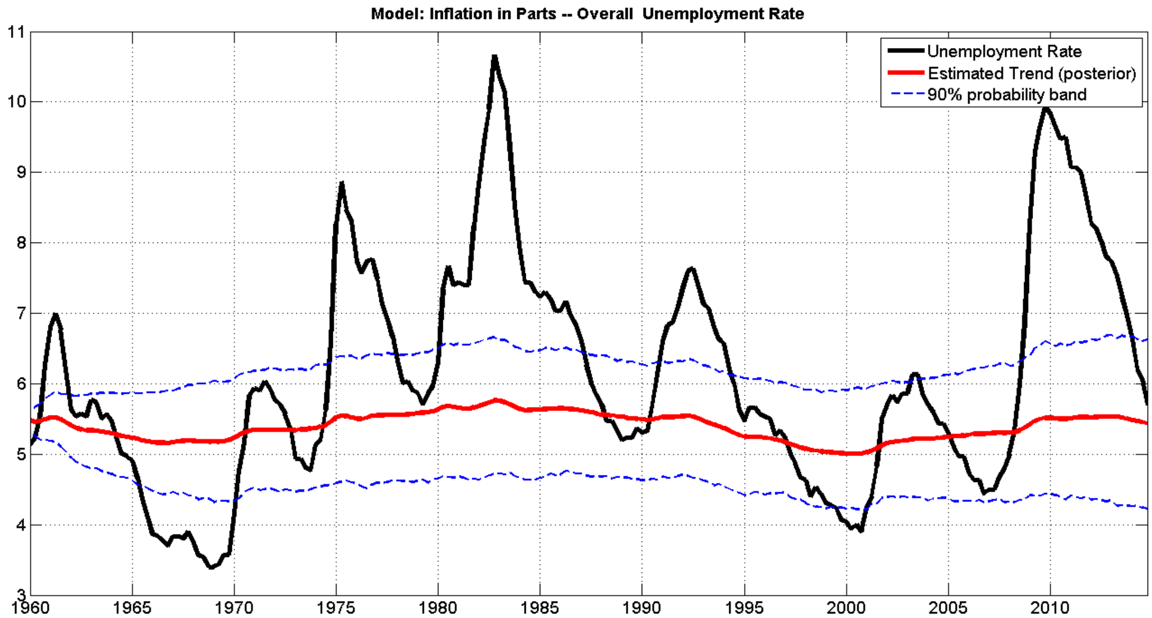
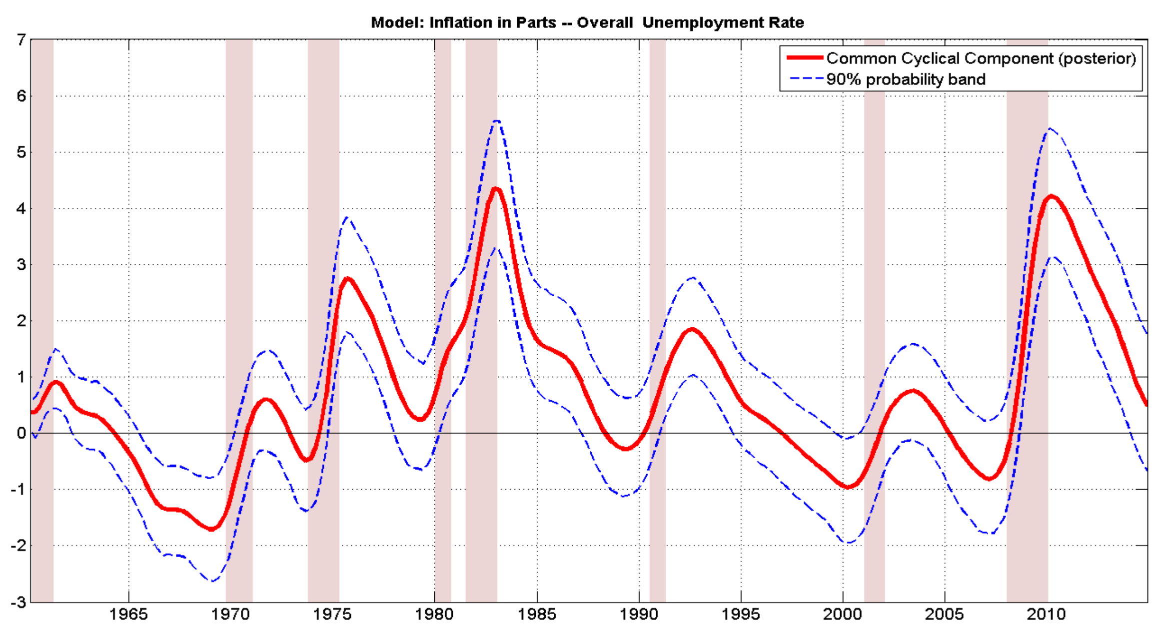


Figure 2: Cyclical UR – common cyclical component



Notes: The estimates above are smoothed, reflecting information based on full-sample from 1960:Q1 through 2014:Q4.

Figure 3: PCE Services Inflation and the Estimated Trend

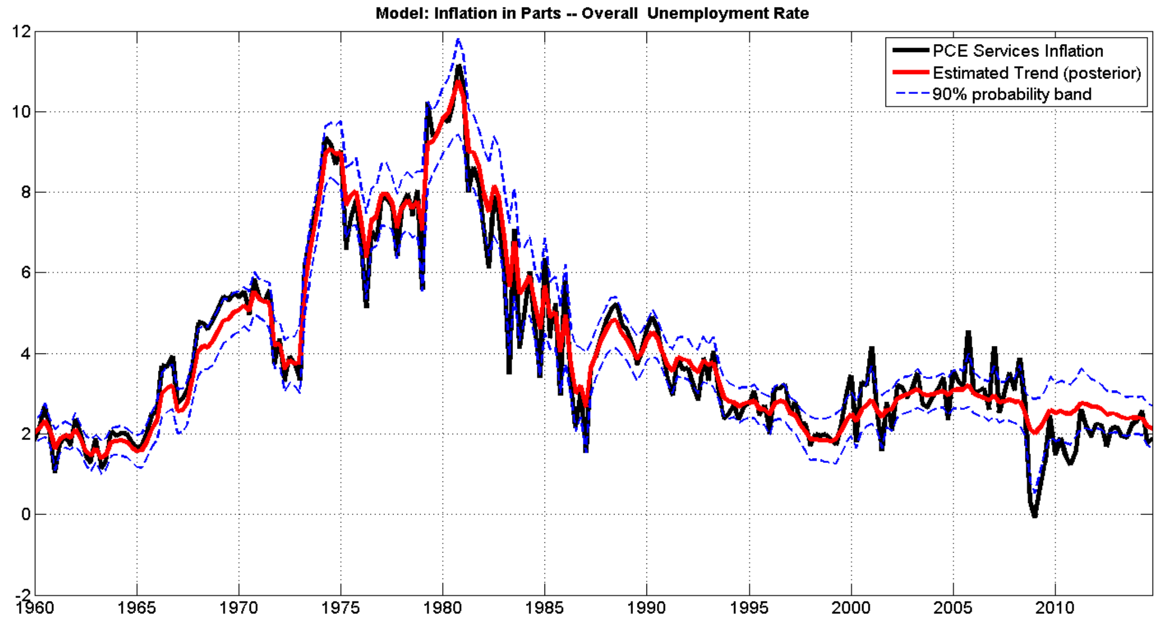
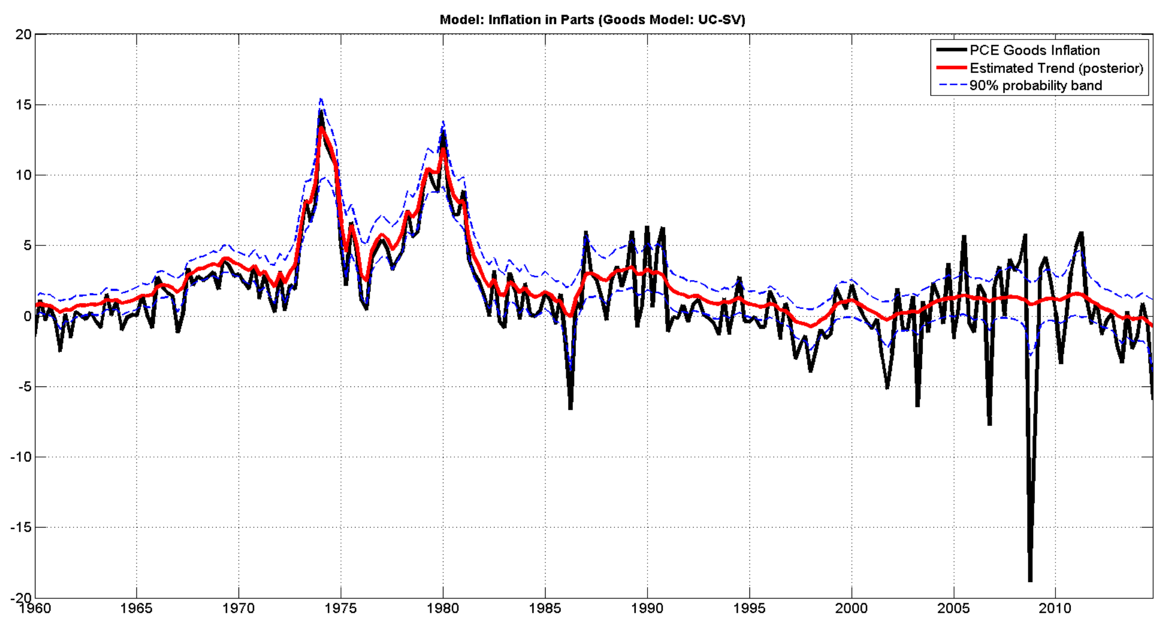
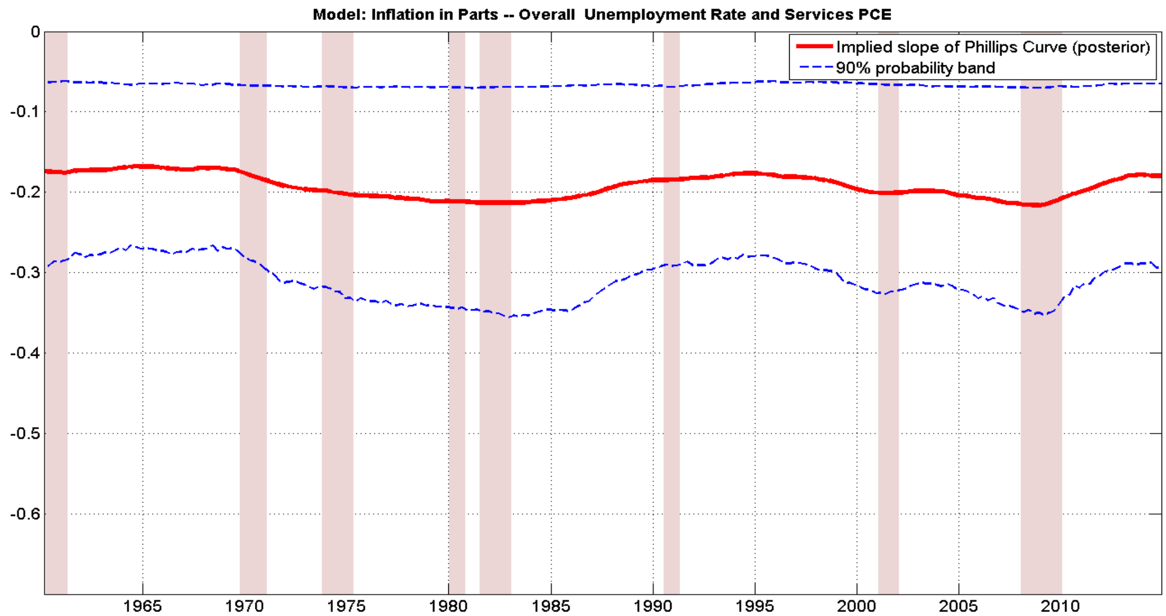


Figure 4: PCE Goods Inflation and the Estimated Trend



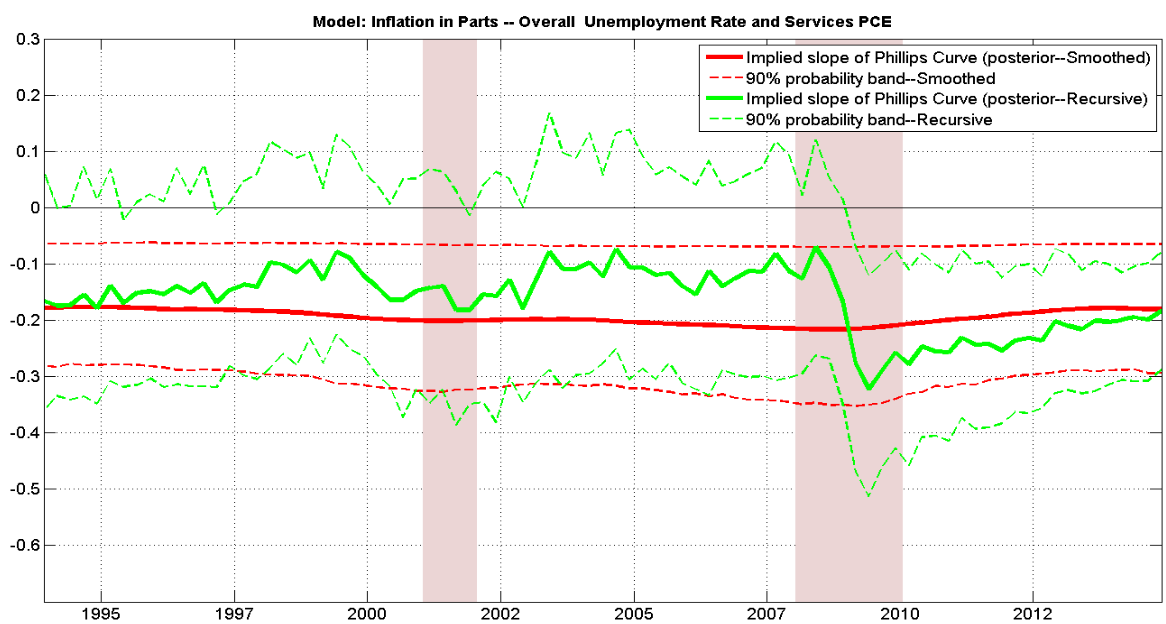
Notes: The estimates above are smoothed, reflecting information based on full-sample from 1960:Q1 through 2014:Q4.

Figure 5: Smoothed Time-varying estimate of implied slope of Phillips curve



Notes: The estimates above are smoothed, reflecting information based on full-sample from 1960:Q1 through 2014:Q4.

Figure 6: Recursive versus Smoothed Time-varying estimate of implied slope of Phillips curve



Notes: While the smoothed estimate is based on full sample information (i.e. 1960:Q1 to 2014:Q4), the recursive estimate is based on the estimating the model at different points in time (in this case recursively beginning from 1993:Q4 to 2014:Q4).

Table 1: PCE Inflation Out-of-sample Forecasting Performance of Inflation in Parts Model

Full Sample (Recursive evaluation: 1994.Q1-2014.Q4)						
Model	h=1Q	h=4Q	h=5Q	h=8Q	h=10Q	h=12Q
Inflation in Parts	1.463	1.576	1.545	1.649	1.630	1.630
Relative RMSE						
Inflation in Parts	1.000	1.000	1.000	1.000	1.000	1.000
Stella and Stock (2015) PC	1.032*	1.053**	1.065***	1.050	1.079***	1.039*
AR4	1.053	1.170*	1.146*	1.167**	1.234**	1.233**
AR1 gap	1.021	1.041	1.020	0.978	1.013	1.025
RW (Atkeson and Ohanian)	1.098**	1.057	1.056**	1.059	1.066	1.046**
Stock and Watson (2007)	1.001	1.006	1.007	0.984	1.007	0.995
Three variable BVAR	1.115	1.226*	1.181**	1.148**	1.220**	1.218**
SS-SV-BVAR	0.972	1.079*	1.038	1.002	1.017	0.963*

Pre-Crisis Sample (Recursive evaluation: 1994.Q1-2007.Q3)						
Model	h=1Q	h=4Q	h=5Q	h=8Q	h=10Q	h=12Q
Inflation in Parts	0.959	1.151	1.137	1.244	1.192	1.159
Relative RMSE						
Inflation in Parts	1.000	1.000	1.000	1.000	1.000	1.000
Stella and Stock (2015) PC	1.058*	1.039	1.075*	1.087**	1.099**	1.096**
AR4	1.090	1.108	1.140	1.213	1.289	1.361
AR1 gap	1.159	1.001	1.039	1.024	1.080	1.153
RW (Atkeson and Ohanian)	1.052*	0.979	1.037	1.029	1.009	1.011
Stock and Watson (2007)	1.013	1.009	1.029	1.046**	1.037***	1.009
Three variable BVAR	1.131	1.142*	1.209*	1.256	1.319	1.373
SS-SV-BVAR	1.003	1.009	1.015	1.021	1.031	0.978

Notes for Table: The first row reports the Root Mean Square Errors from the Inflation in Parts Model. All other rows report the ratios of the Root Mean Square Errors of the various models relative to the Inflation in Parts OU-Spec (i.e. UC model that uses total UR). So a ratio of more than 1, indicates that the Inflation in Parts model does better than the particular model. The forecast performance is based on recursive estimation, i.e. expanding sample. The table reports statistical significance based on Diebold-Mariano test (*10 percent, **5 percent, and ***1 percent significance levels respectively). Stock-Watson (2007) refers to univariate UC-SV model. Atkeson and Ohanian (2001) quarterly random walk model forecasts in our exercise are the quarterly average of the lagged four quarter available PCE inflation annualized rates. Three variable BVAR consists of unemployment rate, services inflation, and goods inflation, and is equipped with the Minnesota and Sum of Coefficient priors. The hyper values of the prior are set to one.

Table 2: PCE Inflation Out-of-sample Forecasting Performance of Inflation in Parts Model

Forecasting Average Inflation

Full Sample (Recursive evaluation: 1994.Q1-2014.Q4)						
Model	h=1Q	h=4Q	h=5Q	h=8Q	h=10Q	h=12Q
Inflation in Parts	1.463	1.033	0.953	0.883	0.848	0.797
Relative RMSE						
Inflation in Parts	1.000	1.000	1.000	1.000	1.000	1.000
Stella and Stock (2015) PC	1.032*	1.086*	1.111**	1.145**	1.160**	1.168***
AR4	1.053	1.305	1.349	1.407*	1.477**	1.553**
AR1 gap	1.021	1.099	1.086	1.009	1.005	1.009
RW (Atkeson and Ohanian)	1.098**	1.159	1.179	1.216	1.240	1.268
Stock and Watson (2007)	1.001	0.999	1.005	0.995	1.002	1.016
Three variable BVAR	1.115	1.405	1.453	1.475*	1.531**	1.602**
SS-SV-BVAR	0.972	1.109	1.117	1.099	1.111	1.101

Notes for Table: The first row reports the Root Mean Square Errors from the Inflation in Parts Model. All other rows report the ratios of the Root Mean Square Errors of the various models relative to the Inflation in Parts OU-Spec (i.e. UC model that uses total UR). So a ratio of more than 1, indicates that the Inflation in Parts model does better than the particular model. The forecast performance is based on recursive estimation, i.e. expanding sample. The table reports statistical significance based on Diebold-Mariano test (*10 percent, **5 percent, and ***1 percent significance levels respectively). Stock-Watson (2007) refers to univariate UC-SV model. Atkeson and Ohanian (2001) quarterly random walk model forecasts in our exercise are the quarterly average of the lagged four quarter available PCE inflation annualized rates. Three variable BVAR consists of unemployment rate, services inflation, and goods inflation, and is equipped with the Minnesota and Sum of Coefficient priors. The hyper values of the prior are set to one. The composite forecast of aggregate inflation is computed at recursive forecast evaluation round by combining the goods and services inflation forecasts using the actual weights as of the forecast origin date.

Table 3: PCE Inflation Out-of-sample Density Forecasting Performance

Relative log predictive score

Full Sample (Recursive evaluation: 1994.Q1-2014.Q4)						
	h=1Q	h=4Q	h=5Q	h=8Q	h=10Q	h=12Q
Relative Log Score: LS of Parts Model - LS of Benchmark Model						
Inflation in Parts	0.000	0.000	0.000	0.000	0.000	0.000
Stella and Stock (2015) PC	7.099	3.340	4.392	1.404	3.582	9.245
Stock and Watson (2007)	4.265	-0.658	-0.155	-1.245	-1.089	4.463
SS-SV-BVAR	-7.880	3.193	2.623	4.186	7.813	11.834

Notes for Table: All the numbers reported in the table are sum of log score from the inflation in parts model minus the sum of log score corresponding to the benchmark listed in the row. So a **positive number** suggests that inflation in parts model is more accurate compared to the particular benchmark model. The forecast performance, i.e. sum of log score is based on recursive estimation, i.e. expanding sample.

Table 4: PCE Inflation Out-of-sample Density Forecasting Performance

Relative CRPS

Full Sample (Recursive evaluation: 1994.Q1-2014.Q4)						
	h=1Q	h=4Q	h=5Q	h=8Q	h=10Q	h=12Q
Relative CRPS: CRPS of Parts Model - CRPS of Benchmark Model						
Inflation in Parts	0.000	0.000	0.000	0.000	0.000	0.000
Stella and Stock (2015) PC	-2.891	-3.422	-4.448	-3.823	-5.238	-4.566
Stock and Watson (2007)	0.742	-0.035	-0.266	1.200	-0.658	-0.072
SS-SV-BVAR	1.272	-5.097	-4.019	-1.977	-4.729	-3.905

Notes for Table: All the numbers reported in the table are sum of CRPS from the inflation in parts model minus the sum of CRPS corresponding to the benchmark listed in the row. So a **negative number** suggests that inflation in parts model is more accurate compared to the particular benchmark model. The forecast performance, i.e. sum of CRPS is based on recursive estimation, i.e. expanding sample.