International Journal of Forecasting 35 (2019) 1770–1789

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

DSGE forecasts of the lost recovery[☆]

Michael Cai[a,*], Marco Del Negro[a,*], Marc P. Giannoni[b], Abhi Gupta[c,*], Pearl Li[d], Erica Moszkowski[e]

[a] Research Department, Federal Reserve Bank of New York, United States
[b] Research Department, Federal Reserve Bank of Dallas, United States
[c] UC Berkeley, United States
[d] Stanford University, United States
[e] Harvard Business School, United States

ARTICLE INFO

Keywords:
DSGE models
Real-time forecasts
Great recession
Financial frictions

ABSTRACT

The years following the Great Recession were challenging for forecasters. Unlike other deep downturns, this recession was not followed by a swift recovery, but instead generated a sizable and persistent output gap that was not accompanied by deflation as a traditional Phillips curve relationship would have predicted. Moreover, the zero lower bound and unconventional monetary policy generated an unprecedented policy environment. We document the actual real-time forecasting performance of the New York Fed dynamic stochastic general equilibrium (DSGE) model during this period and explain the results using the pseudo real-time forecasting performance results from a battery of DSGE models. We find the New York Fed DSGE model's forecasting accuracy to be comparable to that of private forecasters, and notably better for output growth than the median forecasts from the FOMC's Summary of Economic Projections. The model's financial frictions were key in obtaining these results, as they implied a slow recovery following the financial crisis.

1. Introduction

The years since the Great Recession have been quite challenging from a forecasting point of view. Unlike previous post-war recessions, the deep recession was not followed by a swift recovery, but instead generated a persistent output gap. However, this large gap was not associated with negative inflation, as a traditional Phillips curve relationship would have predicted, resulting in what Stock (2011) called the "missing disinflation" (see also Ball & Mazumder, 2011; Coibion & Gorodnichenko, 2015; Del Negro, Giannoni, & Schorfheide, 2015; Hall, 2011). At the same time, the federal funds rate was stuck at near-zero levels for several years. This prompted the central bank to use tools that had never been used before, such as quantitative easing (henceforth, QE) and forward guidance. On top of all this, the U.S. economy found itself in the middle of both a demographic transition caused by the retirement of baby boomers, and a secular downward shift in the growth rate of total factor productivity, at least according to some authors (see, among others, Fernald, 2015; Fernald, Hall, Stock, & Watson, 2017; Gordon, 2015).

This combination of unusual, far-from-steady-state conditions presented a challenging environment for any econometric model, but in particular for dynamic stochastic general equilibrium (DSGE) models in the tradition of Smets and Wouters (2003, 2007), due to their rigid structure and tight cross-equation restrictions. Over the past decade, these models have become part of many central banks' forecasting and policy analysis toolboxes, and the post-Great Recession setting provided an important real-time test of their predictive accuracy. So how did they fare?

☆ Prepared for the November 2017 *Central Bank Forecasting Conference* held at the Federal Reserve Bank of St. Louis.

* Corresponding authors.
E-mail addresses: michael.cai@ny.frb.org (M. Cai), marco.delnegro@ny.frb.org (M. Del Negro), marc.giannoni@dal.frb.org (M.P. Giannoni), abhi.gupta@berkeley.edu (A. Gupta), pearlzli@stanford.edu (P. Li), emoszkowski@hbs.edu (E. Moszkowski).

https://doi.org/10.1016/j.ijforecast.2018.12.001

Against this backdrop, this paper pursues two objectives. The first objective addresses the above question as far as the Federal Reserve Bank of New York's DSGE model (henceforth NY Fed DSGE) is concerned. Specifically, Section 2 of the paper documents how the NY Fed DSGE model fared in terms of real-time forecasting accuracy relative to forecasters such as those surveyed in the Blue Chip survey or the Survey of Professional Forecasters (henceforth SPF), as well as to the Federal Reserve System's Summary of Economic Projections (henceforth SEP), and how researchers using the model coped with the difficulties discussed above. We should stress that the forecasting comparison exercise performed in Section 2 is performed using *real* real-time forecasts—that is, forecasts that were generated at that time.[1] The advantage of this feature of our exercise is that, by construction, there is no look-ahead bias in the choice of either the model or the observables. The disadvantage is that the results are necessarily based only on the available sample of forecasts. Section 2 also discusses how the model changed to incorporate financial frictions and began to use financial data as observables.

The second objective of the paper complements this real real-time forecasting exercise with a *pseudo* real-time analogue. The main goal of this exercise, which is pursued in Section 3, is to determine what model features, and observables, explain the performance of the NY Fed DSGE model. In addition, this exercise extends the historical forecast accuracy comparisons of Del Negro and Schorfheide (2013) and Edge and Gürkaynak (2010) in terms of both the period and the models considered. These earlier comparisons did not focus on the post-Great Recession years, with (Edge & Gürkaynak, 2010) not considering them at all and Del Negro and Schorfheide (2013) barely including them (their sample ends in early 2011). Moreover, Edge and Gürkaynak (2010) only consider the (Smets & Wouters, 2007) model, while (Del Negro & Schorfheide, 2013) focus mainly on the performances of close variants of this model. Here, the centerpiece of our analysis will be models with financial frictions (e.g., Christiano, Motto, & Rostagno, 2014; Del Negro et al., 2015; Del Negro, Hasegawa, & Schorfheide, 2016) that incorporate corporate bond spreads as observables.[2]

We find that in the short and medium run — from one to eight quarters ahead — the NY Fed DSGE model's root mean squared errors (henceforth, RMSEs) are comparable to those of the median forecasts of both the Blue Chip and SPF surveys. Relative to the median of the FOMC's SEP, the NY Fed DSGE model performs much better in terms of the accuracy of output growth forecasts, especially at longer horizons (three years ahead). The NY Fed DSGE model's inflation forecast performs worse than the median SEP up to a two-year horizon, but better at a three-year horizon and beyond. The results of the pseudo real-time forecasting exercise show that financial frictions play a major role, especially in terms of the projections for economic activity, as they imply a slow recovery from financial crisis — a result reminiscent of the findings of Reinhart and Rogoff (2009).

The forecasts in this paper are generated by a microfounded structural model. This implies that they can always be explained in terms of the "impulse and propagation" of structural shocks. Over the course of this paper, we will sometimes take advantage of this feature and describe the DSGE forecasts in these terms, using shock-decompositions and impulse response functions. Some readers may find this commingling of story-telling and forecasting confusing, as forecasting papers usually do not concern themselves with explaining the model's forecasts. However, it could be argued that this is a strength of forecasting with DSGE models — the story and the forecast go together, which implies that we can learn which model features may have resulted in inaccurate forecasts. We will elaborate on this further in the remainder of the paper.

2. Real real-time forecasts of the NY Fed DSGE model

This section begins with a brief description of the main features of the NY Fed DSGE model and of their evolution over time. For the sake of brevity, this description acts as a broad-level overview, with all of the technical details relegated to the appendix and to other sources. The section then continues by documenting the model's forecasting accuracy from 2011, which was the first year in which the model was used to produce regular projections.

2.1. A short history of the New York Fed DSGE model

The New York Fed DSGE model came to existence around 2004 as a three-equation New Keynesian model (see Sbordone, Tambalotti, Rao, & Walsh, 2010). At that stage, the model was used for a variety of policy analysis exercises, but not for forecasting. In 2008, that model was replaced by a medium-scale (that is, similar to the model of Smets & Wouters, 2007, in terms of features) New Keynesian DSGE model built along the lines of Del Negro and Schorfheide (2008) and estimated by means of Bayesian methods using five time series: real GDP growth, core PCE inflation, hours, the labor share, and the federal funds rate.[3]

[1] In this sense, the exercise is similar to those conducted in several papers studying either official central bank forecasts or regularly published model-based forecasts, such as those from the FRB/US model of the Federal Reserve's Board of Governors (e.g., Alessi, Ghysels, Onorante, Peach, & Potter, 2014; Groen, Kapetanios, & Price, 2009; Romer & Romer, 2000, 2008; Tetlow & Ironside, 2007). Edge, Kiley, and Laforte (2010) compare the accuracies of real real-time forecasts from the Board of Governors' Greenbook (the staff forecasts) and FRB/US to those of projections from EDO, the DSGE model used at the Board. In their case, however, the DSGE forecasts are constructed in a pseudo real-time environment. The study by Iversen, Laseen, Lundvall, and Söderström (2016) is the closest to this paper, as it performs a truly real real-time exercise when comparing the forecasts of the Riksbank's DSGE model to the judgmental forecasts published by the Riksbank and to those of a Bayesian vector autoregression for the period 2007–2013.

[2] In addition to the articles we have already mentioned, there are several other papers that have assessed pseudo real-time forecasts of DSGE models, some of which are used in central banks. Examples are those by Adolfson, Andersson, Lindé, Villani, and Vredin (2007), Christoffel, Coenen, and Warne (2011), Fawcett, Kórber, Masolo, and Waldron (2015), Kilponen, Orjasniemi, Ripatti, and Verona (2016), Kolasa and Rubaszek (2015), Kolasa, Rubaszek, and Skrzypczyński (2012), Lees, Matheson, and Smith (2011) and Wieland and Wolters (2012). Fair (2018) provides a recent study of the information content of DSGE forecasts, including those presented here.

[3] Del Negro and Schorfheide (2008) and Del Negro et al. (2013) provide detailed descriptions of the model, priors, data, and estimation procedure.

In mid-2010, the model began to be used internally for forecasting the U.S. economy, and from the end of 2010 onward, the model's forecasts have been produced systematically almost every FOMC cycle and incorporated into internal policy documents. At that time (and for the next six years), the zero lower bound on nominal interest rates (henceforth ZLB) was an important constraint on monetary policy. We incorporated this constraint into the DSGE forecasts by augmenting the measurement equation with federal funds rate expectations obtained from financial markets, following the approach described by Del Negro and Schorfheide (2013) and Del Negro, Giannoni, and Patterson (2012). This approach amounted to forcing the model's expectations for the policy instrument to coincide with market expectations. Since the latter of course take the ZLB into account, the DSGE projections perforce did likewise. In order to enhance the model with the ability to accommodate federal funds rate expectations, the policy rule in the model was augmented with anticipated policy shocks, as used by Laseen and Svensson (2011). These policy "news" shocks capture constraints on future policy, whether they are contractionary (i.e., when the anticipated policy rate is higher than predicted by the reaction function) or stimulative (i.e., when the anticipated policy rate is lower than that predicted by the reaction function, as under a "forward guidance" policy).

In 2010, the model was transformed further by the addition of financial frictions, following the work of Christiano, Motto, and Rostagno (2003) and Christiano et al. (2014). In the aftermath of the financial crisis, we felt that this addition was overdue (Section 3.2 makes the case that this was definitely a good idea from a forecasting performance perspective in the years following the crisis). Specifically, the model incorporated a financial accelerator à la Bernanke, Gertler, and Gilchrist (1999), implying that firms' abilities to invest are constrained by their leverage, and more broadly by financial market conditions. In order to capture financial conditions quantitatively, we added the spreads between the yields of Baa corporate bonds and Treasuries to the model's set of observables. In June 2011, the NY Fed DSGE forecasts obtained from the model with financial frictions became part of a memo[4] produced four times a year for the FOMC (Dotsey, Del Negro, Sbordone, & Sill, 2011; see also page 2 of the June 2011 FOMC Minutes[5]).

The model built in 2010, which is described in some detail by Del Negro et al. (2013), continued to be the main workhorse for DSGE projections and policy analysis at the NY Fed until the end of 2014. It was then replaced by another New Keynesian model with financial frictions – referred to henceforth as the SWFF model and used by Del Negro and Schorfheide (2013) and Del Negro et al. (2015). Relative to the financial friction model introduced in 2010, SWFF was closer to the original (Smets & Wouters, 2007) model in terms of the specification of the household's utility function and other modeling details. Importantly, its forecasting accuracy, especially in periods of financial stress such as the financial crisis, was demonstrated by Del Negro and Schorfheide (2013) and Del Negro et al. (2016). In addition, it had the advantage of adding investment and consumption to the set of observables.[6] This addition was the main rationale behind the change.

The SWFF model itself was never actually used in production at the NY Fed. Instead, we adopted a variant of this model, which we will call SWFF^{+}. This was partly because the SWFF model used in academic papers measured inflation using the GDP deflator, whereas the core PCE deflator was a more relevant measure for policy purposes. We therefore added this variable to the set of observables under the assumption that inflation in the model is the common component between these two empirical measures of inflation.[7] Moreover, at the time a debate on a possible secular decline in productivity growth beginning in the early 2000s was raging (e.g., Fernald, 2015; Gordon, 2015). Given the important policy implications of this debate, we also added John Fernald's measure of total factor productivity growth (henceforth, TFP) to the data on which the model was estimated. In order to give the DSGE a chance to capture secular shifts in productivity growth, we modeled TFP as the sum of two components: a trend-stationary one (as per Smets & Wouters, 2007) and a non-stationary component with growth rates that follow an autoregressive process. As the autocorrelation coefficient approaches one, in principle the latter component can capture very persistent shifts in TFP growth. Furthermore, we also added the 10-year Treasury yield to the set of observables in order to capture changes in financial conditions stemming from both quantitative easing operations and forward guidance. Finally, in 2016 we included GDI as an additional measure of output, following the work of Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2016). We refer to this most recent model as Model SWFF^{++}.[8]

Starting in September 2014,[9] the NY Fed DSGE model forecasts have been made public on the Liberty Street Economics blog (http://libertystreeteconomics.newyorkfed.org), initially twice a year but changing to four times a year at the beginning of 2017 (specifically, they were made available in May[10] and December[11] in 2015, in

[4] https://www.federalreserve.gov/monetarypolicy/files/FOMC20110609memo02.pdf.

[5] https://www.federalreserve.gov/monetarypolicy/files/fomcminutes20110622.pdf.

[6] SWFF is estimated on the same observables as the model of Smets and Wouters (2007) (namely the growth rates in GDP, consumption, investment, and wages, all expressed in real terms, the level of hours, GDP deflator inflation, and the federal funds rate), plus spreads and long-run inflation expectations obtained from the SPF. The latter are included because (Del Negro & Schorfheide, 2013) found that they improve the model's accuracy for forecasting inflation even when the prior on the steady-state inflation parameter is relaxed substantially relative to Smets and Wouters' paper.

[7] This choice was inspired by the work of Boivin and Giannoni (2006) and Justiniano, Primiceri, and Tambalotti (2013).

[8] The appendix provides all of the equilibrium conditions, the prior specification, and data definitions for models SWFF, SWFF^{+} and SWFF^{++}. As was mentioned earlier, Del Negro et al. (2013) provide this information for the early financial friction model.

[9] http://libertystreeteconomics.newyorkfed.org/2014/09/the-frbny-dsge-model-forecast.html.

[10] http://libertystreeteconomics.newyorkfed.org/2015/05/the-frbny-dsge-model-forecast-april-2015.html.

[11] http://libertystreeteconomics.newyorkfed.org/2015/12/the-frbny-dsge-model-forecastnovember-2015.html.

May[12] and November[13] in 2016, and in February,[14] May,[15] August[16] and November[17] in 2017). The current model specification is also available online,[18] as is the Matlab code for the early financial friction model and SWFF^{+}, and the Julia code for SWFF^{++}.[19]

2.2. NY Fed DSGE forecasts

This section examines the performances of NY Fed DSGE forecasts of real GDP growth and core PCE inflation, focusing on the forecasts made for each FOMC cycle from 2011Q1 to 2016Q1. We start by considering the RMSEs of the DSGE model's real output growth and core PCE inflation forecasts relative to the output forecasts of the Blue Chip Economic Indicators (henceforth BCEI) monthly survey and the output and inflation forecasts of the SPF and the FOMC's SEP.[20] We do not show the federal funds rate projections because the NY Fed DSGE forecasts during this period were conditional on external forecasts of this variable in order to take the ZLB and forward guidance into account, as was discussed earlier. Second, we examine the evolution of the NY Fed DSGE model's forecasts for output and inflation and compare the forecasts to contemporaneous SEP forecasts and realized data in order to explain some of the differences in forecast accuracy. The NY Fed DSGE forecasts considered in this comparison cover the period from March 2011 to March 2016.

We compute RMSEs by creating a sample of comparable NY Fed DSGE forecasts for each survey forecast. For a given survey forecast, we search for the nearest *preceding* DSGE model forecast with the same first forecast quarter (in the case of the SEP, we use the NY Fed DSGE forecast produced for the same FOMC meeting). If we cannot find such a forecast, we drop that observation from the sample.[21] This matching scheme ensures that the DSGE forecasts are not given an informational advantage.

The BCEI forecasts are reported as quarter-to-quarter (henceforth Q/Q) percentage changes, and are released monthly. We consider the April, July, October and January forecasts, as these are the last ones that are released prior to the release of the Q1, Q2, Q3, and Q4 GDP measurements. Under our matching scheme, these forecasts are typically paired with the forecasts produced for the March, June, September, and December FOMC meetings, respectively, whenever available (Table A-2 in the online appendix contains the list of all forecast vintages used in the BCEI, SPF, and SEP RMSE comparisons). The Blue Chip survey asks respondents to forecast from the current quarter until the end of the next calendar year, which means that the forecast horizon ranges from nine quarters in January (beginning in Q4 of the previous year) to six quarters in October. We follow the literature and compare the NY Fed DSGE forecast with the average BCEI projection, which is often referred to as the Consensus Blue Chip forecast.

The SPF survey is conducted by the Philadelphia Fed's Real-Time Data Research Center, and is released at the beginning of the second month of each quarter. It is therefore matched with DSGE forecasts from the January, April, July, and October FOMC meetings, whenever possible. Note that this alignment implies that the SPF forecasters have an informational advantage relative to the DSGE, as they have one additional quarter of NIPA data (the preliminary NIPA releases occur at the very end of January, April, July, and October). The SPF forecasts for core PCE inflation and real GDP growth are also reported in Q/Q percentage changes. Their forecast horizon is consistently five quarters. We compare the NY Fed DSGE forecast with the median SPF projection.[22]

Lastly, the SEP is released every other FOMC meeting beginning with the March meeting (the January meeting until 2013). SEP participants project Q4/Q4 (that is, the growth rate over the four quarters of the year being forecast) real GDP growth rates and core PCE inflation rates for the current year and up to three subsequent years. We compare the DSGE forecasts with the median SEP projections.[23] Since DSGE forecasts are also produced in anticipation of each FOMC meeting, the corresponding DSGE forecasts are a natural match for the SEP projections. Note that while both Blue Chip and SPF surveys produce "fixed horizon" projections (that is, they are always released a fixed time before the quarter being forecast), the SEP are "fixed target": in each year, there are four SEP releases that share the same first forecast year, but are made using different information sets.

The three sets of RMSE comparisons shown in Fig. 1 illustrate that the NY Fed DSGE projections are broadly competitive with the survey forecasts over the 2011–2016 period in terms of accuracy. The left panel of Fig. 1 shows that the NY Fed DSGE and BCEI RMSEs for output growth are

12 http://libertystreeteconomics.newyorkfed.org/2016/05/the-frbny-dsge-model-forecastmay-2016.html.

13 https://www.newyorkfed.org/medialibrary/media/research/blog/2017/LSE_dsge-forecast-appendix_Aug-2017.pdf.

14 https://libertystreeteconomics.newyorkfed.org/2016/11/the-frbny-dsge-model-forecastnovember-2016.html.

15 http://libertystreeteconomics.newyorkfed.org/2017/05/the-new-york-fed-dsge-model-forecast-may-2017.html.

16 http://libertystreeteconomics.newyorkfed.org/2017/09/the-new-york-fed-dsge-model-forecast-august-2017.html.

17 http://libertystreeteconomics.newyorkfed.org/2017/11/the-new-york-fed-dsge-model-forecast-november-2017.html.

18 https://github.com/FRBNY-DSGE/DSGE.jl/blob/master/docs/DSGE_Model_Documentation_1002.pdf.

19 The code for the three models is available at https://github.com/FRBNY-DSGE in the `DSGE-2014-Sep` (https://github.com/FRBNY-DSGE/DSGE-2014-Sep), `DSGE-2015-Apr` (https://github.com/FRBNY-DSGE/DSGE-2015-Apr), and `DSGE.jl` (https://github.com/FRBNY-DSGE/DSGE.jl) repositories, respectively.

20 We cannot compare the historic DSGE inflation forecasts to the BCEI forecasts because the latter reports GDP deflator inflation instead of core PCE inflation.

21 Although historically we ran DSGE forecasts at least once or twice each quarter, the times when they were run within the quarter were not always consistent. For this reason, there is not always a suitable DSGE forecast preceding a survey forecast.

22 The Philadelphia Fed's website (https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/2018/survq118) uses the median as the headline number rather than the mean.

23 When the median is not available, we use the average of the upper and lower limits of the SEP central tendency, a range that excludes the three highest and three lowest forecasts of each variable in each year.

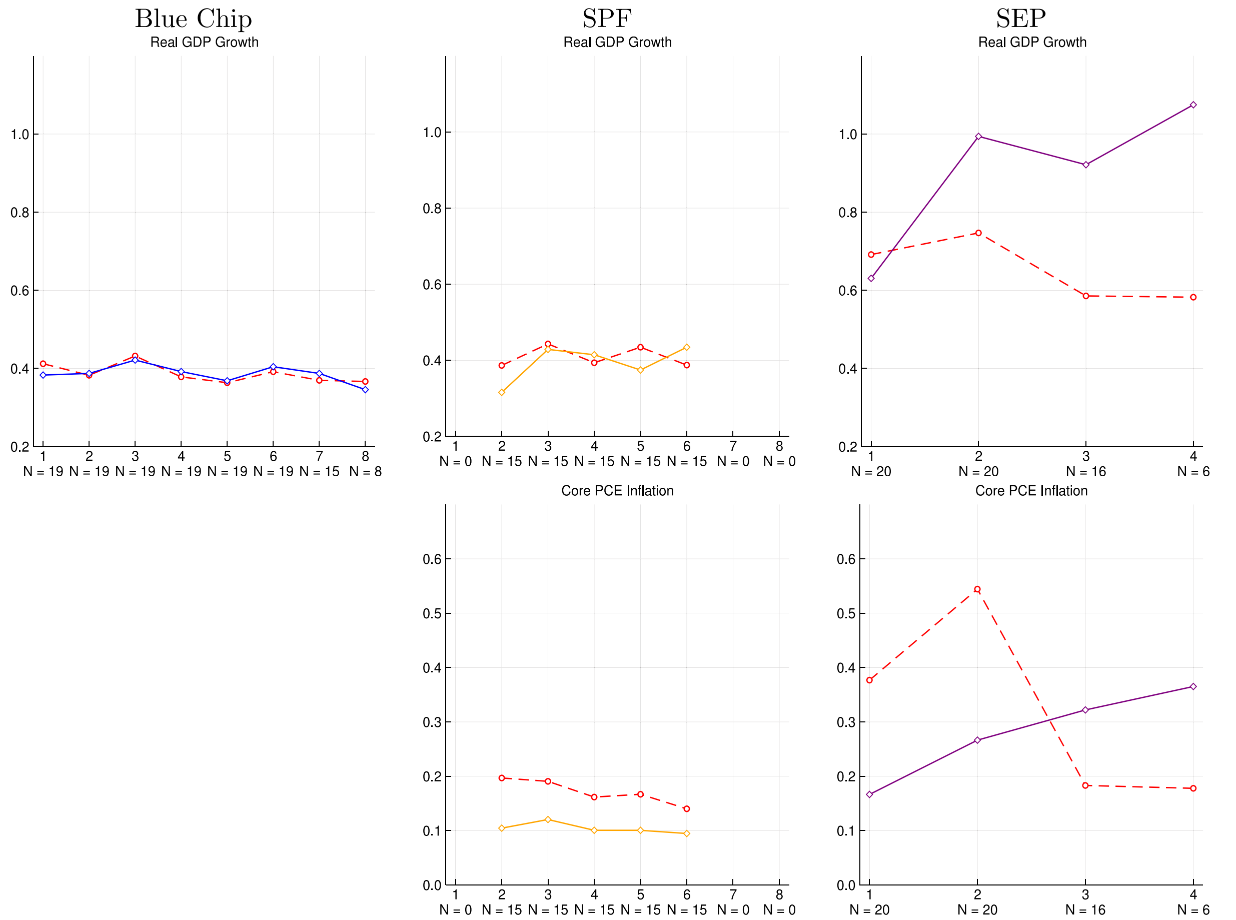

Fig. 1. Historic RMSEs for NY Fed DSGE model forecasts. *Note*: These panels compare the RMSEs for NY Fed DSGE model forecasts (red circles) of real GDP growth and core PCE inflation from March 2011 to March 2016 with those of the Blue Chip Economic Indicators survey (blue diamonds, left), the Survey of Professional Forecasters (SPF) (yellow diamonds, center), and the Summary of Economic Projections (SEP) (purple diamonds, right). The Blue Chip and SPF forecasts are given in terms of Q/Q percentage rates and the SEP forecasts are expressed in Q4/Q4 average rates. When computing RMSEs, each external forecast is matched to the nearest preceding DSGE forecast in order to ensure comparability of the results. We indicate the number of observations below each horizon. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

virtually the same throughout the forecast horizon.[24] The DSGE model's forecasts for output growth are also comparable to the SPF forecasts in terms of accuracy (middle panels; note that we show RMSEs from period 2 onward, given that the SPF has a one-quarter informational advantage over the DSGE). The DSGE core PCE inflation forecasts are somewhat worse than the SPF forecasts, confirming Faust and Wright's (2013) finding that private survey forecasts are hard to beat for inflation. However, the results in Section 3.4 indicate that SPF's informational advantage may be playing an important role for inflation forecasts. The NY Fed DSGE model performs notably better than the SEP's output forecasts over horizons from two to four years ahead (note that we have only six four-year-ahead observations), and performs only marginally worse at the first year horizon. In terms of inflation, the median SEP is more accurate than the DSGE for one to two years ahead, but slightly less accurate for three to four years ahead.

We should stress that the comparison here is between the predictions of a single model — the NY Fed model — and those of forecast *combinations* such as the Consensus Blue Chip. It is well known that forecast combinations, or pools, are often more accurate than their individual components (e.g., Timmermann, 2006), so the fact that a single model is performing as well as these pools is worth noting.

Next, Fig. 2 shows the NY Fed DSGE forecasts of four-quarter average real GDP growth and core PCE inflation made in the first quarter of each year from 2011 to 2016, and provides some context for the RMSEs discussed previously. For comparison, we also include both the realized data series as of November 2017 and contemporaneous SEP projections (we show the SEP's "central tendency", which includes all SEP participants except the top and bottom three). Early in 2011, we see that the SEP projected that

24 It may seem surprising that the first quarter ahead forecasts (that is, the nowcasts) from the DSGE model are as accurate as the BCEI's, given the latter's informational advantage. This result is driven by the fact that the NY Fed DSGE model conditions its projections on judgmental nowcasts from the staff in order to improve the short-run accuracy of its forecasts (see Del Negro & Schorfheide, 2013). Section 3.4 elaborates on this issue.

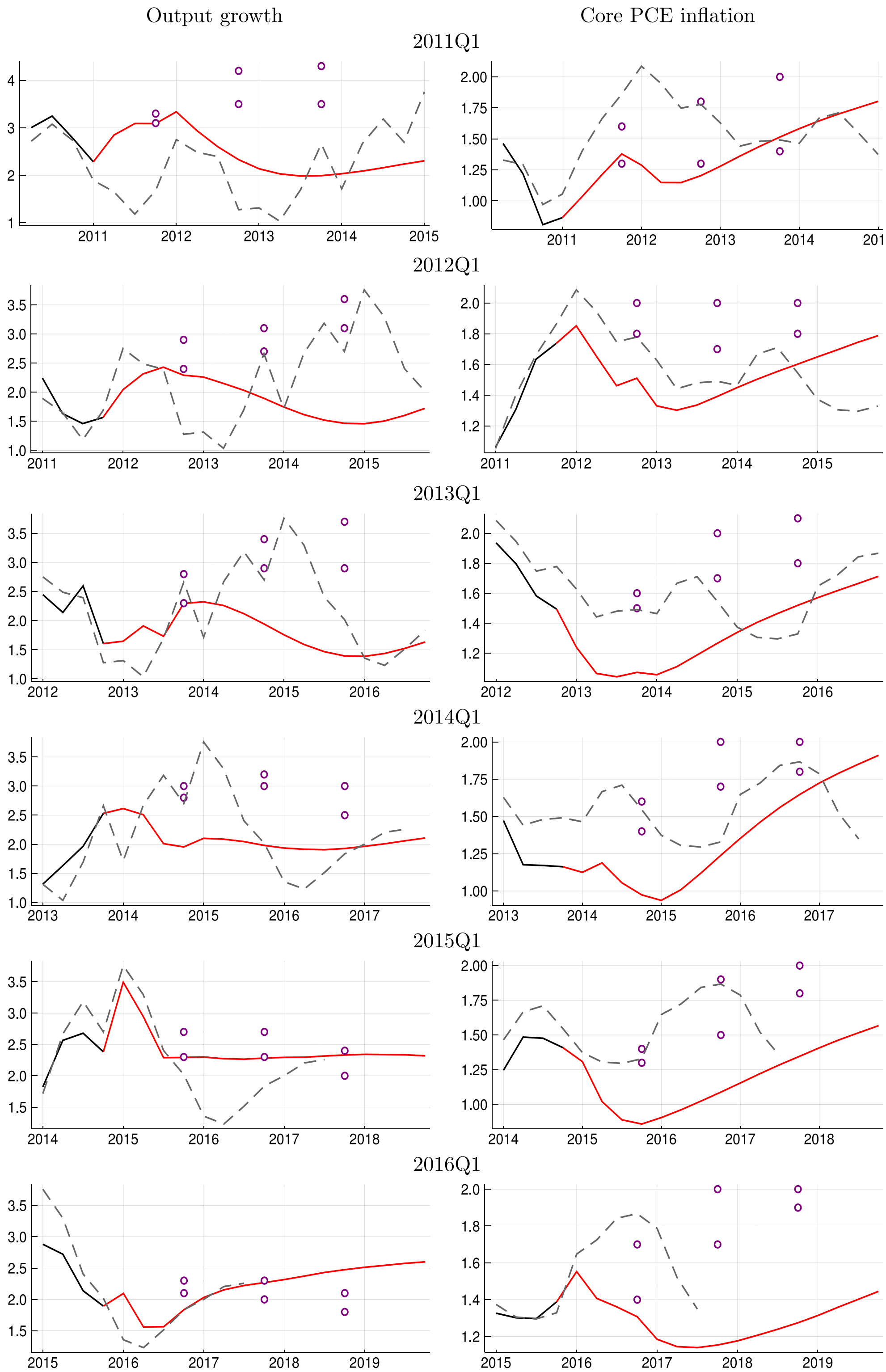

Fig. 2. Evolution of the NY Fed DSGE model forecasts *Note*: The figure shows NY Fed DSGE model forecasts of the four-quarter average real GDP growth (left column, red lines) and core PCE inflation (right column, red lines) produced for the April 2011 and April 2012 FOMC meetings, as well as those from March 2013, March 2014, March 2015, and March 2016. In addition, it also shows the realized data as of the forecast date (solid black lines), the revised series as of November 1, 2017 (dashed black lines), and the upper and lower bounds of the central tendency of the Summary of Economic Projections (SEP) forecasts (purple circles) from the corresponding FOMC meetings. The April SEP projections are still considered “Q1” because the Q1 NIPA data were not yet available at the time when the forecasts were made. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the recovery from the Great Recession would be relatively quick, with growth rates above four percent. The NY Fed DSGE model, on the other hand, projects a very slow recovery from the financial crisis, a finding that echoes the results of Reinhart and Rogoff (2009), although it is obtained in a completely different setting. As we now know, the more pessimistic forecasts of the NY Fed DSGE model were much closer to the realized growth rates through 2013. As is discussed at length in Section 3.2, the model's financial frictions play a key role in these projections. The DSGE model's inflation projections are also very subdued, meaning that they miss the spike in inflation associated with the so-called Arab Spring in late 2011–2012. However, they are quite well in line with the low inflation experienced after 2013.

In the latter half of the sample, that is, from 2014 onward, the DSGE model's forecasts are less accurate over the short run, but still reasonably accurate over the medium and long term. It is worth noting that, by 2015, the SEP and DSGE output growth forecasts have largely aligned. For inflation, the DSGE model's forecasts are often more downbeat than the SEP's, predicting only a gradual return of inflation to the FOMC's long-run goal of 2%. Especially in later years, the DSGE tends to systematically under-predict inflation, while the SEP tends to over-predict it, as it always projects it returning to 2% inflation within a couple of years.

3. Pseudo real-time forecasts

This section uses the results of a *pseudo* real-time forecasting exercise to help us to understand which model features and observables explain the performance of the NY Fed DSGE model. While a *real* real-time environment means that we only have the forecasts from the specific model used at that time, a pseudo real-time setting offers the possibility of running counterfactual experiments, such as: What forecasts would we have obtained if we had stripped financial frictions from the model (Section 3.2)? What would happen if we did not condition the forecast on external expectations for the policy rate (Section 3.3)? What if we did not condition on the nowcast (Section 3.4)? The remaining subsections expand the forecast accuracy comparison in terms of both the models under consideration and the sample size. Section 3.5 compares the accuracy of the DSGE projections to those of simple univariate models and other standard benchmarks. While the forecasts discussed in Section 2 pertain only to the period since 2011, which implies that the evaluation sample is quite short, in a pseudo-real time setting we can investigate the models' performances from 1992 onward (this is the beginning of the sample used by Edge & Gürkaynak, 2010, and Del Negro & Schorfheide, 2013). This is done in Section 3.6. Lastly, we ask whether the addition of model features and data series in the current version of the NY Fed model, SWFF^{++}, helped or hindered the forecasting performance relative to the baseline SWFF model used by Del Negro and Schorfheide (2013), Del Negro et al. (2015), and Del Negro et al. (2016) (Section 3.7). The next section provides some details regarding the construction of the real-time dataset and of the DSGE model forecasts.

Table 1
Data series used in each model.

Variable	SW	SWFF	SWFF^{+}	SWFF^{++}
GDP growth	X	X	X	X
Consumption growth	X	X	X	X
Investment growth	X	X	X	X
Real wage growth	X	X	X	X
Hours worked	X	X	X	X
GDP deflator inflation	X	X	X	X
Federal funds rate	X	X	X	X
10y inflation expectations		X	X	X
Spread		X	X	X
Core PCE inflation			X	X
10y bond yield			X	X
TFP growth			X	X
GDI growth				X

3.1. Real-time dataset and DSGE forecasts setup

The models used in this section are the prototypical (Smets & Wouters, 2007) model (henceforth SW), which does not have financial frictions; the SWFF model; and the two "descendants" of SWFF that were mentioned in Section 2.1, SWFF^{+} and SWFF^{++}. This section starts by discussing the data series used for these models (shown below in Table 1) and the process of constructing a real-time dataset. Next, we discuss the construction of the Blue Chip forecasts dataset, our benchmark for evaluating the accuracy of the DSGE forecasts. Our construction of both the real-time and Blue Chip forecasts datasets follows the approach of Del Negro and Schorfheide (2013, Section 4.1) and Edge and Gürkaynak (2010). Finally, we discuss the DSGE forecast setup.

3.1.1. Data series

Data on nominal GDP (GDP), nominal GDI (GDI), the GDP deflator (GDPDEF), core PCE inflation (JCXFE), nominal personal consumption expenditures (PCEC), and nominal fixed private investment (FPI) are produced at a quarterly frequency by the Bureau of Economic Analysis and included in the National Income and Product Accounts (NIPA). Average weekly hours of production and nonsupervisory employees for total private industries (AWHNONAG), civilian employment (CE16OV), and the civilian non-institutional population (CNP16OV) are produced by the Bureau of Labor Statistics (BLS) at a monthly frequency. The first of these series is obtained from the Establishment Survey, and the remainder from the Household Survey. Both surveys are released in the BLS Employment Situation Summary. Since our models are estimated on quarterly data, we take averages of the monthly data. Compensation per hour for the non-farm business sector (COMPNFB) is obtained from the Labor Productivity and Costs release, and produced by the BLS at a quarterly frequency.

The federal funds rate (the remainder of the paper will sometimes use the acronym FFR) is obtained from the Federal Reserve Board's H.15 release at a business day frequency. Long-run inflation expectations (average CPI inflation over the next 10 years) are available from the SPF from 1991Q4 onward. Prior to 1991Q4, we use the 10-year expectations data from the Blue Chip survey to construct a long time series that begins in 1979Q4.[25] Since the BCEI and the SPF measure inflation expectations in terms of the average CPI inflation whereas we use the GDP deflator and/or core PCE inflation as observables for inflation, we follow (Del Negro & Schorfheide, 2013) and subtract 0.5 from the survey measures, which is roughly the average difference between CPI and GDP deflator inflation across the whole sample. We measure interest-rate spreads as the difference between the annualized Moody's Seasoned Baa Corporate Bond Yield and the 10-Year Treasury Note Yield at constant maturity. Both series are available from the Federal Reserve Board's H.15 release.

Lastly, the TFP growth is measured using John Fernald's TFP growth series, unadjusted for changes in utilization. We use his estimate of $(1-\alpha)$ to convert it into labor-augmenting terms. The details of the data transformations are given in Section A.6 of the appendix.

3.1.2. *Blue chip forecasts*

We compare our pseudo real-time forecasts primarily with contemporaneous ones from the BCEI and the Blue Chip Financial Forecasts (BCFF) survey. The latter contains business economists' projections for financial variables, while the BCEI focuses mainly on macroeconomic variables. This paper is interested in forecasts of real GDP growth and (GDP deflator) inflation from the BCEI and forecasts of the federal funds rate from the BCFF. The RMSE comparisons below compare our DSGE model forecasts to the mean BCEI GDP growth and inflation forecasts and the median BCFF federal funds rate forecast. The BCEI survey is published on the 10th of each month, using data that were available at the beginning of the month, while the BCFF survey is published on the 1st of each month. Though both surveys are released on a monthly basis, we restrict our attention to the January, April, July, and October forecasts. These are the months in which the last forecast for each quarter is made.

For example, the BEA publishes the first estimate of fourth-quarter GDP at the end of January, and the first estimate of first-quarter GDP at the end of April. Hence, the Blue Chip surveys released in February, March, and April contain forecasts in which the first forecasted quarter is Q1. The April Blue Chip survey is the last one to forecast Q1, and choosing it gives the Blue Chip forecasters the greatest informational advantage, as they have access to all of the information released during Q1, and can potentially incorporate higher-frequency financial and other data into their forecasts.

The sample that we consider contains the Blue Chip forecasts from January 1991 to April 2016 (the same sample as in Section 2). Within this sample, we construct real-time datasets using the data vintages available on the 10th of January, April, July, and October of each year. We use the St. Louis Fed's ALFRED database as our primary source of vintaged data. Hourly wage vintages are only available on ALFRED beginning in 1997; for pre-1997 vintages, we use data from Edge and Gürkaynak (2010). The GDP, GDP deflator, PCE, investment, hours, and employment series have vintages available for the entire sample. The earliest available vintages for the core PCE index and GDI are July 29th, 1999, and December 20th, 2012, respectively. Before these dates, we use the earliest available vintage of each series. John Fernald's capital share and TFP growth series are not available on ALFRED. Thus, though there do seem to be revisions, particularly to the TFP growth estimates, we treat these two series as unrevised, using the February 28th, 2017, vintage.[26] The financial variables and the population series are not revised. For each real-time vintage, we use the Hodrick-Prescott filter on the population data observations available as of the forecast date.

When we compare the RMSEs of the DSGE model and Blue Chip forecasts below, we only use as many DSGE forecast horizons as are available in the corresponding Blue Chip release. As was mentioned in Section 2.2, the BCEI respondents submit quarterly forecasts through the end of the next calendar year, so that they forecast nine quarters in January (beginning with Q4 of the previous year) but only six quarters in October. For the majority of our sample (beginning in April 1997), the BCFF respondents submit forecasts for six quarters in the months of January, April, July, and October and for five quarters in all other months.[27] The RMSEs are computed using data downloaded on November 1st, 2017.

3.1.3. *DSGE forecast setup*

Our baseline setup conditions on external interest rate forecasts following Section 5.4 of Del Negro and Schorfheide (2013), because this was the approach taken when generating the NY Fed DSGE model forecasts. We augment the measurement equation to add

$$R^e_{t+k|t} = R_* + \mathbb{E}_t R_{t+k}, \qquad k = 1, \ldots, K,$$

where we use the median k-period-ahead forecast from the BCFF for the observed series $R^e_{t+k|t}$, $\mathbb{E}_t R_{t+k}$ is the model-implied k-period-ahead interest rate expectation, and R_* is the steady-state interest rate. (See Section A in the appendix for additional details.) We provide the model with the ability to accommodate federal funds rate expectations by augmenting the policy rule in the model with anticipated policy shocks, as was discussed in Section 2.1. We take the number of anticipated shocks K to be six, which is the maximum number of BCFF forecast quarters (excluding the observed quarterly average that we impute in the first forecast period).

Specifically, in a given quarter t, the interest rate expectations observables $R^e_{t+1|t}, \ldots, R^e_{t+K|t}$ come from the BCFF

[25] Since the Blue Chip survey reports long-run inflation expectations only twice a year, we treat these expectations in the remaining quarters as missing observations and adjust the measurement equation of the Kalman filter accordingly.

[26] Note that model SWFF does not use core PCE, GDI, or TFP as observables, so the lack of real-time data for these variables is only an issue for $SWFF^+$ and $SWFF^{++}$.

[27] Before April 1997, BCFF provides forecasts for five quarters in January, April, July, and October and for four quarters in all other months. Unlike the macroeconomic variables forecast in the BCEI, which are released with a lag, the quarterly averages for the financial variables in the BCFF are observed immediately at the end of each quarter. To maintain consistency with the output growth and inflation forecasts, we impose that the first forecasted period for the interest rate is the previous quarter, which is forecasted perfectly to be the observed quarterly average. This gives us a total FFR forecast horizon of seven quarters.

Table 2
Summary of $T+1$ conditioning information.

Variable	Source
GDP growth$_{T+1}$	BCEI forecast of $T+1$ GDP growth
GDP deflator inflation$_{T+1}$	BCEI forecast of $T+1$ GDP deflator inflation
Spread	Observed Data
R_{T+1}	Observed Data
$R_{T+2\|T+1}$	$R^e_{T+2\|T+1}$
$\vdots$	$\vdots$
$R_{T+K+1\|T+1}$	$R^e_{T+K+1\|T+1}$

forecast released in the first month of quarter $t+1$.[28] For example, for $t = $ 2008Q4, we use the January 2009 BCFF forecasts of interest rates. We start by using interest rate expectations data beginning in 2008Q4 and continue their use through liftoff, reflecting the post-financial crisis era of central bank forward guidance. Unlike (Del Negro & Schorfheide, 2013), after 2008Q4 we use the expanded dataset containing interest rate forecasts in both estimation and forecasting — again, because this was the approach taken in forecasting with the NY Fed DSGE. However, rather than estimating a separate standard deviation $\sigma_{r^m,k}$ for each of the K anticipated shocks, we impose the restriction $\sigma^2_{r^m,k} = \frac{\sigma^2_{r^m}}{K}$, which implies that the sum of the variances of the anticipated shocks equals the variance of the contemporaneous shock $\sigma^2_{r^m}$. We do this because we have too few observations at the beginning of the ZLB period to be able to estimate these variances independently.[29]

Furthermore, we follow Section 5.3 of Del Negro and Schorfheide (2013) in conditioning on nowcasts — forecasts of the current quarter — of GDP growth, GDP deflator inflation, and financial variables. We accomplish this by appending an additional period of partially observed-data for period $T+1$ (the current quarter, given our timing convention).[30] Specifically, for each real-time forecast vintage, we condition on the corresponding BCEI release's mean forecasts of GDP growth and GDP deflator inflation in period $T+1$. Our choice of forecast origin months means that the entire first forecast quarter has already elapsed by the time the forecast is made, so quarterly averages of financial variables have been observed in their entirety. Finally, we use the BCFF interest rate forecast $R^e_{T+2:T+K+1|T+1}$ as observed expectations of future interest rates in quarter $T+1$. Table 2 summarizes the $T+1$ conditioning information. Note that we do not use any of this $T+1$ information when estimating the model parameters. The models are estimated using only time T information. In fact, we do not reestimate the DSGE model in every quarter in the pseudo real-time forecasting exercise, but only once a year using the January vintage.

3.2. The importance of financial frictions

This section investigates the importance of financial frictions for the forecasting performance of the DSGE models during the recovery. It does so by comparing the forecasting performance of the prototypical SW model with that of SWFF, a version of that model that has been augmented with financial frictions.

The top and bottom panels of Fig. 3 compare the RMSEs of SW (top row, red circles) and SWFF (bottom row, red circles) with those of Blue Chip (blue diamonds) for output growth, inflation, and interest rate projections one to eight quarters ahead, computed from April 2011 to April 2016. For both models, the forecasts are conditional on the BCFF forecasts for the federal funds rate and the BCEI nowcasts for output growth and inflation. (We do this because conditioning on external forecasts for the policy instrument and nowcasts was the standard procedure for the NY Fed DSGE projections during this period, as was discussed above.)

Fig. 3 shows that the accuracies of the SWFF projections for output growth and inflation are comparable to those of the BCEI median forecasts. In fact, the output growth RMSEs for SWFF (lower left panel) are also very similar to those of the NY Fed DSGE model shown in Fig. 1. However, the accuracy of the forecasts from the SW model is considerably worse, especially for output. SWFF differs from SW in both the addition of financial frictions (and spreads as observables) and the use of long run inflation expectations (and a time-varying inflation target). Fig. A-1 in the appendix shows that the key difference between the two models in terms of forecasting performance is the financial frictions: the SW model with long run inflation expectations — called SWπ by Del Negro and Schorfheide (2013) — performs as poorly as SW for output during this period (although it does perform slightly better for inflation, consistent with the findings of Del Negro & Schorfheide, 2013).

Fig. 4 helps us to understand why the SWFF model's forecasts are so much more accurate than the SW's by showing the two models' forecasts computed using the January 2012 vintage. The top and bottom rows show the forecasts for the SW and SWFF models, respectively. Specifically, for output, inflation, and the interest rate, the figure shows the DSGE model forecasts (red solid line); the January 2012 Blue Chip forecasts (blue solid line); the real-time data (black solid); and the revised final data from November 1st, 2017 (gray dashed). Similarly to the SEP forecasts shown in Fig. 2, the SW model forecasts a fast recovery after the Great Recession. Like the NY Fed DSGE model, the SWFF model instead projects a slow recovery, with its forecasts being even more subdued than the BCEI projections. The January 2012 inflation projections from SW are also further off the mark than those from SWFF.[31]

The differences in forecasts between SW and SWFF are not surprising if we consider the different explanations

[28] Since the BCFF survey is released during the first few days of the month, the information set of BCFF forecasters is effectively t; that is, they have no information about quarter $t+1$.

[29] This restriction was also imposed when producing the NY Fed DSGE projections.

[30] Unlike (Del Negro & Schorfheide, 2013), we treat the nowcast for $T+1$ as a perfect signal of y_{T+1}, a specialization of both the *Noise* and *News* assumptions in that paper, in which we set $\eta_{T+1} = 0$. This is also what we do in the production of the NY Fed DSGE forecasts, although we usually rely on the staff's nowcast rather than the BCEI's.

[31] This is explained in part by the fact that the degree of nominal rigidities is lower in SW than in SWFF, as is documented in Table A-1. Hence, inflation depends more on current marginal costs and less on future marginal costs (see the discussion by Del Negro et al., 2015). Since, in terms of levels, the output gap is also still open in 2012 for the SW model, the current marginal costs are still low and the inflation projections are lower.

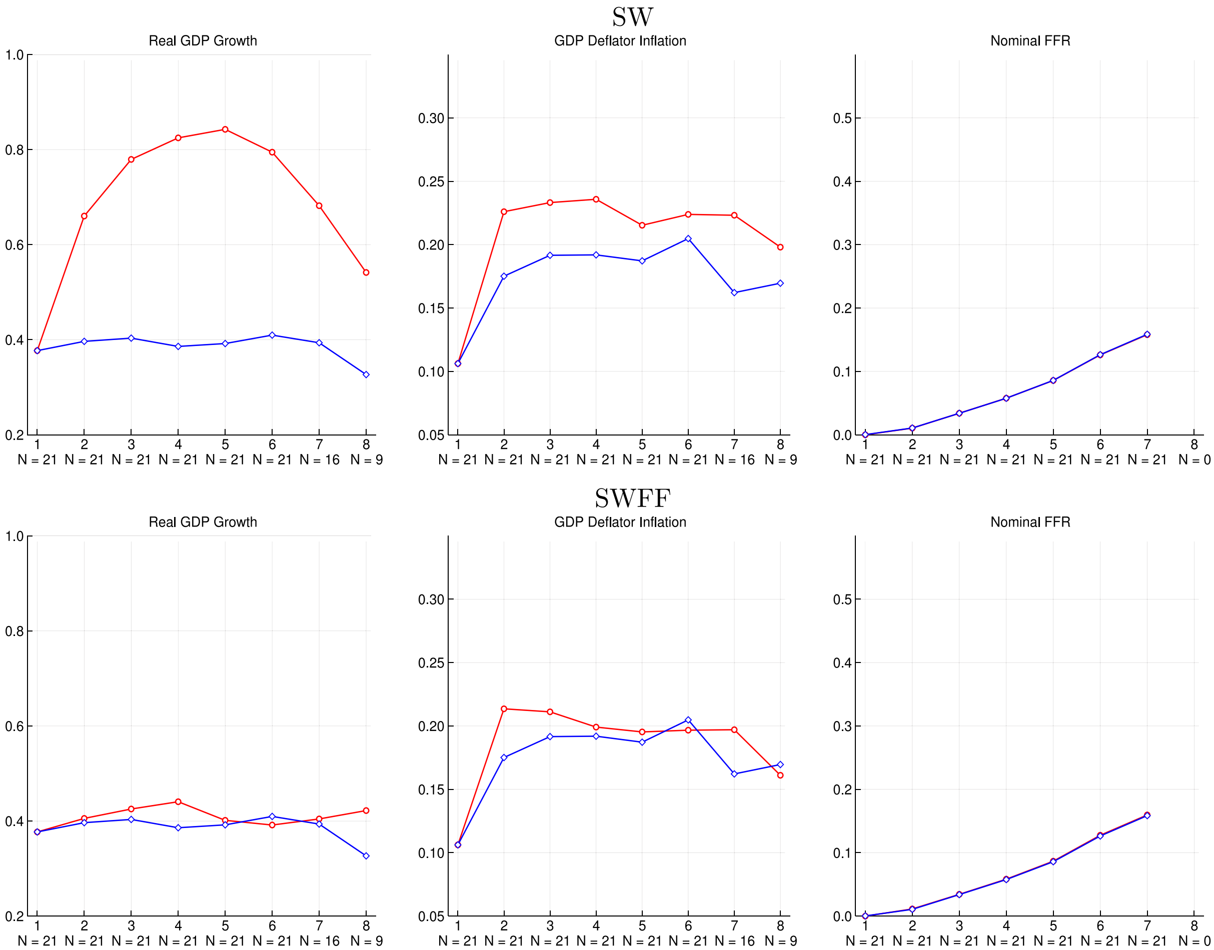

Fig. 3. RMSEs for the SW and SWFF models. *Notes*: The top and bottom panels compare the RMSEs for the SW (top row, red circles) and SWFF (bottom row, red circles) DSGE models with those of the Blue Chip (blue diamonds) for one to eight quarters ahead for output growth, inflation, and interest rates. Output growth and inflation are expressed in Q/Q percentage terms, whereas interest rates are given in quarterly percentage points. The $N = n$ labels under each x-axis tick indicate the number of observations available for both the BCEI and DSGE forecasts at that horizon. The forecasts included in these calculations are from April 2011 to April 2016. The DSGE forecasts are conditional on the BCFF forecasts for the federal funds rate, and the BCEI nowcasts for output growth and inflation. Section 3.2 provides the details of the forecast comparison exercise. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

these two models have for the Great Recession. Fig. 5 decomposes the history of real GDP growth, as of 2012, into the various disturbances that affect the economy in the two models. The SWFF model (right panel) attributes the Great Recession almost exclusively to financial shocks, mostly the so-called "risk premium" shocks. (These are the shocks labeled b in Fig. 5, represented by blue bars.) The impulse responses in Fig. 6 (bottom panel) show that these risk premium shocks have a very persistent effect on the economy: they have a negative effect on *growth rates* for almost 12 quarters, implying that the *level* of GDP begins to recover only after three years.

The SW model also attributes the Great Recession in part to risk premium shocks. (See the left panel of Fig. 5.) However, the role of these shocks is not as important as in SWFF, partly because the SW model does not use spreads as observables. Moreover, because the SW model lacks financial frictions, the impulse responses to these shocks are far less persistent (top panel of Fig. 6), with growth rebounding only a few quarters after the shock. In that model, the Great Recession is driven largely by policy shocks (which capture the ZLB constraint; yellow bars in the left panel of Fig. 5) and by the marginal efficiency of investment shocks (these are the so-called MEI shocks emphasized by Justiniano, Primiceri, & Tambalotti, 2010; they are labeled μ in Fig. 5 and are represented by light blue bars). Fig. 6 shows that both of these shocks have much less persistent effects on GDP growth than the risk premium shocks in SWFF.

In conclusion, the SW model attributes the Great Recession to disturbances of which the effects on the economy are relatively transitory, in contrast to the SWFF model, in which financial shocks have much more persistent effects on output growth. This implies that the SW model expects a faster return of the economy to its steady state, and therefore high growth rates of the economy. In addition, when these high growth rates do not materialize in the aftermath of the recession, the model attributes these forecast misses to *additional* temporary negative shocks that are likewise followed by a quick recovery. As the effects of these shocks

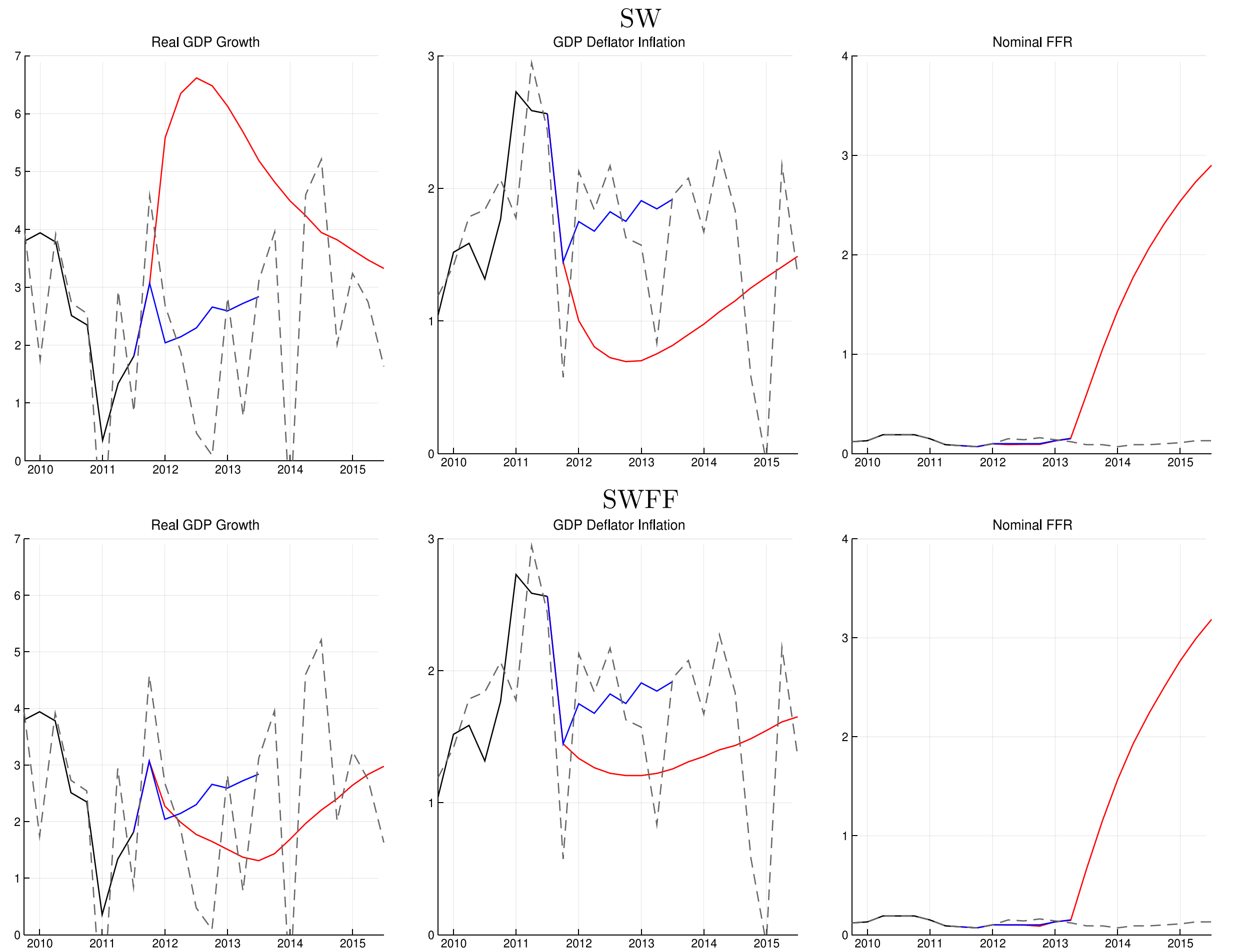

Fig. 4. SW and SWFF forecasts using January 2012 data. *Notes*: For output, inflation, and the interest rate, the panels show the DSGE forecasts (red solid) obtained using data available as of January 2012; the January 2012 Blue Chip forecast (blue solid); real-time data (black solid); and revised final data from November 1st, 2017 (gray dashed). The DSGE forecasts are conditional on the BCFF forecasts for the federal funds rate and the BCEI nowcasts for output growth and inflation. The top and bottom rows show the forecasts for the SW and SWFF model, respectively. Output growth and inflation are expressed in Q/Q percent annualized terms, whereas interest rates are in quarterly annualized percentage points. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

compound, SW ends up predicting very high growth rates for the economy, as Fig. 4 shows.

Does SWFF predict a slow recovery after *every* recession? Fig. 7 reveals that such is not the case. The figure shows the real GDP growth projections using the April 1982 data vintage — that is, at the trough of the 1982 recession.[32] The SWFF model predicts a very fast recovery after the 1982 recession, and its predictions are broadly in line with *ex-post* outcomes. This is the case because the model attributes the recession to disturbances, such as monetary policy shocks, which have effects on the economy that are more transient than those of financial shocks.

3.3. Conditioning on FFR expectations

As was discussed earlier, our baseline analysis conditions on interest rate forecasts from the BCFF in both the estimation and forecasting steps in order to incorporate additional information that is available in the era of central bank forward guidance. This section investigates the impact of that choice. Fig. 8 shows the RMSEs of the SW and SWFF models when we do not use BCFF interest rate forecasts.[33] The sample is the same as in Fig. 3 (April 2011 to April 2016), and we continue to condition on the BCEI nowcasts of output growth and inflation, as well as on the observed quarterly average interest rate in the first period.

The main takeaway of Fig. 8 is that, in the absence of interest rate expectations data, the RMSEs for output

[32] We use the end-of-sample parameter estimates, but otherwise the forecast is out-of-sample.

[33] We obtain the results in Fig. 8 by continuing to use the parameter estimates from the estimation *with* the FFR expectations data. However, Fig. A-2 in the appendix for RMSEs shows that we obtain very similar results when we do not use FFR expectations data at all, including in the estimation. Even when we do not condition on the expected policy path, the projections for the federal funds rate still respect the ZLB as we follow the algorithm described in Section 6.2 of Del Negro and Schorfheide (2013). Specifically, for each path where the ZLB is violated, we use unanticipated policy shocks to bring the federal funds rate up the ZLB.

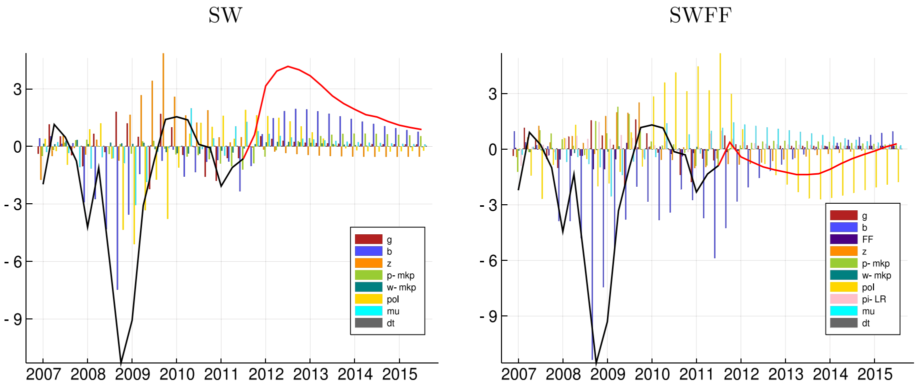

Fig. 5. Shock decompositions of GDP growth. *Notes*: The panels show the SW (left) and SWFF (right) models' shock decompositions of real GDP growth from the January 2012 forecast origin. The solid line (black for realized data, red for the mean forecast) shows the deviation of output growth from its steady state in Q/Q percentage annualized terms. The bars represent the contributions of the various shocks to the deviation from the steady state, computed as the counterfactual values obtained when all other shocks are zero. Some of the shocks have been aggregated in this decomposition. The SWFF shocks are categorized, in order, into aggregate demand, discount factor, financial frictions, productivity, price markup, wage markup, monetary policy, inflation target, and marginal efficiency of investment. The gray bars represent the deterministic trend, the counterfactual values obtained from iterating the initial state vector forward without any shocks. The shock categories for the SW model are a strict subset of the SWFF shock categories. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

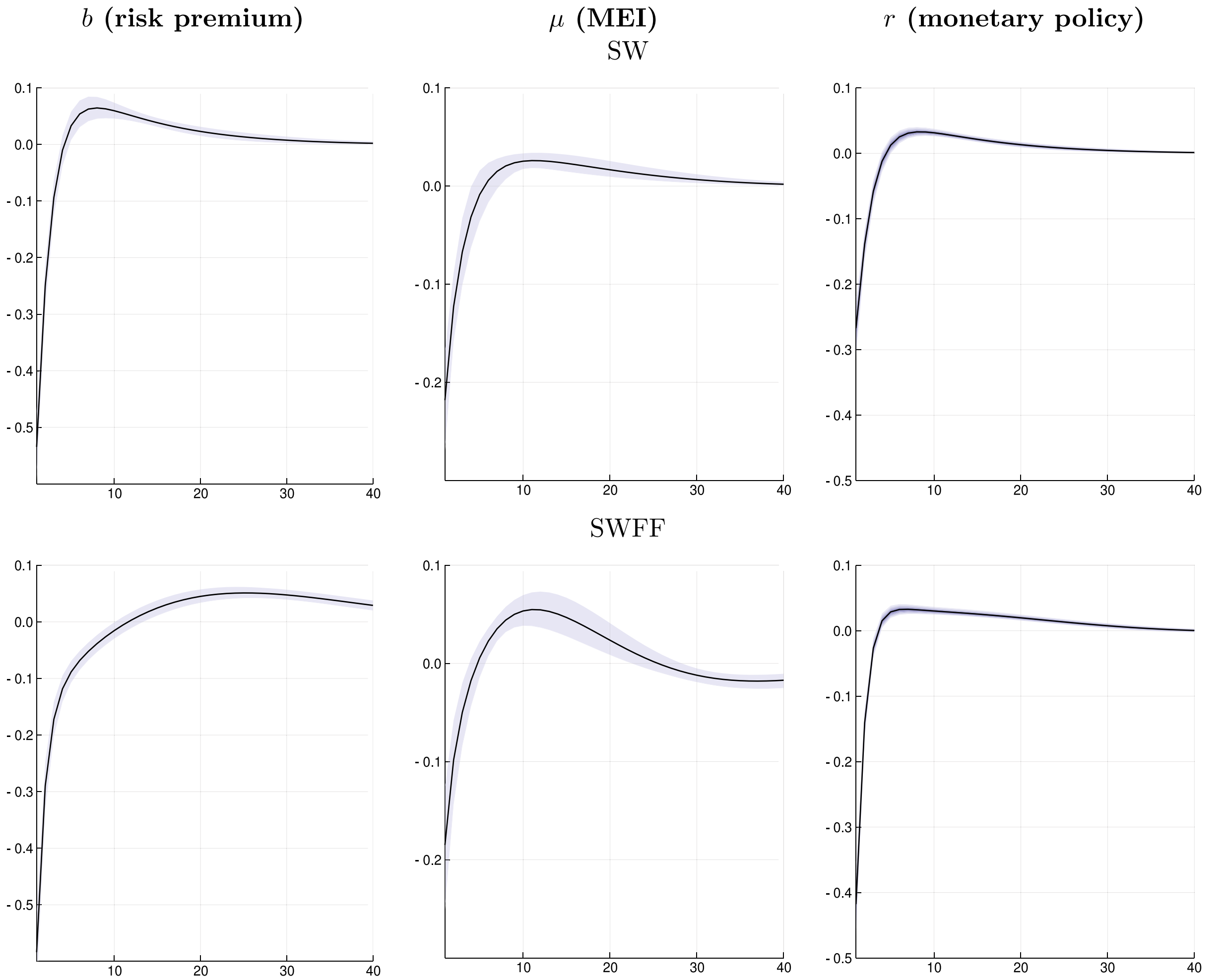

Fig. 6. Impulse responses of real GDP growth. *Notes*: The panels compare the SW (top panels) and SWFF (bottom panels) DSGE models' impulse response functions of real GDP growth to a one-standard-deviation innovation in the discount factor (left), the marginal efficiency of investment (center), and (contemporaneous) monetary policy (right). Parameters estimated using the baseline January 2012 dataset are used. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 7. SWFF forecast of the 1982 recession. *Notes*: For real GDP growth, the figure shows the SWFF forecast beginning in 1982Q1 (red solid); real-time data (black solid); and revised final data from November 1st, 2017 (gray dashed). The forecast was generated using April 1982 data, with the parameters from the January 2016 estimation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

growth and inflation in the SWFF model are very similar to those computed in Fig. 3, though the RMSEs for the federal funds rate deteriorate substantially. Regarding the SW model, the RMSEs for output growth improve somewhat in the absence of interest rate expectations data, but remain sensibly above those of the SWFF model. On the basis of these results, one may conclude that policy transmission is weak in SWFF (the forecasts for the policy rate are very different, but those for output growth and inflation are not), but less weak in SW. This would be the wrong conclusion ((Del Negro et al., 2015) show that the policy transmission in SWFF is quite important). Instead, the explanation for this result can be found in the different ways in which SWFF and SW interpret the conditioning on federal funds rate expectations. The remainder of the section elaborates on this point.

We increase our understanding of the effect of conditioning on FFR expectations on the two models by again focusing on a specific set of forecasts: those computed using the January 2012 vintage. Fig. 9 is analogous to Fig. 4 except that the DSGE projections are computed without using FFR expectations. Clearly, both DSGE models predict an earlier liftoff of the federal funds rate relative to both the BCFF projections and ex-post outcomes. This is not surprising: Blue Chip forecasters are aware of the Federal Reserve's forward guidance, while the DSGE econometrician is not unless we condition on either market or survey expectations (which is why the NY Fed DSGE model conditions on federal funds rate expectations). We also note that SWFF projects a faster liftoff of the policy rate than SW. This is not surprising, given the fact that SW (counterfactually) projects inflation to be lower than SWFF, and that the estimated policy reaction function, which is the basis of the FFR projections for the DSGE models, depends positively on inflation. This observation explains why the RMSEs for the federal funds rate, shown in Fig. 9, are worse for SWFF than for SW.

The differences between the DSGE forecasts for output growth and inflation in Fig. 4 and 9 illustrate the effect of conditioning on FFR expectations. From the perspective of the DSGE econometrician, forward guidance can be interpreted in two different ways, namely as either "Odyssean" or "Delphic" (see Campbell, Evans, Fisher, & Justiniano, 2012). The Odyssean interpretation amounts to anticipated future monetary policy accommodation; that is, the policy "news" shocks discussed in Section 2.1. On the other hand, the Delphic interpretation leads the econometrician to revise her assessment of the state of the economy, which is of course latent in DSGE models. The lower FFR projections are then interpreted as an indication that the state of the economy is worse than had previously been estimated.[34]

Both effects are at play in the DSGE projections. However, the comparison of Fig. 4 and 9 indicates that the Odyssean effect is very strong, particularly for the SW model: the SW projections for output growth in Fig. 9 are still overly optimistic relative to the ex-post outcomes, but much less so than those in Fig. 4. The comparison of Fig. 4 and 9 therefore reveals that the SW model suffers from what Del Negro et al. (2012) called the "forward guidance puzzle": incorporating the accommodation from forward guidance results in overly optimistic projections for the economy. This also explains why the SW RMSEs for real GDP growth shown in Fig. 8 are smaller than those in Fig. 3. For the SWFF model, the differences in both forecasts and RMSEs with and without conditioning on FFR expectations are much more muted than for the SW model. This is partly because SWFF interprets forward guidance as a combination of Odyssean and Delphic signals, which cancel each other out in terms of output growth and inflation projections. In addition, SWFF is affected less by the "forward guidance puzzle" than SW.[35]

3.4. Conditioning on nowcasts

Del Negro and Schorfheide (2013) discuss the challenges that face the DSGE econometrician. One well-understood challenge is model misspecification (see for example Del Negro & Schorfheide, 2004; Del Negro, Schorfheide, Smets, & Wouters, 2007). Another challenge arises from the limitations of the econometrician's information set; that is, the set of observables used in estimating

[34] Some readers may find it confusing that we discuss Delphic forward guidance even though there are no information asymmetries in the model. However, recall that the state of the economy is latent from the perspective of the DSGE econometrician, and therefore there are informational asymmetries from the perspective of the econometrician: she/he does not see the policy shocks (unlike the agents in the DSGE model, who have perfect information on all of the shocks), but needs to make inference on them on the basis of the available information (all the observables, including the expected policy path).

[35] This is because the SWFF model has higher nominal rigidities than the SW model, among other factors (see Del Negro et al., 2015, and the parameter estimates shown in Table A-1 of the appendix). We should note that it is not straightforward either to assess the relative importance of Odyssean and Delphic effects, or to attribute the different responses across models to forward guidance shocks to specific model features. We leave these questions for future research.

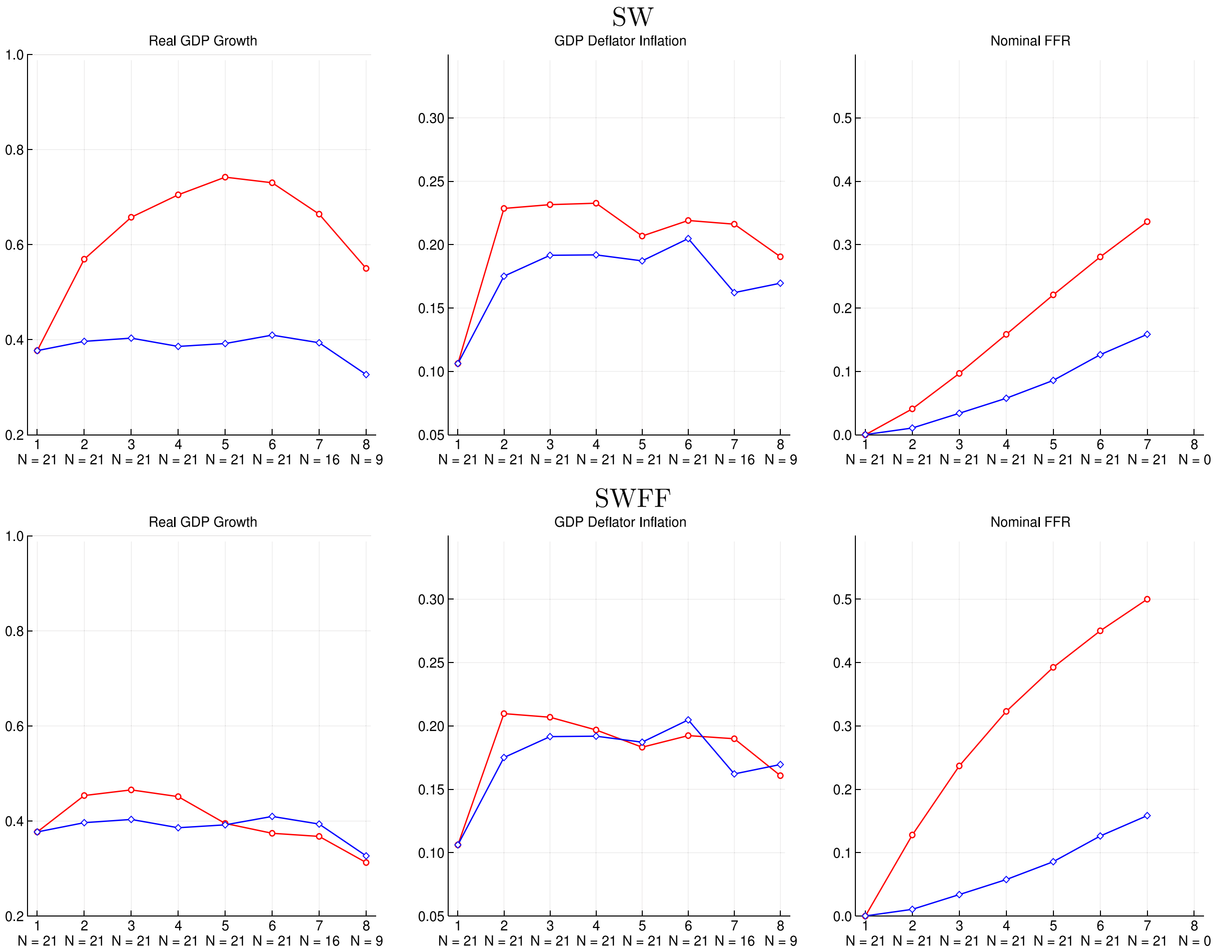

Fig. 8. RMSEs for SW and SWFF vs. Blue Chip, without conditioning on FFR expectations. *Notes*: The top and bottom panels compare the RMSEs of the SW (top row, red circles) and SWFF (bottom row, red circles) DSGE models that do not condition on FFR expectations, with those of the Blue Chip forecasts (blue diamonds) for one to eight quarters ahead for output growth, inflation, and interest rates. Output growth and inflation are expressed in Q/Q percentage terms, whereas interest rates are in quarterly percentage points. The $N = n$ label under each x-axis tick indicates the number of observations available for both the BCEI and DSGE forecasts at that horizon. The forecasts included in these calculations are from April 2011 to April 2016. The DSGE forecasts are conditional on the BCEI nowcasts for output growth and inflation. Section 3.3 provides the details of the forecast comparison exercise. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the model and generating forecasts. Augmenting the set of observables with spreads, for instance, as the SWFF model does, provides valuable information to the econometrician regarding financial conditions. Similarly, conditioning on FFR expectations informs the econometrician about the degree of future policy accommodation. A third challenge is given by the timeliness of the econometrician's information set: the majority of the data series used in the estimation of our model, both "hard" (monthly releases of inflation and consumption) and "soft" (e.g. from surveys, such as the Institute for Supply Management survey, or ISM), become available only at a quarterly frequency, and therefore do not include all of the information that is available at a higher frequency. Blue Chip forecasters use this information to produce nowcasts for output and inflation. For this reason, the DSGE model current-quarter forecasts stand to benefit from conditioning on the nowcasts obtained from the Blue Chip survey. Similarly, the NY Fed forecasts discussed in Section 2 incorporate the nowcasts from in-house forecasters.

How much does incorporating the nowcast improve the DSGE forecasts? Fig. 10 depicts RMSEs for SWFF and the Blue Chip forecasts for output growth, inflation, and the nominal federal funds rate without conditioning on nowcasts. The sample is the same as in Fig. 3 (April 2011 to April 2016), and we continue to condition on the BCFF FFR expectations. Not surprisingly, the Blue Chip nowcasts are much more accurate than the DSGE's for both output growth and inflation. However, for output growth the RMSEs are quite similar to those in Fig. 3 from horizon two onward, while for inflation the improvement associated with including nowcasts persists for about four quarters. We therefore confirm the results of Del Negro and Schorfheide (2013) that the positive effect of conditioning on the nowcast of inflation is much more persistent than the corresponding effect on the GDP, which is not surprising in light of the different levels of persistence in the two series.[36]

36 As was noted in Section 3.1, the nowcast is treated simply as $T + 1$ data, as opposed to a noisy measurement of the forecast variables at time

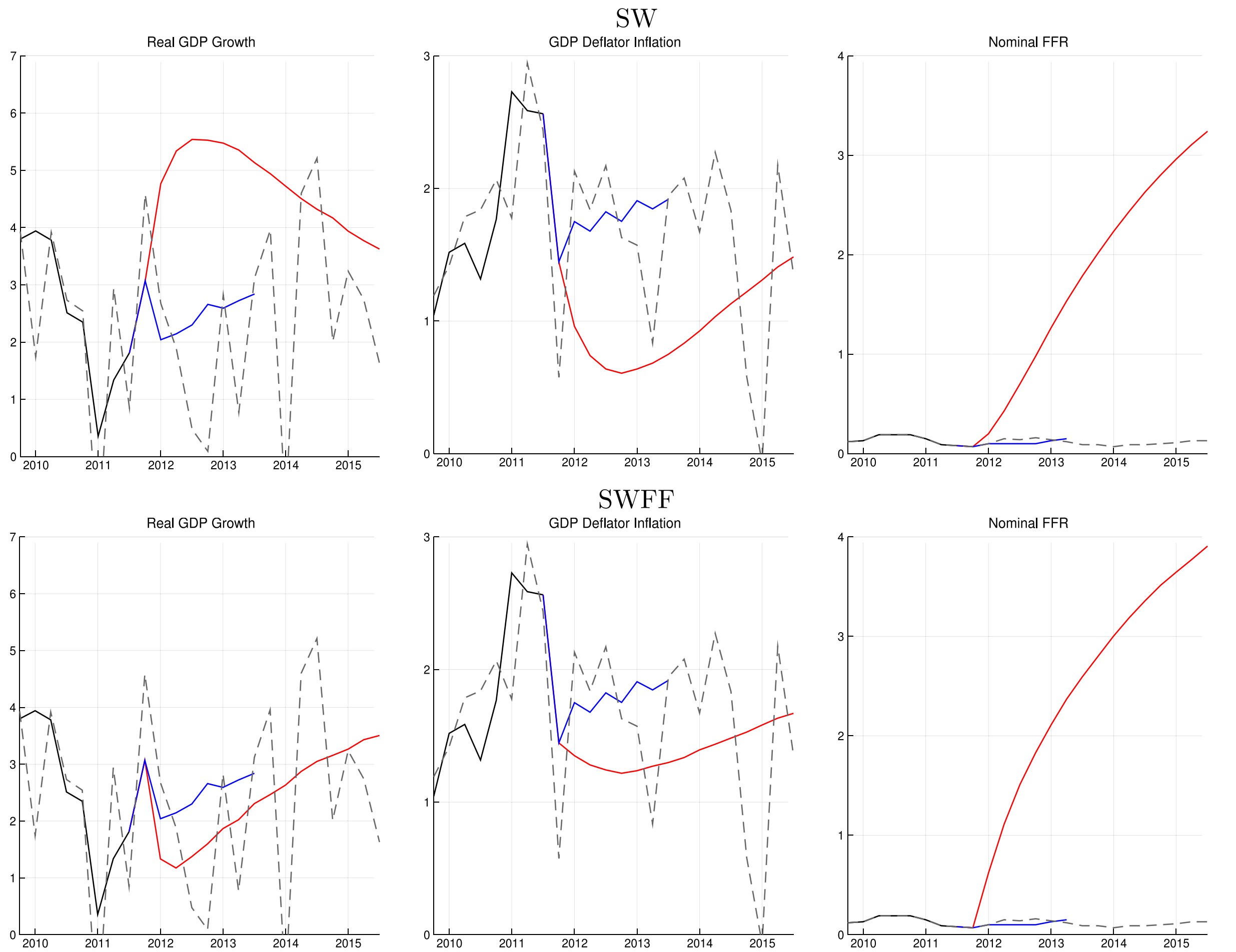

Fig. 9. SW and SWFF forecasts using the January 2012 data, without conditioning on FFR expectations. *Notes*: For output, inflation, and the interest rate, the figure shows the DSGE forecasts obtained using data available as of January 2012 (red solid); the January 2012 Blue Chip forecasts (blue solid line); the real-time data (black solid); and the revised final data as of November 1st, 2017 (gray dashed). The DSGE forecasts are conditional on the BCEI nowcasts for output growth and inflation. The top and bottom rows show the forecasts for the SW and SWFF models, respectively. The output growth and inflation are expressed in Q/Q percentage annualized terms, whereas the interest rates are in quarterly annualized percentage points. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.5. Comparison with naive forecasts/AR models

Edge and Gürkaynak (2010) show that naive predictions obtained using the sample mean for output growth and inflation and the random walk for interest rates perform about as well in their sample as the forecasts from Smets and Wouters' DSGE model. Gürkaynak, Kısacıkoğlu, and Rossi (2013) find that simple models, such as univariate autoregressive (henceforth, AR(p) denotes an autoregressive model with p lags and the constant) or small vector autoregressive models, perform at least as well as Smets and Wouters' model, if not better. In general, the literature has found that either naive or simple AR forecasts are hard to beat for both output (e.g. Chauvet & Potter, 2013) and inflation (e.g. Atkeson & Ohanian, 2001). In light of this, we thought that it would be useful to compare the accuracy of the SWFF forecasts to those of naive and AR(2) forecasts (the results for AR(1) forecasts are nearly identical) for the sample we are interested in. We use the same naive forecasts as (Edge & Gürkaynak, 2010) for output growth and interest rates, but for inflation we use the random walk forecasts based on a four-quarter moving average of past data, which is usually considered as a standard benchmark for this variable in the literature (see Surico, Giannone, & D'Agostino, 2006).[37]

Fig. 11 compares the RMSEs from the SWFF model (the same red circles as in Fig. 3) to those obtained from the

$T+1$ as per (Del Negro & Schorfheide, 2013). We do this because this is the approach taken when producing the NY Fed DSGE forecasts.

37 Edge and Gürkaynak (2010) seem to use the *ex-post* sample mean over the forecast evaluation period as their benchmark, which of course is not available ex-ante. We instead use the sample mean with real-time data for GDP growth, which is what the literature generally uses as a benchmark (again, see Surico et al., 2006). We also considered a random walk forecast based on the last quarterly observation for both output growth and inflation, and, not surprisingly, obtained very poor results which we do not report. Finally, note also that since the SWFF model takes advantage of the nowcast, we let the AR(2) model do that as well and treat it as an observable.

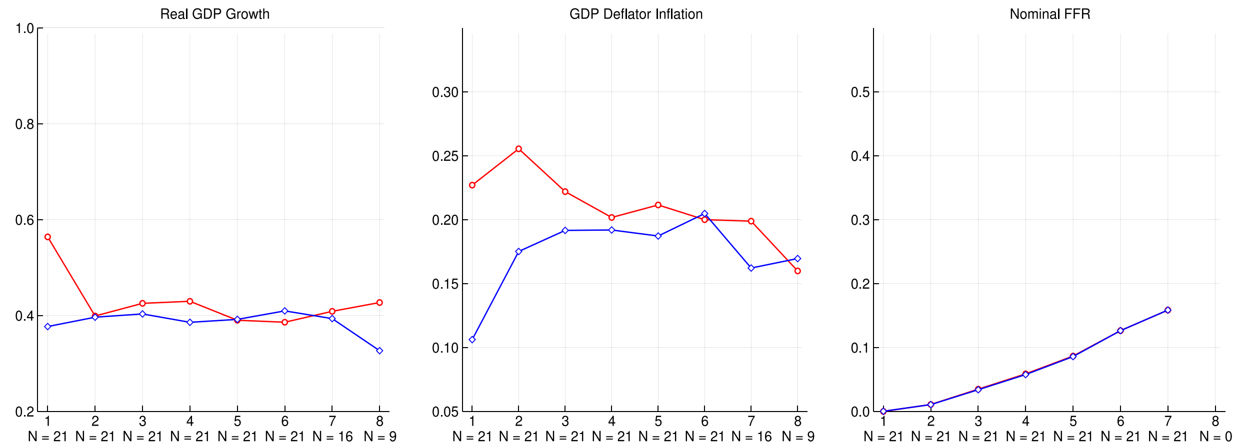

Fig. 10. RMSEs for SWFF vs. Blue Chip, without conditioning on nowcasts. *Notes*: The panels compare the RMSEs of SWFF (red circles) with those of the Blue Chip (blue diamonds) for one to eight quarters ahead for output growth, inflation, and interest rates. Output growth and inflation are expressed in Q/Q percentage terms, whereas interest rates are in quarterly percentage points. The $N = n$ label under each x-axis tick indicates the number of observations available for both the BCEI and DSGE forecasts at that horizon. Only forecast origins from April 2011 to April 2016 are included in these calculations. Section 3.4 provides the details of the forecast comparison exercise. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

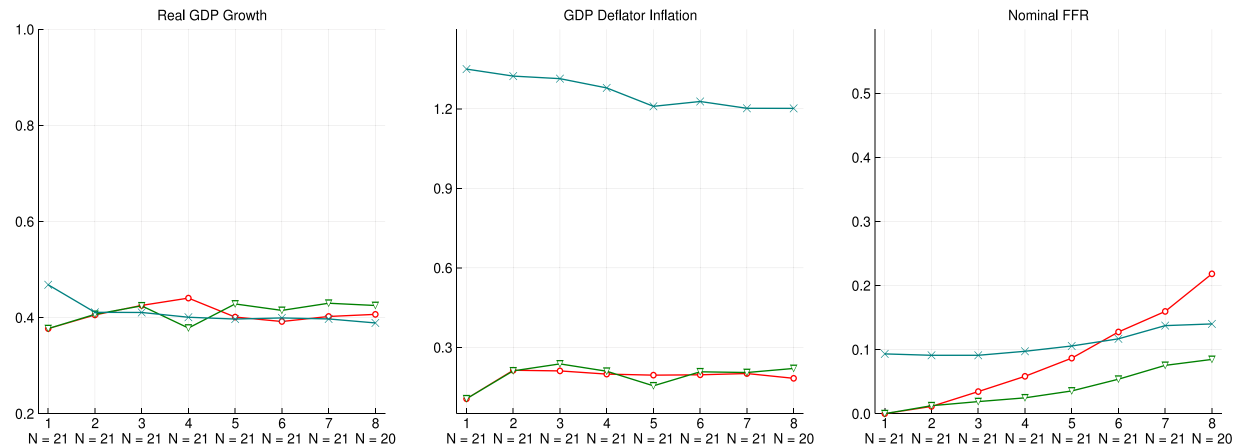

Fig. 11. RMSEs for SWFF, AR(2), and a naive forecast. *Notes*: The figure compares the RMSEs of the SWFF (red circles) DSGE model with those of an AR(2) (green triangles) and a set of naive forecasts (teal crosses) for one to eight quarters ahead for output growth, inflation, and interest rates. The naive forecast for real GDP growth is the sample mean of the data until the first forecast horizon. The naive forecasts for the GDP deflator and the nominal rate are random walks averaged over four quarters. All variables are expressed in Q/Q percentage terms. Only forecast origins from April 2011 to April 2016 are included in these calculations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

AR(2) (green triangles) and naive (teal crosses) forecasts. The accuracy of the AR(2) model is very similar to that of SW for both output and inflation (and more accurate for the interest rate forecasts, but those are really the Blue Chip's forecasts since the DSGE projections are conditional on the expected policy path). The naive forecasts are also as accurate as those of the DSGE for output, but far less accurate for inflation (and somewhat less accurate for the interest rate, at least up to five quarters).

Except for inflation, where Atkeson and Ohanian's (2001) benchmark performs very poorly, these results confirm the findings in the literature.[38] In light of these results, a skeptic could ask, "What is the point of forecasting using the DSGE models if they cannot improve upon simple ARs and naive forecasts (and nor can the Blue Chip, by the way)?" At least to us, the answer seems pretty obvious: try to perform policy analysis or to understand the forces driving the economy using an AR model if you can! We view forecasting chiefly as a test for DSGEs, rather than their main goal. We will elaborate on this point further in the our conclusion.

3.6. *Whole sample vs. post-great recession*

Thus far, and in much of the paper, the results have focused on forecasting during the recovery from the Great Recession, because this is the period of interest and the one for which we have forecasts from the NY Fed DSGE model.

38 The RMSEs obtained using the sample mean of inflation as a naive benchmark, which we do not report, are all above 0.5%, which is considerably worse than those of the DSGE model.

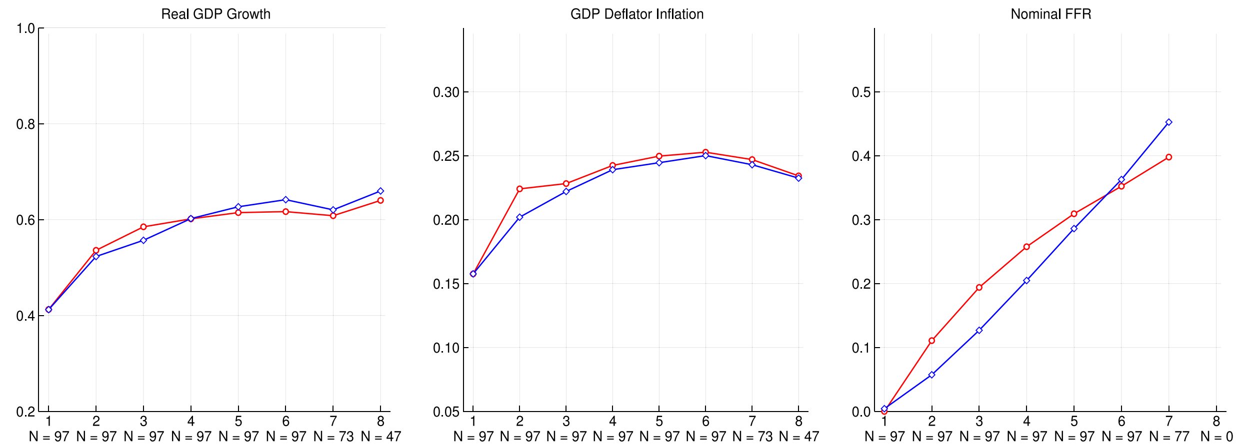

Fig. 12. RMSEs for SWFF vs. Blue Chip, computed from the whole sample (January 1992 to April 2016). *Notes*: The figure compares the RMSEs for SWFF (red circles) with those of the Blue Chip (blue diamonds) for one to eight quarters ahead for output growth, inflation, and interest rates. Output growth and inflation are expressed in Q/Q percentage terms, whereas interest rates are in quarterly percentage points. The $N = n$ label under each x-axis tick indicates the number of observations available for both the BCEI and DSGE forecasts at that horizon. The forecasts included in these calculations are from January 1991 to April 2016. The DSGE forecasts are conditional on the BCFF forecasts for the federal funds rate, and on the BCEI nowcasts for output growth and inflation. Section 3.6 provides the details of the forecast comparison exercise. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

This section turns to the question of how the DSGE models fared across our entire available sample of 1992–2017, for the sake of comparison with the previous literature on the accuracy of DSGE model forecasts for the U.S. As in the previous sections, we condition on time $T + 1$ BCEI forecasts of output and inflation. Interest rate expectations are incorporated starting in 2008Q4, to match the beginning of the ZLB period.

Fig. 12 shows that the SWFF model's performance is remarkably similar to that of the Blue Chip forecasts across all horizons and variables. As far as output and inflation are concerned, this finding is in line with that of Del Negro and Schorfheide (2013). Interest rate projections are moderately worse in the short- to medium-run, but overall are comparable in performance. This last point is notable given the lack of interest rate expectations from 1992–2008Q3, and indicates that the model is capable of producing reasonable interest rate forecasts away from the zero lower bound.

Edge and Gürkaynak's (2010) results showed that the accuracy of the DSGE models' forecasts is comparable to those of private forecasters. However, one could dismiss those findings on the grounds that they applied to the Great Moderation period, an easy period to forecast.[39] The results shown here are notable because they document that the accuracy of the DSGE models' forecasts is comparable to that of private forecasters, even though almost half of the sample includes periods that are particularly difficult for DSGE models, such as the Great Recession and its aftermath.

3.7. SWFF vs. its descendants

As was described in Section 2.1, the main models used in producing the various internal policy materials and forecasts were built on top of SWFF, chiefly by adding more observables (and more features to accommodate these observables).[40] This section asks to what extent these choices changed the DSGE's forecasting accuracy. Comparing the RMSEs from SWFF in Fig. 3 to the RMSEs shown in Fig. 13, we see that the near- and medium-term output growth forecast performances declined slightly from SWFF to SWFF$^+$ and from SWFF$^+$ to SWFF^{++}, whereas the long-term forecasting performance improved a bit at horizons seven and beyond, even outperforming the Blue Chip forecasts at that horizon. The near- and medium-term forecasts of inflation remained largely on a par between SWFF and its descendants, but in a similar fashion to the output growth forecasts, the long-term performance improved at horizons six and beyond.

4. Conclusions

The paper documents the accuracy of the projections of the NY Fed DSGE model during the recovery from the financial crisis. We find that our DSGE model's RMSEs are comparable to those obtained from the mean and median forecasts of the Blue Chip and SPF surveys, respectively, in the short and medium run (from one to eight quarters ahead). However, the NY Fed DSGE model performed much better than the median of the FOMC's Summary of Economic Projections in terms of the accuracy of its output growth forecasts, especially at longer horizons. For inflation, the DSGE performed worse than the median SEP up to

39 (Del Negro & Schorfheide, 2013) showed that this is still true if the sample is extended to 2011. Edge and Gürkaynak (2010) also find all of the forecast methods to be inaccurate in an R^2 sense, in that there were few forecastable fluctuations in the Great Moderation period.

40 The technical details of the additional features included in models SWFF$^+$ and SWFF^{++} are provided in Sections A.4 and A.5 of the appendix, respectively.

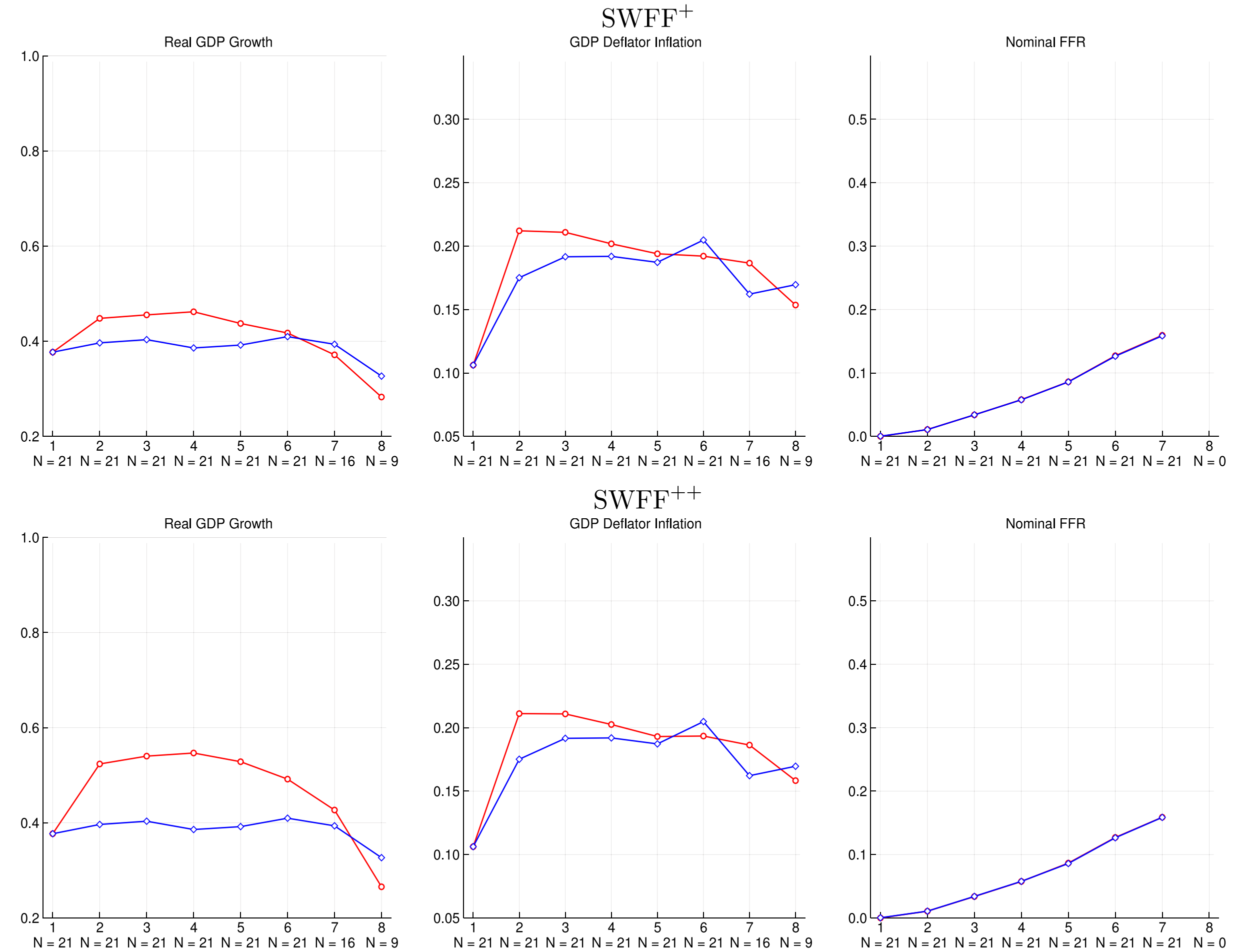

Fig. 13. RMSEs for SWFF$^+$ and SWFF^{++} vs. Blue Chip. *Notes*: The top and bottom panels compare the RMSEs for the SWFF$^+$ (top row, red circles) and SWFF^{++} (bottom row, red circles) DSGE models with those of the Blue Chip (blue diamonds) for one to eight quarters ahead for output growth, inflation, and interest rates. Output growth and inflation are expressed in Q/Q percentage terms, whereas interest rates are in quarterly percentage points. The $N = n$ label under each x-axis tick indicates the number of observations available for both the BCEI and DSGE forecasts at that horizon. The forecasts included in these calculations are from April 2011 to April 2016. The DSGE forecasts are conditional on the BCFF forecasts for the federal funds rate, and on the BCEI nowcasts for output growth and inflation. Section 3.7 provides the details of the forecast comparison exercise. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a two-year horizon, but better at a three-year horizon. The paper then uses a pseudo real-time forecasting exercise to assess which model features explain the results. It finds that financial frictions play a major role, especially in terms of the projections for economic activity, as they imply a slow recovery from financial crises.

The work of Otrok (2001), Schorfheide (2000) and Smets and Wouters (2003, 2007) more than ten years ago contained an implicit promise, namely that the macroeconomic profession could count on theory-based models that are flexible enough to fit the data, not just in sample but also *out of sample*. This paper shows that medium-scale DSGE models have kept some of their promises as far as out-of-sample forecasting accuracy is concerned. In order to do so, though, they had to change and incorporate financial frictions. Our prediction is that they will have to change again in the near future, both to keep up with the frontier of macroeconomics research (e.g., heterogeneous agents models as per Kaplan, Moll, & Violante, 2018, or non-linear models as per Brunnermeier & Sannikov, 2014) and to maintain and perhaps even improve their forecasting performances.

In closing, we should stress that we do not consider forecasting to be the primary objective of DSGE models, even though it is the focus of this paper. The out-of-sample forecasting accuracy is not important in itself, but only as an indirect test of model misspecification. DSGEs are used in many central banks for quantitative policy analysis. While good forecasting performance is no guarantee that the model's answers will be correct, one can at least say that bad forecasting performance is an indication that something is wrong with the model. In that case, its users should at the very least be aware of it.

Acknowledgments

We thank participants of the Central Bank Forecasting Conference for helpful comments. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Dallas, New York or the Federal Reserve System.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2018.12.001.

References

Adolfson, M., Andersson, M. K., Lindé, J., Villani, M., & Vredin, A. (2007). Modern forecasting models in action: improving macroeconomic analyses at central banks. *International Journal of Central Banking*, *3*(4), 111–144.

Alessi, L., Ghysels, E., Onorante, L., Peach, R., & Potter, S. (2014). Central bank macroeconomic forecasting during the global financial crisis: the european central bank and federal reserve bank of new york experiences. *Journal of Business & Economic Statistics*, *32*(4), 483–500.

Aruoba, S. B., Diebold, F. X., Nalewaik, J., Schorfheide, F., & Song, D. (2016). Improving gdp measurement: A measurement-error perspective. *Journal of Econometrics*, *191*(2), 384–397.

Atkeson, A., & Ohanian, L. E. (2001). Are phillips curves useful for forecasting inflation?. *Federal Reserve Bank of Minneapolis Quarterly Review*, *25*(1), 2–11.

Ball, L., & Mazumder, S. (2011). Inflation dynamics and the great recession. *Brookings Papers on Economic Activity*, *Spring*, 337–402.

Bernanke, B., Gertler, M., & Gilchrist, S. (1999). The financial accelerator in a quantitative business cycle framework. In J. B. Taylor, & M. Woodford (Eds.), *Handbook of macroeconomics: Vol. 1C*. North Holland, Amsterdam.

Boivin, J., & Giannoni, M. (2006). DSGE models in a data-rich environment. *Technical report*, National Bureau of Economic Research.

Brunnermeier, M. K., & Sannikov, Y. (2014). A macroeconomic model with a financial sector. *American Economic Review*, *104*(2), 379–421.

Campbell, J. R., Evans, C. L., Fisher, J. D., & Justiniano, A. (2012). Macroeconomic effects of fomc Forward Guidance. *Brookings Papers on Economic Activity*, *44*(1), 1–80.

Chauvet, M., & Potter, S. (2013). Forecasting output. In *Handbook of economic forecasting: Vol. 2* (pp. 141–194). Elsevier.

Christiano, L. J., Motto, R., & Rostagno, M. (2003). The great depression and the Friedman-Schwartz Hypothesis. *Journal of Money, Credit and Banking*, *35*, 1119–1197.

Christiano, L. J., Motto, R., & Rostagno, M. (2014). Risk shocks. *American Economic Review*, *104*(1), 27–65.

Christoffel, K., Coenen, G., & Warne, A. (2011). Forecasting with DSGE Models. In M. Clements, & D. Hendry (Eds.), *Handbook of Economic Forecasting* (pp. 89–127). Oxford University Press.

Coibion, O., & Gorodnichenko, Y. (2015). Is the phillips curve alive and well after all? Inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics*, *7*(1), 197–232.

Del Negro, M., Eusepi, S., Giannoni, M. P., Sbordone, A. M., Tambalotti, A., Cocci, M., Hasegawa, R., & Linder, M. H. (2013). The FRBNY DSGE model. *FRB of New York Staff Report No. 647*.

Del Negro, M., Giannoni, M. P., & Patterson, C. (2012). The forward guidance puzzle. *FRBNY Staff report*.

Del Negro, M., Giannoni, M. P., & Schorfheide, F. (2015). Inflation in the great recession and new keynesian Models. *American Economic Journal: Macroeconomics*, *7*(1), 168–196.

Del Negro, M., Hasegawa, R. B., & Schorfheide, F. (2016). Dynamic prediction pools: an investigation of financial frictions and forecasting performance. *Journal of Econometrics*, *192*(2), 391–405.

Del Negro, M., & Schorfheide, F. (2004). Priors from general equilibrium models for vars. *International Economic Review*, *45*(2), 643–673.

Del Negro, M., & Schorfheide, F. (2008). Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics*, *55*(7), 1191–1208.

Del Negro, M., & Schorfheide, F. (2013). DSGE Model-Based Forecasting. In G. Elliott, & A. Timmermann (Eds.), *Handbook of economic forecasting: Vol. 2*. Elsevier.

Del Negro, M., Schorfheide, F., Smets, F., & Wouters, R. (2007). On the fit of new keynesian Models. *Journal of Business & Economic Statistics*, *25*(2), 123–162.

Dotsey, M., Del Negro, M., Sbordone, A. M., & Sill, K. (2011). System dsge Project Documentation. *FOMC memo*, *June*.

Edge, R., & Gürkaynak, R. (2010). How useful are estimated dsge Model Forecasts for Central Bankers?. *Brookings Papers of Economic Activity*, *41*(2), 209–259.

Edge, R. M., Kiley, M. T., & Laforte, J. -P. (2010). A comparison of forecast performance between federal reserve staff forecasts, simple reduced-form models, and a dsge model. *Journal of Applied Econometrics*, *25*(4), 720–754.

Fair, R. (2018). Information content of dsge Forecasts. *Mimeo*, Yale University.

Faust, J., & Wright, J. H. (2013). Forecasting inflation. In *Handbook of economic forecasting: Vol. 2* (pp. 2–56). Elsevier.

Fawcett, N., Kórber, L., Masolo, R., & Waldron, M. (2015). Evaluating uk point and density forecasts from an estimated DSGE model: The role of off-model information over the financial crisis. *Staff Paper 538*, Bank of England.

Fernald, J. G. (2015). Productivity and potential output before, during, and after the great recession. *NBER Macroeconomics Annual*, *29*(1), 1–51.

Fernald, J. G., Hall, R. E., Stock, J. H., & Watson, M. W. (2017). The disappointing recovery of output after 2009. *Brookings Papers on Economic Activity*, *2017*(1), 1–58.

Gordon, R. J. (2015). Secular stagnation: a supply-side view. *American Economic Review*, *105*(5), 54–59.

Groen, J. J., Kapetanios, G., & Price, S. (2009). A real time evaluation of bank of england forecasts of inflation and growth. *International Journal of Forecasting*, *25*(1), 74–80.

Gürkaynak, R. S., Kısacıkoğlu, B., & Rossi, B. (2013). Do dsge Models Forecast More Accurately Out-Of-Sample than VAR Models? The views expressed in this article are those of the authors.. In *VAR models in macroeconomics — New developments and applications: essays in honor of Christopher A. Sims* (pp. 27–79). Emerald Group Publishing Limited.

Hall, R. E. (2011). The long slump. *American Economic Review*, *101*, 431–469.

Iversen, J., Laseen, S., Lundvall, H., & Söderström, U. (2016). Real-time forecasting for monetary policy analysis: the case of sveriges riksbank. *Discussion Paper 11203*, CEPR.

Justiniano, A., Primiceri, G., & Tambalotti, A. (2010). Investment shocks and business cycles. *Journal of Monetary Economics*, *57*(2), 132–145.

Justiniano, A., Primiceri, G. E., & Tambalotti, A. (2013). Is there a trade-off between inflation and output stabilization? *American Economic Journal: Macroeconomics*, *5*(2), 1–31.

Kaplan, G., Moll, B., & Violante, G. L. (2018). Monetary policy according to hank. *American Economic Review*, *108*(3), 697–743.

Kilponen, J., Orjasniemi, S., Ripatti, A., & Verona, F. (2016). The aino 2.0 model. *Research Discussion Paper 16*, Bank of Finland.

Kolasa, M., & Rubaszek, M. (2015). Forecasting with dsge Models with Financial Frictions. *International Journal of Forecasting*, *31*(1), 1–19.

Kolasa, M., Rubaszek, M., & Skrzypczyński, P. (2012). Putting the new keynesian dsge model to the real-time forecasting test. *Journal of Money, Credit and Banking*, *44*(7), 1301–1324.

Laseen, S., & Svensson, L. E. (2011). Anticipated alternative policy-rate paths in policy simulations. *International Journal of Central Banking*, *7*(3), 1–35.

Lees, K., Matheson, T., & Smith, C. (2011). Open economy forecasting with a dsge-var: Head to head with the RBNZ published forecasts. *International Journal of Forecasting*, *27*(2), 512–528.

Otrok, C. (2001). On measuring the welfare costs of business cycles. *Journal of Monetary Economics*, *45*(1), 61–92.

Reinhart, C. M., & Rogoff, K. S. (2009). *This time is different: eight centuries of financial folly*. Princeton University Press.

Romer, C. D., & Romer, D. H. (2000). Federal reserve information and the behavior of interest rates. *American Economic Review*, *90*(3), 429–457.

Romer, C. D., & Romer, D. H. (2008). The fomc versus the Staff: Where Can Monetary Policymakers Add Value?. *The American Economic Review*, *98*(2), 230.

Sbordone, A. M., Tambalotti, A., Rao, K., & Walsh, K. J. (2010). Policy analysis using dsge models: an introduction. *FRBNY Economic Policy Review*, *October*, 23–43.

Schorfheide, F. (2000). Loss function-based evaluation of dsge Model. *Journal of Applied Econometrics*, *15*(6), 645–670.

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

Smets, F., & Wouters, R. (2007). Shocks and frictions in us Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, *97*, (3), 586–606.

Stock, J. (2011). Discussion of ball and mazumder's "inflation Dynamics and the Great Recession". *Brookings Papers on Economic Activity*, *Spring*, 387–402.

Surico, P., Giannone, D., & D'Agostino, A. (2006). (Un)Predictability and macroeconomic stability. *Working Paper Series 605*, European Central Bank.

Tetlow, R. J., & Ironside, B. (2007). Real-Time model uncertainty in the united states: The Fed, 1996–2003. *Journal of Money, Credit and Banking*, *39*(7), 1533–1561.

Timmermann, A. (2006). Forecast combinations. In G. Elliott, C. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting: Vol. 1* (pp. 135–196). Elsevier.

Wieland, V., & Wolters, M. H. (2012). Forecasting and policy making. In G. Elliott, & A. Timmermann (Eds.), *Handbook of Economic Forecasting: Vol. 2*. Elsevier.

Michael Cai is a Research Analyst at the New York Fed.

Marco Del Negro is a Vice President in the Macroeconomics and Monetary Studies Function of the Research and Statistics Group. Mr. Del Negro's research focuses on the use of general equilibrium models in forecasting and policy analysis. Before joining the Bank, he was a research economist and associate policy adviser with the Macro group of the research department of the Federal Reserve Bank of Atlanta. Mr. Del Negro has published work in the American Economic Review, American Economic Journal: Macroeconomics, Journal of Econometrics, Journal of Applied Econometrics, Journal of International Economics, Journal of Monetary Economics, Journal of Money, Credit and Banking, International Economic Review, Journal of the European Economic Association, and the Review of Economic Studies.

Marc P. Giannoni is Senior Vice President and Director of Research. He joined the Dallas Fed in September 2017. Giannoni previously was research Economist and Assistant Vice President in the Macroeconomic and Monetary Studies Function of the New York Fed. He is a native of Switzerland and began his career as an economist with the Swiss National Bank in Zurich in 1992. He joined the New York Fed as an economist in 2000 before leaving to begin an academic career at the Columbia University Graduate School of Business in 2002. Giannoni rejoined the New York Fed in 2011 while continuing as an adjunct professor of finance and economics at Columbia. He holds BA and MA degrees in economics from the University of Geneva in Switzerland and MA and Ph.D. degrees in economics from Princeton University.

Abhi Gupta was a Research Analyst at the New York Fed and is now a graduate student at UC Berkeley.

Pearl Li was a Research Analyst at the New York Fed and is now a graduate student at Stanford University.

Erica Moszkowski was a Research Analyst at the New York Fed and is now a graduate student at Harvard Business School.