# Alternative Tests for Correct Specification of Conditional Predictive Densities

Barbara Rossi and Tatevik Sekhposyan
ICREA-UPF and Bank of Canada

George Washington University

February 13, 2014

*The views expressed here are those of the authors and not necessarily those of the Bank of Canada.*

# Motivation

- Recent economic events (great recession, unconventional monetary policy, fiscal cliff, etc.) sparked great interest in understanding uncertainty and its macroeconomic impact.

# Motivation

- Recent economic events (great recession, unconventional monetary policy, fiscal cliff, etc.) sparked great interest in understanding uncertainty and its macroeconomic impact.

- There has been increased emphasis in trying to correctly characterize uncertainty.

# Motivation

- Recent economic events (great recession, unconventional monetary policy, fiscal cliff, etc.) sparked great interest in understanding uncertainty and its macroeconomic impact.
- There has been increased emphasis in trying to correctly characterize uncertainty.
- Our paper contributes to the literature on evaluating the proper calibration of the predictive uncertainty.

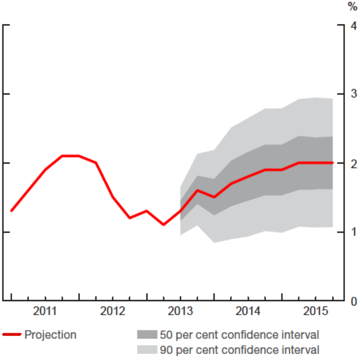
# Motivation

- Recent economic events (great recession, unconventional monetary policy, fiscal cliff, etc.) sparked great interest in understanding uncertainty and its macroeconomic impact.
- There has been increased emphasis in trying to correctly characterize uncertainty.
- Our paper contributes to the literature on evaluating the proper calibration of the predictive uncertainty.
- We also argue that the framework would be useful to describe what has become known as macroeconomic uncertainty or policy uncertainty.

# Goal

**Chart 33: Projection for core CPI inflation**

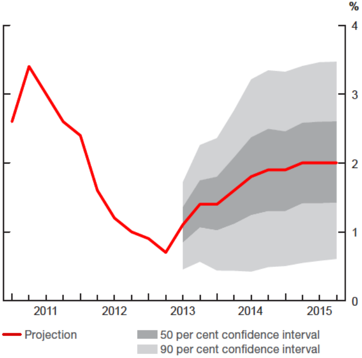
Year-over-year percentage change, quarterly data




Source: Bank of Canada

**Chart 34: Projection for total CPI inflation**

Year-over-year percentage change, quarterly data

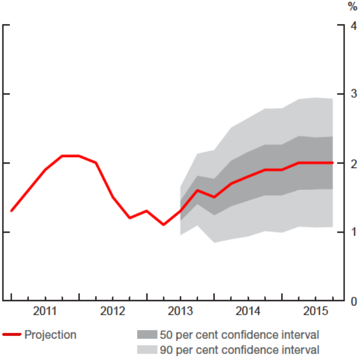



Source: Bank of Canada

# Goal

**Chart 33: Projection for core CPI inflation**

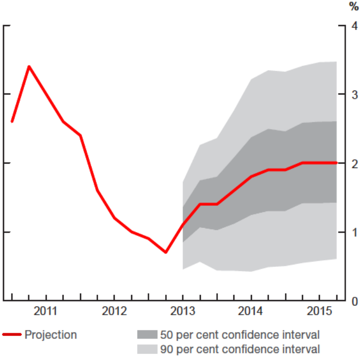
Year-over-year percentage change, quarterly data




Source: Bank of Canada

**Chart 34: Projection for total CPI inflation**

Year-over-year percentage change, quarterly data




Source: Bank of Canada

Test whether density forecasts are correctly specified

# Contributions of Our Paper

- Proposes alternative tests for correct specification of predictive densities
  - existing tests focus on null hypothesis in terms of model's population parameters
  - our proposed test aims at finite sample evaluation

# Contributions of Our Paper

- Proposes alternative tests for correct specification of predictive densities
  - existing tests focus on null hypothesis in terms of model's population parameters
  - our proposed test aims at finite sample evaluation
- Monte Carlo results show that our tests have good size and power in detecting misspecification in predictive densities

# Contributions of Our Paper

- Proposes alternative tests for correct specification of predictive densities
  - existing tests focus on null hypothesis in terms of model's population parameters
  - our proposed test aims at finite sample evaluation

- Monte Carlo results show that our tests have good size and power in detecting misspecification in predictive densities

- We apply our tests to evaluate the proper calibration of
  - mean probability forecasts of Survey of Professional Forecasters in real time
  - predictive densities of Smets and Wouters' (2007) DSGE model

# Outline

1. Contributions
2. Overview of the Literature
3. Our Proposed Tests
4. Monte Carlo Results
5. Empirical Applications
6. Conclusions

# Overview: Predictive Density Evaluation

1. Obtaining conditional predictive densities

- parametric: model + errors

  Classical: $y_{t+h} = bx_t + \epsilon_{t+h}$ and $\epsilon_t \sim iidN(0, \sigma^2)$, then conditional predictive density is: $\phi_{t+h}(.) \sim N(\hat{b}x_t, \hat{\sigma}^2)$

  Bayesian: $p_{t+h}(y_{t+h}|y_{1:t}) = p(y_{t+h}|\theta, y_{1:t})p(\theta|y_{1:t})d\theta$

- non-parametric
  - obtain predictive densities directly, e.g. SPF
  - bootstrap
  - historical MSFE

# Overview: Predictive Density Evaluation

1. Obtaining conditional predictive densities

- parametric: model + errors

  Classical: $y_{t+h} = bx_t + \epsilon_{t+h}$ and $\epsilon_t \sim iidN(0, \sigma^2)$, then conditional predictive density is: $\phi_{t+h}(.) \sim N(\hat{b}x_t, \hat{\sigma}^2)$

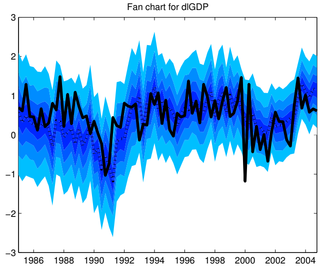
  Bayesian: $p_{t+h}(y_{t+h}|y_{1:t}) = p(y_{t+h}|\theta, y_{1:t})p(\theta|y_{1:t})d\theta$

- non-parametric
  - obtain predictive densities directly, e.g. SPF
  - bootstrap
  - historical MSFE

The sequence of $P$ out-of-sample conditional predictive densities is $(\{\phi_{t+h}(.)\}_{t=R}^{T})$

# Overview: Predictive Density Evaluation

1. Obtaining conditional predictive densities: example



# Overview: Predictive Density Evaluation

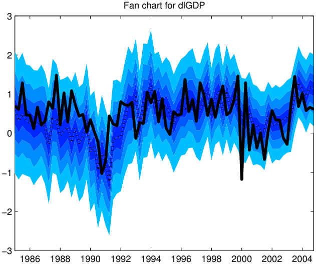
2. Evaluating conditional predictive densities

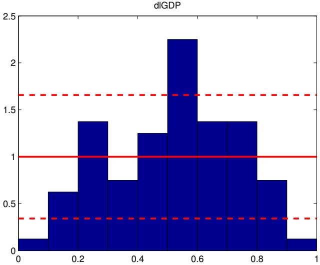
- compare them with the realized value
- Diebold, Gunther, and Tay (1998), Diebold, Tay, and Wallis (1999) propose the Probability Integral Transform (PIT)
- PIT = the cumulative density function of the density $\phi_{t+h}(.)$ evaluated at the realized value $y_{t+h}$

$$z_{t+h} = \int_{-\infty}^{y_{t+h}} \phi_{t+h}\left(u|\mathcal{F}_t\right) du \equiv \Phi_{t+h}\left(y_{t+h}|\mathcal{F}_t\right)$$

# Overview: Predictive Density Evaluation

2. Evaluating conditional predictive densities: example




$$z_{t+h} = \int_{-\infty}^{y_{t+h}} \phi_{t+h}(u|\mathcal{F}_t)\, du \equiv \Phi_{t+h}(y_{t+h}|\mathcal{F}_t)$$

# Overview: Predictive Density Evaluation

- Diebold, Gunther & Tay (1998); Diebold, Tay & Wallis (1999)

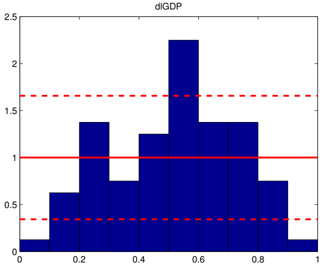
$$H_0 : \Phi_{t+h}\left(y_{t+h}|\mathcal{F}_t\right) = \Phi_0\left(y_{t+h}|\mathcal{F}_t\right)$$

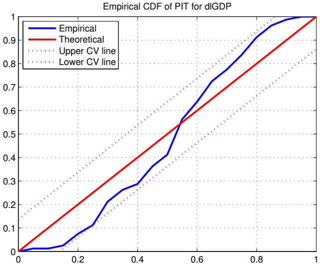
When $h = 1$, PITs are independently and identically distributed with a $U(0,1)$.

In case of multi-step-ahead forecasts, i.e. $h > 1$, PITs are distributed with $U(0,1)$.

# Overview: Predictive Density Evaluation

- Diebold, Gunther & Tay (1998); Diebold, Tay & Wallis (1999)



# Overview: Predictive Density Evaluation

Diebold, Gunther & Tay (1998); Diebold, Tay & Wallis (1999) suggest testing the various properties of the PITs

# Overview: Predictive Density Evaluation

Diebold, Gunther & Tay (1998); Diebold, Tay & Wallis (1999) suggest testing the various properties of the PITs

1. uniformity
   - plots of empirical cumulative distribution functions, histograms
   - Kolmogorov-Smirnov; Cramér-von Mises; etc. tests

# Overview: Predictive Density Evaluation

Diebold, Gunther & Tay (1998); Diebold, Tay & Wallis (1999) suggest testing the various properties of the PITs

1. uniformity
   - plots of empirical cumulative distribution functions, histograms
   - Kolmogorov-Smirnov; Cramér-von Mises; etc. tests

2. independence
   - autocorrelation functions
   - Box-Pierce; Ljung-Box; Brock, Dechert, Scheinkman tests

# Overview: Predictive Density Evaluation

Diebold, Gunther & Tay (1998); Diebold, Tay & Wallis (1999) suggest testing the various properties of the PITs

1. uniformity
   - plots of empirical cumulative distribution functions, histograms
   - Kolmogorov-Smirnov; Cramér-von Mises; etc. tests

2. independence
   - autocorrelation functions
   - Box-Pierce; Ljung-Box; Brock, Dechert, Scheinkman tests

Parameter estimation error due to sampling uncertainty and dynamic-misspecification due to ignoring part of the relevant information are not considered

# Overview: Predictive Density Evaluation

Corradi and Swanson (2006) takes both of these into account

- let $\mathcal{F}_t$ be the true information set
- let $\Im_t \subset \mathcal{F}_t$ be the information set used by the forecaster

$$H_0 : \Phi_{t+h}\left(y_{t+h}|\Im_t\right) = \Phi_0\left(y_{t+h}|\Im_t, \theta^{\dagger}\right)$$

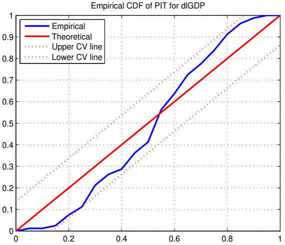
# Overview: Predictive Density Evaluation

Corradi and Swanson (2006) takes both of these into account

- let $\mathcal{F}_t$ be the true information set
- let $\Im_t \subset \mathcal{F}_t$ be the information set used by the forecaster

$$H_0 : \Phi_{t+h}\left(y_{t+h}|\Im_t\right) = \Phi_0\left(y_{t+h}|\Im_t, \theta^{\dagger}\right)$$

Consider




$$\Psi_P(r) = P^{-1/2}\sum_{t=R}^{R+P}\left(1\{z_{t+h} \leq r\} - r\right)$$

# Overview: Predictive Density Evaluation

Corradi and Swanson (2006) show that

- $\Psi_P(r)$ converges to the Gaussian process $\Psi^{\circ}(.,.)$, with mean zero and autocovariance function

$$E[\Psi(r_1)\Psi(r_2)] = \Omega(r_1, r_2)$$

- where $\Omega(r_1, r_2)$ has a complicated expression

# Overview: Predictive Density Evaluation

Corradi and Swanson (2006) show that

- $\Psi_P(r)$ converges to the Gaussian process $\Psi^\circ(.,.)$, with mean zero and autocovariance function

$$E\left[\Psi(r_1)\Psi(r_2)\right] = \Omega(r_1, r_2)$$

- where $\Omega(r_1, r_2)$ has a complicated expression

$$\begin{aligned}\Omega(r_1, r_2) \equiv\ & S_{\Phi\Phi}(r_1, r_2) \\ & + 2\Pi E\left[\nabla_\theta \Phi\left(\kappa(r_1)|\Im_t, \theta^\dagger\right)\right]' A\left(\theta^\dagger\right) S_{qq} A\left(\theta^\dagger\right) E\left[\nabla_\theta \Phi\left(\kappa(r_2)|\Im_t, \theta^\dagger\right)\right] \\ & - \Pi E\left[\nabla_\theta \Phi\left(\kappa(r_1)|\Im_t, \theta^\dagger\right)\right]' A\left(\theta^\dagger\right)\left(S_{q\Phi}(r_1)\right) \\ & - \Pi S'_{q\Phi}(r_2) A\left(\theta^\dagger\right) E\left[\nabla_\theta \Phi\left(\kappa(r_2)|\Im_t, \theta^\dagger\right)\right]\end{aligned}$$

# Overview: Predictive Density Evaluation

Complication arises from the fact that

$$
\begin{aligned}
\Psi_P(r) &\equiv \frac{1}{\sqrt{P}} \sum_{t=R}^{R+P} \left(1\left\{\Phi\left(y_{t+h}|\Im_t\right) \leq r\right\} - r\right) \\
&= \frac{1}{\sqrt{P}} \sum_{t=R}^{R+P} \left(1\left\{\Phi\left(y_{t+h}|\Im_t, \theta^{\dagger}\right) \leq r\right\} - r\right) \\
&- E\left[\nabla_\theta \Phi\left(\kappa(r)|\Im_{t-1}, \theta^{\dagger}\right)\right]' \, \frac{1}{\sqrt{P}} \sum_{t=R}^{R+P} \left(\widehat{\theta}_{t,R} - \theta^{\dagger}\right) + o_p(1)
\end{aligned}
$$

# Outline

1. Contributions
2. Overview of the Literature
3. Our Proposed Tests
4. Monte Carlo Results
5. Empirical Applications
6. Conclusions

# Proposed Tests

- let $\mathcal{F}_t$ be the true information set
- let $\Im_t \subset \mathcal{F}_t$ be the information set used by the forecaster
- let $\Im_{t-R+1}^{t}$ be info set based on time $(t-R+1)$ to time $t$

# Proposed Tests

- let $\mathcal{F}_t$ be the true information set
- let $\Im_t \subset \mathcal{F}_t$ be the information set used by the forecaster
- let $\Im_{t-R+1}^{t}$ be info set based on time $(t-R+1)$ to time $t$

$$H_0 : \Phi_{t+h}\left(y_{t+h}|\Im_{t-R+1}^{t}\right) = \Phi_0\left(y_{t+h}|\mathcal{F}_t\right)$$

# Proposed Tests

- let $\mathcal{F}_t$ be the true information set
- let $\Im_t \subset \mathcal{F}_t$ be the information set used by the forecaster
- let $\Im_{t-R+1}^t$ be info set based on time $(t - R + 1)$ to time $t$

$$H_0 : \Phi_{t+h} \left( y_{t+h} | \Im_{t-R+1}^t \right) = \Phi_0 \left( y_{t+h} | \mathcal{F}_t \right)$$

- We test whether the distribution is correctly specified given the forecasting model and its estimation technique.

# Proposed Tests

- let $\mathcal{F}_t$ be the true information set
- let $\Im_t \subset \mathcal{F}_t$ be the information set used by the forecaster
- let $\Im_{t-R+1}^{t}$ be info set based on time $(t-R+1)$ to time $t$

$$H_0 : \Phi_{t+h}\left(y_{t+h}|\Im_{t-R+1}^{t}\right) = \Phi_0\left(y_{t+h}|\mathcal{F}_t\right)$$

- We test whether the distribution is correctly specified given the forecasting model and its estimation technique.

Our test statistics also relies on:

$$\Psi_P(r) = P^{-1/2}\sum_{t=R}^{R+P}\left(1\{z_{t+h} \leq r\} - r\right)$$

# Proposed Tests

- let $\mathcal{F}_t$ be the true information set
- let $\Im_t \subset \mathcal{F}_t$ be the information set used by the forecaster
- let $\Im_{t-R+1}^{t}$ be info set based on time $(t - R + 1)$ to time $t$

$$H_0 : \Phi_{t+h}\left(y_{t+h}|\Im_{t-R+1}^{t}\right) = \Phi_0\left(y_{t+h}|\mathcal{F}_t\right)$$

- We test whether the distribution is correctly specified given the forecasting model and its estimation technique.

Our test statistics also relies on:

$$\Psi_P(r) = P^{-1/2}\sum_{t=R}^{R+P}\left(1\{z_{t+h} \leq r\} - r\right)$$

- similar to Corradi and Swanson (2006), but asymptotic distribution is different!

# More on the Null Hypothesis

*Example 1:* Let $y_t = \theta + \eta_t$, where $\eta_t \sim iidN(0,1)$.

The forecaster uses the correctly specified model, but estimates $\theta$ with $R = 2$: $\hat{\theta}_{t,2} = (y_t + y_{t-1})/2$; $\sigma_\eta$ assumed to be known for simplicity.

- $\phi_0 \left( y_{t+1} | \mathcal{F}_t \right) \sim N(\theta, 1)$
- $\phi_{t+1} \left( y_{t+1} | \Im_{t-2}^{t} \right) \sim N((y_t + y_{t-1})/2, 1)$

Then, our null hypothesis does not hold.

# More on the Null Hypothesis

*Example 2:* Conversely, let $y_{t+1} = \frac{1}{2}(y_t + y_{t-1}) + \eta_t$, where $\eta_t \sim iidN(0,1)$.

The forecaster uses a mis-specified model with a constant and $R = 2$, $\sigma_\eta$ assumed to be known for simplicity.

- $\phi_0 (y_{t+1}|\mathcal{F}_t) \sim N((y_t + y_{t-1})/2, 1)$
- $\phi_{t+1} \left(y_{t+1}|\Im_{t-2}^t\right) \sim N((y_t + y_{t-1})/2, 1)$

Then, our null hypothesis holds even though the estimated model is mis-specified.

# Proposed Tests

## Assumptions

i. $\{y_{t+h}, X_t'\}_{t=R}^{T}$ is mixing and $\{y_{t+h}\}_{t=R}^{T}$ has a cumulative distribution function $\Phi_0(.)$ that is continuous, differentiable and has a well defined inverse;

ii. $\{\Phi_{t+h}(y_{t+h}|\Im_t)\}_{t=R}^{T}$ has non-zero Jacobian with continuous partial derivatives; or

ii. $Pr(z_{t+h} \leq r_1, z_{t+h+d} \leq r_2) = F_d(r_1, r_2), Pr(z_{t+h} \leq r) = F(r)$, where $F_d(.,.)$ and $F(.)$ are distribution functions, $F(.)$ is continuous;

iii. $R < \infty$ as $P, T \to \infty$

(iii) implies that parameter estimation error does not vanish under the null hypothesis, as in Giacomini & White (2006) and Amisano & Giacomini (2007).

# Proposed Tests

- $h = 1$ and $\Im_{t-R}^{t} \subset \mathcal{F}_t$ such that $y_{t+1}|\Im_{t-R}^{t} = y_{t+1}|\mathcal{F}_t$
  Theorem 1: (i) $\{z_{t+h}\}_{t=R}^{T}$ is iid $U(0,1)$; (ii) $\Psi_P(r)$ weakly converges to a Gaussian process $\Psi(.)$, with mean zero and auto-covariance $E[\Psi(r_1)\Psi(r_2)] = [\inf(r_1, r_2) - r_1 r_2]$.

- $h > 1$ or $\Im_{t-R}^{t} \subset \mathcal{F}_t$ such that $y_{t+1}|\Im_{t-R}^{t} \neq y_{t+1}|\mathcal{F}_t$
  Theorem 2: (i) $\{z_{t+h}\}_{t=R}^{T}$ is $U(0,1)$ (ii) $\Psi_P(r)$ weakly converges to a Gaussian process $\Psi(.)$, with mean zero and auto-covariance $E[\Psi(r_1)\Psi(r_2)] = \sigma(r_1, r_2)$, where
  $\sigma(r_1, r_2) = \sum_{d=-\infty}^{\infty} [F_d(r_1, r_2) - F(r_1)F(r_2)]$.

# Tests of Correct Specification (I)

- Kolmogorov-Smirnov-type statistic (sup)
$\kappa_P^{CS} = \sup_{r \in [0,1]} \Psi_P(r)^2$

- Cramér-von Mises-type statistic (average)
$C_P^{CS} = \int_0^1 \Psi_P(r)^2 \, dr$

Under maintained assumptions and $H_0$:

$$\kappa_P^{CS} \Rightarrow \text{ functionals of Gaussian processes}$$

and

$$C_P^{CS} \Rightarrow \text{ functionals of Gaussian processes}$$

# Tests of Correct Specification (II)

The tests have non-standard, but usable asymptotic distributions.

Critical values for $h = 1$, Kolmogorov-Smirnov type test, $\kappa_{\alpha;P}^{CS}$

| | | 0.01 | 0.05 | 0.10 |
|---|---|---|---|---|
| Whole | $r \in [0, 1]$ | 2.25 | 1.51 | 1.19 |
| Right Tail | $r \in [0, 0.25]$ | 1.16 | 0.70 | 0.52 |
| Right Half | $r \in [0, 0.50]$ | 1.96 | 1.25 | 0.98 |
| Left Half | $r \in [0.50, 1]$ | 2.04 | 1.31 | 1.01 |
| Left Tail | $r \in [0.75, 1]$ | 1.34 | 0.82 | 0.61 |
| Center | $r \in [0.25, 0.75]$ | 2.21 | 1.48 | 1.16 |
| Tails | $r \in \{[0, 0.25] \cup [0.75, 1]\}$ | 1.45 | 0.95 | 0.74 |

$\kappa_{\alpha;P}^{CS}$ could be used to provide confidence intervals for the empirical cdf of the PIT.

# Recap for the Tests of Correct Specification

- Asymptotic distributions of KS and CM tests are nuisance parameter free under the conditions that guarantee independence of the PITS.
- When the null is rejected, it is not obvious whether it is due to violations of independence, uniformity or both.
- Multitude of tests of independence robust to violations of uniformity, ex. Ljung - Box, BDS, etc.

- We provide with a modified test of correct specification robust to violation of independence
  - However, for $h > 1$ horizon, the asymptotic distributions for the tests are not nuisance parameter free and depend on the serial correlation (HAC or bootstrap)
  - Other solutions - sample reduction, Bonferroni methods, etc.

# Other tests in the same framework

- Inverse Normal Tests of Berkowitz (2001): Let $\Phi^{-1}(.)$ denote the inverse of the standard normal distribution. Under the maintained assumptions and the null $\Phi^{-1}(z_{t+1})$ is iid $N(0,1)$.

- Could also be extended to the Continuous Rank Probability Scores of Gneiting and Raftery (2007). However, in this case the asymptotic distributions are not nuisance parameter free under the considered assumptions.

# Outline

1. Contributions
2. Overview of the Literature
3. Our Proposed Tests
4. Monte Carlo Results
5. Empirical Applications
6. Conclusions

# Monte Carlo Results: Size (I)

- DGP S1: The data are generated according to

$$\widetilde{y}_{t+1} = R^{-1} \sum_{j=t-R+1}^{t} y_j + \sigma_t \eta_{t+1}$$

- $y_t = \mu + \varepsilon_t$, $\mu = 5$, and $\varepsilon_t, \eta_{t+1} \sim iid\ N(0,1)$
- $\sigma_t^2 = R^{-1} \sum_{j=t-R+1}^{t} (y_j - R^{-1} \sum_{j=t-R+1}^{t} y_j)^2$
- Estimate a model with a constant
- Consider $R = [50, 100, 200]$ and $P = [100, 200, 500, 1000]$ to evaluate the performance in finite samples

# Monte Carlo Results: Size (I)

| | DGP S1 | | | | | |
|---|---|---|---|---|---|---|
| | $\kappa_P^{CS}$ | | | $C_P^{CS}$ | | |
| $P$ $R:$ | 50 | 100 | 200 | 50 | 100 | 200 |
| 100 | 0.05 | 0.06 | 0.06 | 0.05 | 0.06 | 0.05 |
| 200 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| 500 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| 1000 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |

Nominal size is 5%. The number of Monte Carlo replications is 5,000.

# Monte Carlo Results: Size (II)

- DGP S3: $\widetilde{y}_{t+1} = R^{-1} \sum_{j=t-R+1}^{t} y_j + u_{t+1} + \rho u_t$
- $y_t = \mu + \varepsilon_t + \rho \varepsilon_{t-1}$, $\rho = 0.2$, $u_t, \varepsilon_{t+1} \sim iid\ N\left(0, \sigma^2\right)$
- $\sigma_t^2 = R^{-1} \sum_{j=t-R+1}^{t} (y_j - R^{-1} \sum_{j=t-R+1}^{t} y_j)^2$
- Estimate a model with a constant and a MA(1) term

# Monte Carlo Results: Size (II)

| DGP S4 - Asymptotic Critical Values with HAC Estimates | | | | | | |
|---|---|---|---|---|---|---|
| | $\kappa_P^{CS}$ | | | $C_P^{CS}$ | | |
| $P$ $R$ : | 50 | 100 | 200 | 50 | 100 | 200 |
| 100 | 0.07 | 0.12 | 0.11 | 0.04 | 0.07 | 0.06 |
| 200 | 0.09 | 0.13 | 0.09 | 0.05 | 0.07 | 0.05 |
| 500 | 0.09 | 0.11 | 0.08 | 0.05 | 0.06 | 0.05 |
| 1000 | 0.09 | 0.10 | 0.10 | 0.05 | 0.05 | 0.06 |

Nominal size is 5%. The number of Monte Carlo replications is 5,000.

# Monte Carlo Results: Power

- Power against constant mis-specification over time
- DGP P1. The data are generated from a chi-squared distribution, and the researcher tests whether it is normal:

$$\widetilde{y}_t = R^{-1} \sum_{j=t-R+1}^{t} y_j + (1-c)\,\eta_{1,t} + c\left(\eta_{2,t}^2 - 1\right)\sqrt{2},$$

where $y_j = \mu + \varepsilon_j, \mu = 1$, and $\varepsilon_j, \eta_{1,t}, \eta_{2,t}$ are iid $N(0,1)$.

# Monte Carlo Results: Power

| DGP P1 | | |
|---|---|---|
| $c$ | $\kappa_P^{CS}$ | $C_P^{CS}$ |
| 0 | 0.05 | 0.05 |
| 0.10 | 0.35 | 0.40 |
| 0.15 | 0.80 | 0.91 |
| 0.20 | 0.99 | 1.00 |
| 0.25 | 1.00 | 1.00 |

$R = 40, P = 960$. The number of Monte Carlo replications is 5,000.

# Outline

1. Contributions
2. Overview of the Literature
3. Our Proposed Tests
4. Monte Carlo Results
5. Empirical Applications
6. Conclusions

# Survey of Professional Forecasters

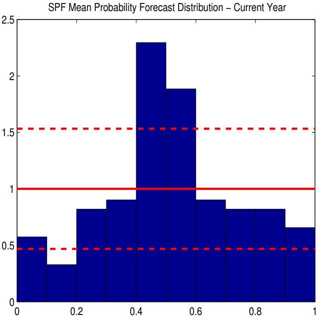
- SPF contains survey forecasts made by professional forecasters
- Focus on density forecasts of real GNP/GDP growth and inflation measured by the GNP/GDP deflator
- SPF reports: (1) the average probability distribution forecasts of the respondents, (2) the cross sectional distribution of the point forecasts across the forecasters
- Evaluated against the first quarterly vintage of Real Time Dataset for Macroeconomists
- Data from 1981:III - 2011:IV, provided by the Philadelphia Fed

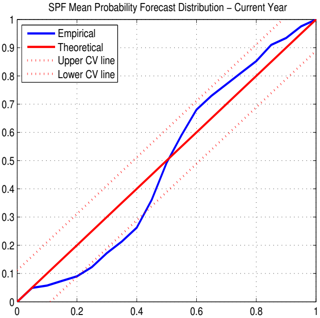
# SPF's Mean Probability Forecast Distribution

| | GDP Growth | | GDP Deflator Growth | |
|---|---|---|---|---|
| | $\kappa_P^{CS}$ | $C_P^{CS}$ | $\kappa_P^{CS}$ | $C_P^{CS}$ |
| h = 0 | 2.10*† | 0.81*† | 15.71*† | 5.12*† |
| h = 1 | 0.59† | 0.12† | 23.40*† | 10.16*† |

# SPF Mean Probability Forecast Distribution

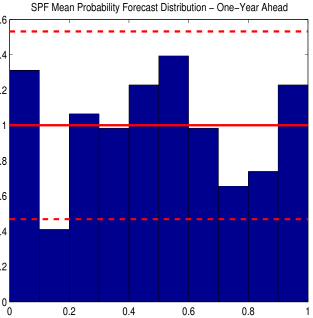
## GDP Growth (1981:III-2011:IV)

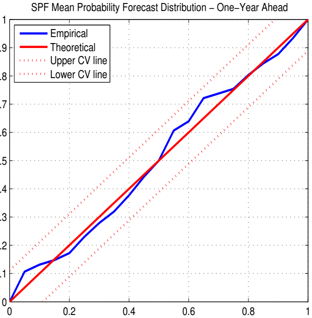


# SPF Mean Probability Forecast Distribution

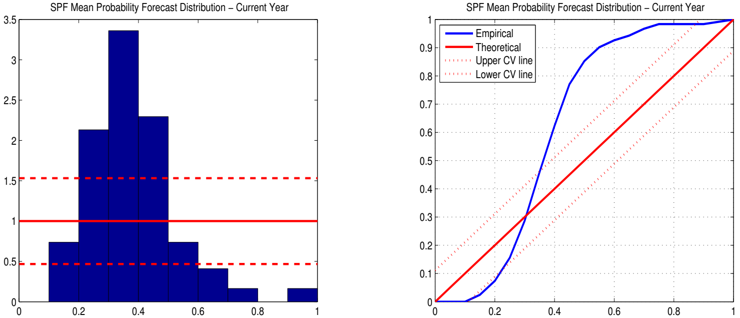
GDP Growth (1981:III-2011:IV)



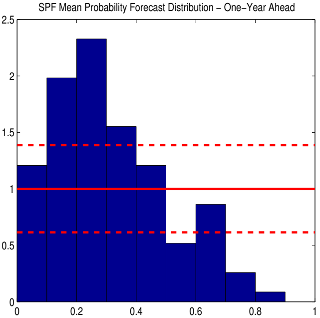
# SPF Mean Probability Forecast Distribution

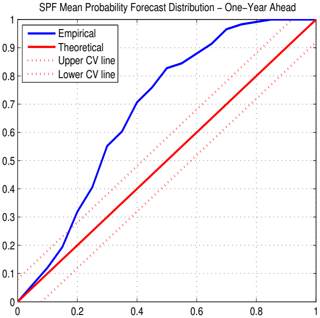
## GDP Deflator Growth (1981:III-2011:IV)



# SPF Mean Probability Forecast Distribution

GDP Deflator Growth (1981:III-2011:IV)



# Smets and Wouters (2007) Forecast Distributions

- 7 variable linearized DSGE model with nominal and real frictions
- log difference of real GDP, GDP deflator, real consumption, real investment and real wage, log hours worked, federal funds rate
- widely used as an attractive benchmark for forecasting and policy analysis
- sample period is 1966:I - 2004:IV, forecast horizon $h = 1$
- estimated with a fixed rolling window of size $R = 80$

# Smets and Wouters (2007) Forecast Distributions

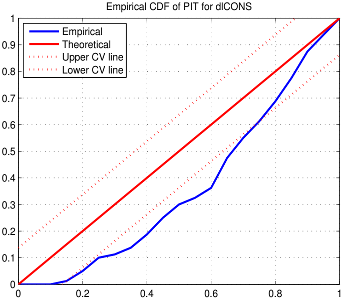
| Variable | $\kappa_P^{CS}$ | $C_P^{CS}$ |
|---|---|---|
| Consumption | 4.42 *† | 1.84 *† |
| Investment | 0.34 † | 0.16 *† |
| Output Growth | 1.46 † | 0.66 *† |
| Inflation | 1.20 † | 0.45 † |
| Hours | 0.65 † | 0.31 † |
| Wages | 1.06 † | 0.27 † |
| Fed Funds Rate | 3.12 *† | 1.37 *† |

There is more evidence of correct calibration for the forecast distributions of inflation, hours and wages.

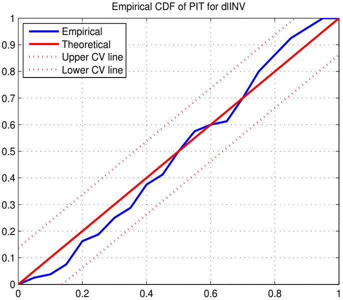
Correct calibration fails for all variables based on tests robust to violations of independence under the null.

# Smets and Wouters (2007) Forecast Distributions

## Consumption

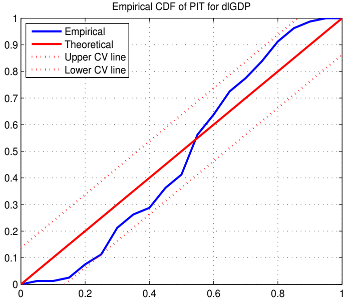



## Investment

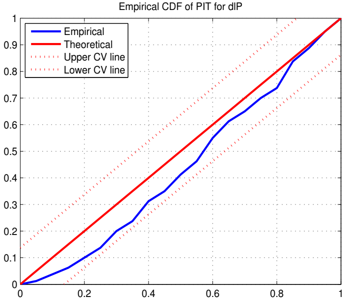


# Smets and Wouters (2007) Forecast Distributions

GDP

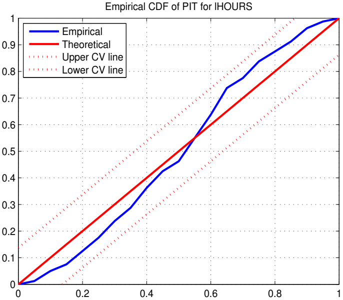



Inflation

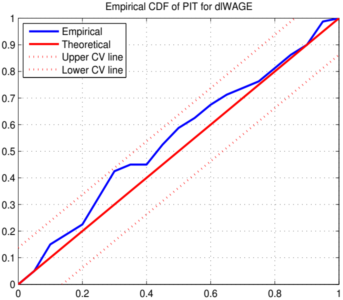


# Smets and Wouters (2007) Forecast Distributions

## Hours

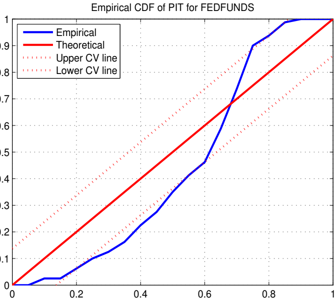



## Wages



# Smets and Wouters (2007) Forecast Distributions

## Fed Funds Rate



# Conclusion

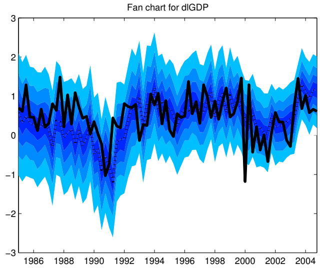
- This paper proposes new methodologies for evaluating density forecasts.
- These methodologies allow researchers to reach empirical conclusions given finite samples.
- Monte Carlo results suggest good size and power for the proposed methodologies.
- When applied to the SPF's density forecasts, our tests detect mis-specification for both output growth and inflation.
- When applied to a benchmark DSGE model of Smets and Wouters (2007), we find mis-specification in all densities.

# Some New Work: Macroeconomic Uncertainty

- Baker, Bloom and Davis (2011) index - economic policy uncertainty index.
- Coenen and Warne (2013) - trace the uncertainty or risk measure given certain quantiles of predictive densities.
- Jurado, Ludvigson and Ng (2013) index - construct it based on MSFE-s, given factor models with stochastic volatility. Obtain an average measure across various series.
- Scotti (2013), Jain, Jo and Sekkel (2013) indexes - still based on MSFE-s, though based on survey forecasts.
- We propose considering PIT bases measure of uncertainty.
- Ultimately, we would need to use it for forecasting as in Bijsterbosch and Guérin (2013).

# Macroeconomic Uncertainty

1. Obtain the predictive density




2. PIT: $U_{t+h} = \int_{-\infty}^{y_{t+h}} \phi_{t+h}\left(y \middle| \Im_t\right) dy$
3. Consider measures of uncertainty

$$\begin{aligned} U_{t+h}^{+} &= \frac{1}{2} + \max\left\{U_{t+h} - \frac{1}{2}, 0\right\} \\ U_{t+h}^{-} &= \frac{1}{2} + \max\left\{\frac{1}{2} - U_{t+h}, 0\right\} \end{aligned}$$

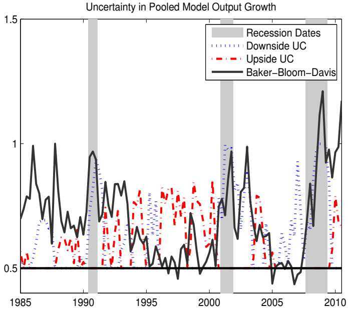
# Macroeconomic Uncertainty

Advantages of our measure

- focuses on uncertainty relative to the predicted outcome
- is a distribution-based measure of uncertainty: it distinguishes between periods of high and low uncertainty measured by probabilities
- is a more general way to measure uncertainty relative to MSFE
- allows for asymmetry, i.e. "positive" or "upside" uncertainty is different than a "negative" or "downside" uncertainty
- could construct a measure of uncertainty for a specific variable, a group of variables, all variables

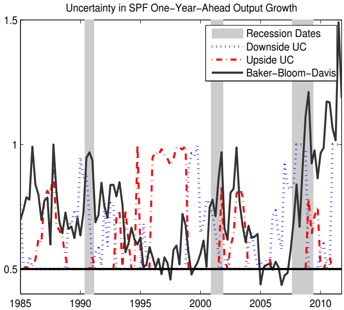
# Macroeconomic Uncertainty

1. from pooled ADL models (with a simple average) of real economic activity, asset prices, wages, prices and money



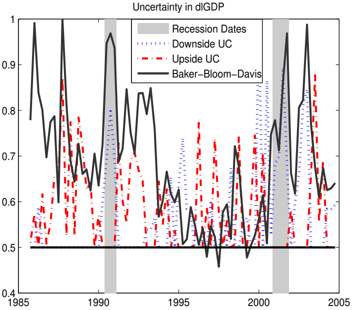
# Macroeconomic Uncertainty

2. from SPF output growth density forecast



# Macroeconomic Uncertainty

3. from DSGE fan charts



# Macroeconomic Uncertainty: Conclusions

- Model-based uncertainty measures seem to track the Baker, Bloom and Davis (2011) index better than SPF.
- All measures point to increased uncertainty during the Great Recession period.
- In general, recession periods could be characterized by "downside" uncertainty.
- There seems to be considerable "upside" uncertainty in late 1990-s, not captured by the Baker, Bloom and Davis (2011) index.