**Readme File for Diebold, Shin and Zhang (2022)**

“On the Aggregation of Probability Assessments:

Regularized Mixtures of Predictive Densities for

Eurozone Inflation and Real Interest Rates”

These codes are frozen in June 2022 and replicate the results of the paper dated June 6, 2022. All codes have been tested using MATLAB 2022a and 2020a.

**Data construction**

The following five files perform data cleaning. Turning the raw excel survey data into the balanced panel. All output files are saved in ~/replication/data. ***This stage requires users running the MATLAB scripts on a Windows system with Microsoft Excel installed and setting the current directory to ~/replication/code.***

**Filename: main\_data\_step01.m**

* This file reads raw survey data from the folder: ~/replication/data/SPF\_individual\_forecasts
* Creates for data cells with following fields

1. Sdate: Survey date (month or quarter)
2. Tdate: Starting date (month or quarter) of forecasting target
3. Id: id of forecasters
4. Point: point prediction
5. Hist: Histogram value of the probability forecasts (row: id, col: probs)
6. Id\_point: ids of point forecasts.
7. Id\_hist: ids of probability forecasts.

* There are four forecasts (ECB refers to these forecasts as “rolling horizon” forecasts, see below)
  + 1-year-ahead inflation (ecbspf\_infl\_1y)
  + 2-year-ahead inflation (ecbspf\_infl\_2y)
  + 1-year-ahead rgdp (ecbspf\_rgdp\_1y)
  + 2-year-ahead rgdp (ecbspf\_rgdp\_2y)
* There are six additional forecasts for the calendar year 1, 2 and 5, which are not the focus of this paper.
* Output: **data\_ecb\_spf\_2019Q4.mat**

**Filename: main\_data\_setp02.m**

* This file adds “actual” values to the constructed dataset (fieldname = actual)
* Actual = “data from most recent vintage” (final)
* Taken from: <https://sdw.ecb.europa.eu/home.do?chart=t1.3>
* Output: **data\_ecb\_spf\_2019Q4\_v02.mat**

**Filename: main\_data\_setp03.m**

* Goal: Produce balanced panel by the extrapolation
* Output: **ecbspf\_infl\_1y\_bp.mat**

**What we did (see details in footnote 17)**

* The filter is such that forecasters with more than four consecutive missing observations are excluded from the panel, as per Genre et al. (2013).
* T=1, we fill in the gaps with average non-missing forecasts
* T>1, we replace with the similar forecasters (ordered by rps, by (Ngr)-groups, based on the previous survey performance). (Ngr=5)

**Filename: main\_data\_step04.m**

* This file deals with two issues
* (1) Equalizing histogram bins over time (by aggregation)
* (2) Add 1% to the zero-prob bins with actual (by distributing probs to others)
* We do this for only inflation (1y)
* Output: **ecbspf\_infl\_1y\_bp\_nozero.mat**

**Filename: main\_data\_step05.m**

* This file deals with nominal interest rate data
* Raw data from FRED: 12-Month London Interbank Offered Rate (LIBOR), based on Euro, Percent, Daily, Not Seasonally Adjusted. Stored in EUR12MD156N.xls
* Output: **data\_nominal\_rate.mat**

**Target variables**

Inflation: The year-on-year percentage change of the Harmonized index of consumer prices (HICP)

**Survey dates** are collected here: SPF\_rounds\_dates.xlsx

**Detailed description of the ECB-SPF data:** <https://www.ecb.europa.eu/stats/prices/indic/forecast/shared/files/SPF_dataset_description.pdf>

**Descriptive analysis**

**Filename:** main\_desc.m

* This file generates the figure 2 that shows individual and average density forecasts at 2004Q4 and 2018Q4.

**Simulation analysis**

* Relevant files are: main\_simul\_dgp1.m, main\_simul\_dgp2.m, main\_simul\_dgp1\_kf.m, main\_simul\_dgp2\_kf.m
* Files without kf is based on incompletely-rational predictive densities. They generate figure 1 and table 1 in the main body.
* Files with kf explore fully-rational forecasts. They generate figure B1 and table B1 in the appendix.
* All simulation results are saved in the subfolder *simul\_results*
* The results presented in the paper are based on different random seeds than those specified in the replication files. As a result, the generated graphs are not precisely identical to those shown in the paper, but qualitatively similar.

**Main empirical analysis**

**Filename:** main\_emp\_fixW.m

* This file implements all proposed methods and generates the tables that are presented in the empirical section of the paper. The original latex tables are stored in the subfolder *latex*.

The following files generate figures for the empirical analysis and store them in the subfolder *graphics*.

* **Filename:** main\_figure\_3d\_inf.m
  + 3D heatmaps for inflation rate (fig3 (top), fig 4, fig 5)
* **Filename:** main\_figure\_3d\_inf.m
  + 3D heatmaps for real interest rate (fig 8)
* **Filename:** main\_figure\_2d\_diff\_inf.m
  + 2D differenced heatmaps for inflation rate (fig 3, fig 6)
* **Filename:** main\_figure\_2d\_diff\_real\_int.m
  + 2D differenced heatmaps for real interest rate (fig 9)
* **Filename:** main\_pit.m
  + PIT analysis (fig 7)