On the Aggregation of Probability Assessments: Regularized Mixtures of Predictive Densities

Francis X. Diebold
University of Pennsylvania

Minchul Shin
Federal Reserve Bank of Philadelphia

July 10, 2020

Abstract: ***

Acknowledgments: For helpful comments we are grateful to Brendan Beare, Graham Elliott, Rob Engle, Domenico Giannone, Christian Hansen, Nour Meddahi, Mike McCracken, Marcelo Medeiros, Mike West, and Ken Wolpin. We are also grateful to participants at NBER Summer Institutes and NBER/NSF Time-Series Conferences. We are especially grateful to Umut Akovali, who made early contributions as a Penn visiting student several years ago. Philippe Goulet Coulombe provided outstanding research assistance. The usual disclaimer applies.

Key words: Density forecasts, forecast combination, survey forecasts, shrinkage, model selection, regularization, partially egalitarian LASSO, model averaging, susbset averaging, trim and average

JEL codes: C2, C5, C8

Contact: fdiebold@upenn.edu, minchul.shin@phil.frb.org

Extended Abstract:

We work in the linear opinion rule framework of Amisano and Geweke (2017), developing regularization methods for density forecast combination. Brodie et al. (2009) recognized that the linear opinion rule must be paired with a simplex constraint for the resulting combined density to be proper. Imposition of the simplex constraint is effectively a particular L^1 (LASSO) regularization, so there is no room for additional tuning of L^1 regularization strength.

We progress by considering two frameworks for predictive density combination, which maintain the L^1 simplex regularization but augment it with additional regularization. One keeps all forecasters in the pool, and the other keeps only a selected subset.

1. *Keeping all forecasters in the pool*:

   (a) First, we consider adding an L^2 (Ridge) shrinkage penalty to the L^1 simplex shrinkage penalty. This turns out to be quite similar to elastic net (***ref), and we explore centering the additional L^2 penalty on 1/K, where K is the number of forecasters, so that we shrink toward equal weights.

   (b) Second, we consider an entropy-based shrinkage penalty that automatically imposes the L^1 simplex constraint but also introduces additional regularization. We show that a Bayesian analysis with a Dirichlet prior, which puts positive probability only on the simplex so that the fundamental L^1 constraint is automatically imposed, and which has mean 1/K so that it shrinks toward equal weights (with a scalar hyperparameter that controls shrinkage strength), produces posterior mixture weights that match those obtained from the entropy-based shrinkage penalty.

2. *Keeping only a selected subset of forecasters in the pool*:

   In addition to keeping all forecasters in the pool but increasing regularization via extra penalty terms centered on $1/K$, we also consider selecting $k<K$ forecasters and then shrinking toward $1/k$. This is the density forecast analog of the partially-egalitarian LASSO and related trimmed-averaging procedures of Diebold and Shin (2019).

We assess approaches 1a, 1b, and 2 relative to each other and relative to the basic simplex-constrained linear opinion rule. We do this with a blend of theory and Monte Carlo, as well as with an empirical application that focuses on survey density forecasts of the growth and inflation factors central to stock and bond pricing.