

DISS. ETH NO. 27509

Essays on Economic Forecasting with Machine Learning

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

Daniel Stefan Ehmman-Wochner

Master of Science (M.Sc.) in Computer Science,
University of Oxford
Master of Arts (M.A.) in Economics and Business Administration,
University of Zurich

born on 05.01.1988

citizen of Eggersriet-Grub SG, Switzerland

accepted on the recommendation of

Prof. Dr. Jan-Egbert Sturm, examiner
Prof. Marcelle Chauvet, Ph.D., co-examiner

2021

Summary of the Thesis¹

This cumulative dissertation includes four essays on economic time series forecasting. It broadly examines how we may meaningfully adopt and adapt machine learning algorithms in data-rich environments for more accurate predictions of (macro-)economic dynamics. All four chapters derive from a consistent data basis: The Federal Reserve Economic Data (FRED) provided by the Federal Reserve Bank of St. Louis.

The first chapter, co-authored with Boriss Siliverstovs, is mainly concerned with the evaluation of macroeconomic forecasts for a wide range of macroeconomic indicators in case of possible instabilities.² This study systematically broadens the relevance of possible model performance asymmetries across business cycles in the spirit of the recent state-dependent forecast evaluation literature (e.g. Chauvet & Potter, 2013) to hundreds of macroeconomic indicators and deepens the forecast evaluation of the recent factor model literature on hundreds of target variables (e.g. Stock & Watson, 2012b) in a state-dependent manner. Our results are consistent with both strands of the literature and generalize the former to over 200 macroeconomic indicators and differentiate the latter across three levels of temporal granularity: We document systematic model performance differences in both absolute and relative terms across business cycles (longitudinal) as well as across variable groups (cross-sectional) and find these performance differences to be robust across several alternative specifications. The cross-sectional prevalence and robustness of state dependency shown in this article encourages economic forecasters to complement model performance assessments with a *state-dependent evaluation of predictive ability*.

The second chapter focuses on predictive macroeconomic modeling with ensembles of trees. While the recent macroeconomic forecasting literature shows that double targeted forests (DTF) can enhance the predictive ability of single targeted forests (STF) (e.g. Borup, Christensen,

¹The summaries provided herein are based on the abstracts of the corresponding essays.

²The remainder of this paragraph corresponds to the abstract in Siliverstovs and Wochner (2021).

Mühlbach, & Nielsen, 2020; Medeiros, Vasconcelos, Veiga, & Zilberman, 2021; Wochner, 2018), I study the merits of triple targeted forests (TTF) as a flexible and rich framework for non-linear tree-based ensemble predictions in macroeconomic forecasting when dealing with high-dimensional and mixed-frequency environments. TTFs offer interesting and novel modeling choices across the different steps of the framework, especially in combination with recent advances put forth in the literature (e.g. novel/traditional factors, supervised/unsupervised factor selection). An extensive out-of-sample forecasting experiment for short-term U.S. output growth shows that optimally tuned and suitably designed TTFs can offer considerable gains in predictive accuracy and may outperform well-established distributed lag and autoregressive models as well as single and double targeted forests. Nevertheless, factor-augmented linear models remain a competitive benchmark.

In the third chapter, I elaborate on forest-based model estimation in the presence of possible state-dependent dynamics and thereby build upon the first two chapters. Specifically, I propose so-called Dynamic Factor Forests (DFF) for macroeconomic forecasting, which extend the model-based *trees* of Zeileis, Hothorn, and Hornik (2008) in the spirit of Garge, Bobashev, and Eggleston (2013) to model-based *forests* and synthesize the recent machine learning, business cycle and dynamic factor model literature within a unified statistical machine learning framework. DFFs are state-dependent, non-linear and smoothed forecasting models that allow us to embed theory-led factor models in powerful tree-based machine learning ensembles conditional on the business cycle state. DFFs incorporate in the spirit of the recent literature four key features: (i) Complex conditioning sets allow to account for generalized and potentially time-varying state-dependent dynamics while (ii) forest-based randomization decorrelates the dynamic factor trees in the ensemble. (iii) The inclusion of Lasso-based state-dependent regularization not only serves as a guard against overfitting but also enables switching mechanisms between structurally distinct state-dependent models. (iv) In addition to traditional principal component-based factors, novel factor transformations are examined. An extensive out-of-sample forecasting experiment for short-term U.S. GDP growth provides promising results in favor of DFFs.

In the fourth chapter, I combine recent advances in the literature for more accurate density predictions of U.S. inflation. Recent work by

Medeiros et al. (2021) shows that point forecasts of Breiman's (2001a) random forest machine learning algorithm can systematically outperform well-established benchmarks at predicting U.S. inflation. This chapter extends their work from point to density forecasts: On the one hand, the predictive densities of well-established random walk, autoregressive and stochastic volatility benchmarks are examined. On the other, three recent forest-based machine learning models are assessed: Namely, Quantile Regression Forests (Meinshausen, 2006), Bayesian Additive Regression Trees (H. A. Chipman, George, & McCulloch, 2010) and Distributional Forests (Schlosser, Hothorn, Stauffer, & Zeileis, 2019). The chapter shows that the most recent Distributional Forest (DIFO) algorithm often achieves the best point and density predictions. DIFOs allow to simultaneously account for locally evolving means, nonlinearities, time-varying volatilities as well as heavy-tailed densities — all of which constitute stylized characteristics of inflation dynamics highlighted in the recent literature, which is shown to be key for their predictive superiority.

Zusammenfassung der Dissertation³

Die vorliegende kumulative Dissertation umfasst vier Aufsätze zu ökonomischen Zeitreihenprognosen. Sie untersucht im Wesentlichen inwiefern Methoden des maschinellen Lernens (“Machine Learning”) im Kontext grosser ökonomischer Datensätze sinnvoll verwendet und adaptiert werden können, um (makro-)ökonomische Entwicklungen genauer zu prognostizieren. Alle vier Kapitel basieren auf einer konsistenten Datengrundlage: Der “Federal Reserve Economic Data (FRED)” der U.S.-Notenbank von St. Louis.

Das erste Kapitel wurde in Zusammenarbeit mit Boriss Siliverstovs erstellt und fokussiert primär auf die Evaluation makroökonomischer Prognosen für eine grosse Anzahl an Indikatoren im Falle möglicher Instabilitäten.⁴ Die Studie erweitert systematisch die Relevanz möglicher Leistungsasymmetrien von Modellen in unterschiedlichen Konjunkturzyklen im Sinne der zyklusabhängigen Evaluationsliteratur (bspw. Chauvet & Potter, 2013) auf mehrere Hundert makroökonomische Zeitreihen und vertieft die Evaluation von Faktormodellen für mehrere Hundert Zielvariablen (bspw. Stock & Watson, 2012b) in einer situationsbedingten bzw. zyklusabhängigen Art und Weise. Unsere Resultate sind konsistent mit beiden Literaturströmungen und generalisieren die Erkenntnisse der Ersteren auf über 200 makroökonomische Indikatoren und vertiefen jene der Letzteren über drei verschiedene temporale Granularitätsstufen: Wir dokumentieren systematische Leistungsunterschiede zwischen Modellen in absoluten sowie relativen Grössen über verschiedene Konjunkturzyklen (Längsschnitt) sowie Variablengruppen (Querschnitt) hinweg und finden robuste Leistungsdifferenzen für zahlreiche Spezifikationen. Die aufgezeigte weite Verbreitung und Robustheit zyklusabhängiger Leistungen,

³Die Zusammenfassungen hierin basieren auf den Kurzfassungen der entsprechenden Arbeiten.

⁴Der Rest dieses Paragraphen entspricht der übersetzten Kurzfassung in Siliverstovs und Wochner (2021).

ermutigten Wirtschaftsprognostiker deren Modell-Leistungsüberprüfungen um *zyklusabhängige Evaluationen* zu ergänzen.

Im zweiten Kapitel, konzentriere ich mich auf die Modellierung makroökonomischer Prognosemodelle mit Verbänden von Entscheidungsbäumen (“ensembles of trees”). Die jüngste makroökonomische Prognoseliteratur zeigt, dass zweifach ausgerichtete Wälder (“Double Targeted Forests”) die Prognosefähigkeit des einfach ausgerichteten Äquivalents (“Single Targeted Forests”) verbessern können (bspw. Borup et al., 2020; Medeiros et al., 2021; Wochner, 2018). Das Kapitel evaluiert die Vorzüge von dreifach ausgerichteten Wäldern (“Triple Targeted Forests”) als flexiblen und ergiebigen Ansatz aus nicht-linearen und Entscheidungsbaum-basierten Verbundprognosen, der sich insbesondere für makroökonomische Prognosen in hochdimensionalen Umgebungen mit gemischten Zeit-Frequenzen eignet. Dreifach ausgerichtete Wälder offerieren interessante und neue Modellierungsmöglichkeiten über die verschiedenen Schritte des Ansatzes hinweg, insbesondere in Kombination mit jüngsten Beiträgen in der Literatur (bspw. traditionelle/neuartige Faktoren, beaufsichtigte/unbeaufsichtigte Faktorbestimmung). In einem umfangreichen (out-of-sample) Vorhersageexperiment für kurzfristige Prognosen des amerikanischen Wirtschaftswachstums wird gezeigt, dass optimal dreifach ausgerichtete Wälder nennenswerte Gewinne in der Prognosegenauigkeit erzielen und etablierte autoregressive Prozesse sowie Modelle mit verteilten Verzögerungen (“distributed lags”) als auch einfach und zweifach ausgerichtete Wälder übertreffen können. Allerdings bleiben lineare Faktormodelle eine kompetitive Modellklasse.

Im dritten Kapitel behandle ich die Spezifizierung makroökonomischer Modelle, um möglichen zyklusbedingten Dynamiken in Wäldern Rechnung zu tragen und baue in diesem Sinne auf den Erkenntnissen der ersten beiden Kapitel auf. Ich schlage sogenannte dynamische Faktorstämme (DFW) (“Dynamic Factor Forests”) für makroökonomische Prognosen vor, welche die Modell-basierten Bäume von Zeileis et al. (2008) im Sinne von Garge et al. (2013) zu Modell-basierten Wäldern erweitern und es erlauben, die drei Literaturströmungen zu maschinellem Lernen, Konjunkturzyklen und Faktormodellen in einem einheitlichen statistischen Machine Learning Modell zusammenzuführen. DFWs sind nicht-lineare, geglättete und zyklusbedingte Prognosemodelle, die es uns erlauben theoriegeleitete Faktormodelle in leistungsstarken Verbänden Entscheidungsbaum-basierter Machine Learning Modelle in Abhängigkeit

des Konjunkturzyklus zu integrieren. Darüber hinaus integrieren DFWs gestützt auf neuere Literatur vier wesentliche Merkmale: (i) Erlauben komplexere Spezifizierungen des Konditionierungssets generalisierte und zeitvariierende, zyklusbedingte Dynamiken. (ii) Ausserdem werden durch Wälder-basierte Randomisierungen die dynamischen Faktorbäume des Verbundes entkorriert. (iii) Die Einbindung zyklusspezifischer Lasso-basierter Modell-Regularisierung reduziert nicht nur das Risiko einer Modellüberanpassung (“overfitting”), sondern gewährt DFWs ausserdem die Möglichkeit von Regimewechseln zwischen strukturell unterschiedlichen Modellen. (iv) Ausserdem werden neben traditionellen Faktoren, neuartige Faktormethoden in DFWs studiert. Ein umfangreiches (out-of-sample) Prognoseexperiment für kurzfristige Entwicklungen des amerikanischen Wirtschaftswachstums attestiert DFWs gute Prognosefähigkeiten.

Im vierten Kapitel kombiniere ich kürzlich erzielte Fortschritte in der Literatur für akkuratere Vorhersagen der amerikanischen Inflationsrate. Jüngste Studien von Medeiros et al. (2021) zeigen, dass die Punktprognosen anhand von Breiman’s (2001a) randomisierten Wäldern (“Random Forests”) systematisch etablierte Benchmark-Modelle übertreffen können. Das vierte Kapitel erweitert deren Arbeit von Punkt- zu Dichteprognosen: Einerseits werden die Dichteprognosen von etablierten Prozessen mit zufälligen Bewegungen (“Random Walk”), autoregressiven Prozessen sowie Prozessen mit stochastischer Volatilität evaluiert. Andererseits werden folgende drei neuen Wald-basierten Machine Learning Methoden untersucht: “Quantile Regression Forests” (Meinshausen, 2006), “Bayesian Additive Regression Trees” (H. A. Chipman et al., 2010) sowie “Distributional Forest” (Schlosser, Hothorn, Stauffer, & Zeileis, 2019). Ich zeige, dass Baum-basierte Algorithmen im Allgemeinen und die neuen Distributional Forests (DIFOs) im Besonderen sehr häufig die besten Punkt- und Dichteprognosen erzielen, denn DIFOs erlauben es gleich mehreren stilisierten Fakten von Inflationsreihen Rechnung zu tragen, die in der Literatur hervorgehoben werden, wie beispielsweise Nichtlinearitäten, lokal evolvierenden Durchschnitten, zeitvariierenden Volatilitäten sowie Verteilungen mit extremen Ereignissen.